Who's Afraid of *Quantum Causation*?

Causal Explanations and Bell Phenomena

MSc Thesis *(Afstudeerscriptie)*

written by

Reinoud Pino

under the supervision of **Prof Dr Sonja Smets**, and submitted to the Examinations Board in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam*.

Date of the public defense:	**Members of the Thesis Committee:**
06/02/2024	Dr Balder ten Cate
	Dr Dingmar van Eck
	Dr Sebastian De Haro Ollé
	Prof Dr Sonja Smets



Institute for Logic, Language and Computation

Abstract

Causal theories of scientific explanation contend that science explains phenomena by giving their causes, or at least that it should. Yet, in quantum mechanics there are phenomena that are not readily framed causally, such as Bell phenomena. There is both a multitude of causal theories of scientific explanation (interventionism, conserved quantities theory of causation(CQTC)) and interpretations of quantum mechanics (Copenhagen, Transactional, Pilot Wave, Many Worlds). After considering Bell phenomena as described in the Copenhagen interpretation from the perspective of both Interventionism and CQTC, it is argued that neither theory considers said description an explanation, but it is also argued that CQTC is most likely to fit with other interpretations. After considering the Transactional Interpretation, Pilot Wave Mechanics and the Many Worlds Interpretation, I argue that the descriptions offered by both the Transactional and Many Worlds Interpretations are qualified as explanations by CQTC.

Contents

1 Scientific Explanation **7**

1.1 The Theory 8

1.2 Insufficiency 9

1.3 Unification as Epistemic Justification 13

2 Causal Explanation **16**

2.1 Interventionism 17

2.1.1 The Theory 17

2.2 Conserved Quantity Theory of Causation 20

2.2.1 Causal Processes 21

2.2.2 The Connection Between a Cause and its Effect 23

2.2.3 Distinguishing Cause and Effect 25

2.3 Causation: concept and empiricity 27

3 Explaining Bell Phenomena in the Copenhagen Interpretation **29**

3.1 Quantum Mechanics 29

3.1.1 The Predictive Recipe 30

3.1.2 Wave-particle Duality 32

3.1.3 The Measurement Problem 36

3.1.4 The Uncertainty Principle 39

3.1.5 Bell Inequalities 40

3.2 Common Causes and Bell phenomena 45
3.3 Copenhagen Interpretation and Interventionism 48
3.3.1 Impossible Interventions 48
3.4 The Copenhagen Interpretation and CQTC 54
3.4.1 Causal Processes 55
3.4.2 Objects 56
3.4.3 World Lines and the Uncertainty Principle 57
3.4.4 Quantum Conserved Quantities 59
3.5 Conclusion 64

4 CQTC and Alternative Quantum Interpretations 66

4.1 Retro-causal Interpretations 67
4.1.1 The Transactional Interpretation 68
4.1.2 Explaining the Transactional Interpretation of Bell Phenomena . . . 70
4.2 Pilot-Wave Theory 75
4.2.1 Pilot-Wave explained 75
4.2.2 PWT and 1/2 spin Bell phenomena 78
4.2.3 PWT and CQTC 79
4.3 The Many Worlds Interpretation 86
4.3.1 MWI and Bell Phenomena 88
4.3.2 MWI and CQTC 91

Introduction

What do we want from science? That is a fundamental and difficult question to answer, but at least part of the answer seems to be "explanations". We experience certain phenomena and we want to understand why these phenomena occur, and turn to science to give us the answer. We see the sea rise and fall, and we wonder why. However, if we ask of science that it gives us explanations, we must have some idea of what scientific explanation actually is. Enter the large and diverse philosophical literature on exactly that question. One subsection of that literature argues that scientific explanation is causal explanation. If I have a phenomenon A and I can point towards phenomena B and C and say "these cause A", then I have given an explanation of A. You ask me: "Why does the sea have tides?" I tell you that the moon and sun exert gravitational force on the earth's waters, causing great waves across our seas and oceans. When the crest of such a wave reaches the shore, it is high tide. When a trough reaches the shore, it is low tide. And thus I have causally explained the tides.

Not everything, however, is so easily cast into a causal framework. There are a number of phenomena in quantum physics that stubbornly resist being understood causally. Any theory of scientific explanation has to explain why the stories we tell about those phenomena are still explanations, or why it is no problem that they are not.

At the same time, fundamental quantum physics is a highly contentious field. We have a highly accurate method for making predictions, but we are unsettled about what the mathematics we use to make those predictions actually means. Different views on what the mathematics mean (and how we might supplement/change it) tend to be called interpreta-

tions, although according to some that is a somewhat denigrating term. I am tempted to agree, but will call them interpretations nonetheless for convention's sake. In any case, if we accept that what we want from science is explanations, then it follows that it is a desiderata on the descriptions that these interpretations offer that there is at least *some* sense in which they are explanations.

And there we have the object of this thesis: to find a match between theories of explanation and interpretations of quantum mechanics. If there is a theory of explanation X according to which the descriptions offered by quantum interpretation Y count as explanation, then that is good news for both X and Y.

Before discussing the structure of this thesis, some preliminaries are in order. First, quantum mechanics is ubiquitous in modern physics, so the list of phenomena is enormous. This thesis focuses on one group of phenomena, and one specific instance in that group. I call that group "Bell phenomena", and include in it all phenomena which highlight quantum physics' seemingly non-local nature. So that includes Bell's experiment figuring three particles, but also Einstein, Podolsky and Rosen's experiment involving two particles. The specific phenomenon I will discuss is that of two entangled anti-correlated 1/2 spin particles, because it is conceptually about as simple as these phenomena get. Second, quantum mechanics is not the only problem of causal theories of explanation, and a lack of a sense of explanation is not the only issue plaguing the interpretations I will discuss, but it is this tension that will be the focus of this thesis.[1]

In the first chapter, I motivate focusing on causal theories of explanation for quantum mechanics, by arguing that the predominant alternative, unificationism, can not be the correct theory of scientific explanation. Unification turns out to be an insufficient condition for explanation and I argue that it should be understood as an epistemic norm on a body of explanations instead.

The second chapter introduces two causal theories of explanation: interventionism and the conserved quantities theory of causality.

[1]Meaning that, at times, I will gloss over issues that do not directly relate to this specific tension.

The third chapter introduces quantum mechanics and some central interpretive issues, in order to give the tools to describe Bell phenomena involving two entangled anti-correlated 1/2 spin particles according to different interpretation, and discuss these descriptions in light of different theories of explanation. Furthermore, the third chapter explains how to understand Bell inequalities from a causal perspective, arguing it rules out common causes for Bell phenomena (at least in the standard interpretation). Finally, the third chapter evaluates both theories of causal explanation in light of the Copenhagen interpretation of Bell phenomena, concluding that the Copenhagen description does not count as explanation according to either theory, but that the conserved quantities theory of causality seems most likely to succeed with other interpretations.

The fourth chapter explores that conclusion. I describe Bell phenomena according to three alternative interpretations: the transactional interpretation, pilot wave mechanics, and the many worlds interpretation. I argue that only the retro-causal interpretation and the many worlds interpretation are compatible with the conserved quantities theory of causality, primarily because they have no wavefunction collapse and avoid direct forms of non-locality, in the sense that non-locality emerges from other aspects of the theories (retro-causality and the splitting of worlds respectively).

Chapter 1

Scientific Explanation

What is scientific explanation? For the first half of the twentieth century and going into the second, the received view was that scientific explanation was something along the lines of Hempel's deductive nomological modal.[1] Briefly put, Hempel says a phenomenon, the explanandum, is explained when there is a set of sentences which forms a deductive argument for the explanandum and at least one of the premises of that argument is a natural law.[2] The model is, however, subject to counterexamples which made it less popular as the twentieth century went on. For a famous example, using the deductive nomological model we can explain why a flagpole casts a shadow, but we can also argue that the shadow is an explanation of the flagpole, which seems a counter-intuitive result. This problem, that this theory of explanation cannot factor in the direction of explanation, is called explanatory asymmetry.[3]

As the second half of the 20th century progressed, causal theories of scientific explanation started to enjoy a predominant place in the literature. Twenty years ago, there was again a certain level of agreement on what scientific explanation was: "Scientific explanation is a matter of providing suitable information about causes of the explanandum phenomenon."[4] So, causal theories of scientific explanation typically argue that to give an

[1] Salmon 2006
[2] Hempel and Oppenheim 1948
[3] Woodward 2005, pp. 358–363
[4] Reutlinger 2017, p. 2

explanation of an explandum, is to place said explandum into a causal structure. Naturally, there is much room for discussion in this view. To start, any theory of causal explanations must have a theory of causation, which in itself is a notoriously difficult question. Furthermore, there has to be a sense in which this concept of causation connects to explanation. That is to say, there is much contention as to *which* causal account is best.[5] In light of the problems causal theories have with quantum mechanics, it is tempting to ask why we bother with them at all. After all, there is an alternative that enjoys a comparable level of popularity and does qualify quantum mechanics as explanatory: unificationism.[6] The aim of this chapter is to show that unificationism can not be the correct theory of scientific explanation, thus motivating the effort to find a causal theory of explanation that is compatible with quantum mechanics.

1.1 The Theory

So, what is unificationism? Unificationism claims that scientific explanation is a matter of offering a unification of a range of phenomena. An explanation, in this theory, is a derivation that can be used for a range of phenomena and, importantly, is part of a set of explanations called the *explanatory store*.[7] In order to say what the explanatory store is, we need to look at a specific version of unificationism.

That version will be Kitcher's, which is arguably the most influential.[8] Let us take a closer look at his formulation. Call the total set of statements endorsed by some specific scientific community K. Furthermore, an *argument pattern* is a triple consisting of a *schematic argument*, *filling instructions* and a *classification*. A *schematic argument* is a sequence of *schematic sentences*, which are sentences in which some non-logical vocabulary are replaced by dummy letters. Such a sentence comes with substitution rules, or *filling instructions*, which dictate how which dummy letters can be replaced by actual terms. A *classification*

[5]Woodward and Ross 2021
[6]Kitcher 1989, pp. 430–431
[7]Kitcher 1989, pp. 430–431
[8]Woodward and Ross 2021

describes which sentences in an argument are premises and which conclusions, and which rules of inference are used where.[9]

Then, call the set of derivations that best systematizes K the explanatory store $E(K)$, where *best systematizes* is supposed to mean best unifies.[10] The set of derivations can be said to have a *generating set*, which is the set of argument patterns such that all derivations in the explantory store are instantations of some argument pattern in the generating set.[11] To be more unifying is to do more with less: if we can get more conclusions from fewer patterns in our generating set, our set of derivations is more unifying.[12] Another metric by which Kitcher measures unification is the stringency of the patterns: "If a pattern sets conditions on instantiations that are more difficult to satisfy than those set by another pattern, then I shall say that the former pattern is more stringent than the latter."[13] The explanatory store is then the set of explanations that best unifies K, and a derivation is an acceptable explanation of some $d \in K$ in virtue of being a member of the explanatory store.

The idea that explanation is strongly connected to unification is quite intuitive, so if we are to argue that we cannot say what explanation is in terms of unification, the question would remain what that relation does consist of. Moreover, how do we show that, although there is a relation between the two, the one is not equal to the other?

I will argue that unification is not the same as explanation, but rather is an epistemic justification for accepting a set of explanations.

1.2 Insufficiency

I will give two arguments against unification's insufficiency for explanation. The first, posed by to Victor Gijsbers, is that it follows from unificationism that for any $E(K)$ we can construct a set of derivations which is equally unifying, but consists entirely of statements

[9]Kitcher 1989, p. 432
[10]Kitcher 1989, p. 431
[11]Kitcher 1989, p. 434
[12]Kitcher 1989, p. 435
[13]Kitcher 1989, p. 433

explaining themselves.[14] The second, taken from James Woodward, is that unificationism does not solve the problem of explanatory asymmetry and thus claims derivations to be explanatory which people generally do not take to be explanations.[15]

The consequence of unificationism that any fact or statement explains itself is called 'spurious unification'.[16] Let's say we have a scientific statement α. Then we can form the argument pattern:

$$\frac{\alpha}{\alpha} \tag{1.1}$$

Prima facie, the set containing only $\frac{\alpha}{\alpha}$ is going to be the most unifying $E(K)$, because it contains the smallest amount of patterns whilst still containing a pattern, namely one, yet gives us maximal conclusions since we can substitute any scientific statement for α, but the α-set is not stringent at all, thus actually not unifying at all.[17]

However, we can construct self-explaining patterns that are nonetheless still stringent.[18] Whenever we have an argument pattern A, we can construct A' such that

- A' has the form of the α-pattern
- α can only be substituted for p such that $p \in C(A)$

$C(X)$ refers to the set of conclusions that is generated by some argument pattern X. Then $C(A) = C(A')$ and the substitution rules for A' mean that they allow for the same amount of substitutions as A, thus A' is as stringent as A. For any set of argument patterns AP, we can take each pattern in the set and apply this procedure, so that we can get a set of patterns AP' which consists of only self explaining patterns, yet is just as unifying and thus just as explanatory as AP.[19]

[14] Gijsbers 2007, p. 483
[15] Woodward 2005, pp. 358–363
[16] Gijsbers 2007, p. 483
[17] Gijsbers 2007, p. 483
[18] Kitcher 1981, p. 527
[19] Gijsbers 2007, p. 484

Kitcher is aware of constructions like this, and he means to block spurious unification by placing a constraint on substitution rules:[20]

> If the filling instructions associated with a pattern P could be replaced by different instructions, allowing for the substitution of a class of expressions of the same syntactic category, to yield a pattern P' and if P' would allow the derivation of *any* sentence, then the unification achieved by P is spurious.

However, this requirement excludes patterns we want and includes patterns we do not. Take for instance this pattern:

$$\frac{\alpha \quad \alpha \to \beta}{\beta} \tag{1.2}$$

And the substitution rules: any sentence for α and β such that $\alpha \to \beta$ is acceptable. Kitcher's restriction implies this pattern is not explanatory, since if we change the substitution rules, e.g. remove the restriction that α and β must make $\alpha \to \beta$ acceptable, then this pattern allows us to conclude anything.[21] So the pattern is excluded by Kitcher's constraint, but it is a pattern we would arguably like to keep, thus the restriction seems too strong.

To see that, simultaneously, the constraint is not strong enough, consider which patterns the restriction excludes/includes. According to Gijsbers, the excluded class is only those patterns which have a 'dummy letter' as their conclusion and nothing more, i.e the conclusion is something to be wholly substituted according to the substitution rules.[22] For if the conclusion is *more* than just a dummy letter, the pattern could never be used to conclude *every* sentence, since it must always include that part which is not substituted.[23][24]

So, let's take a pattern A which has a conclusion that includes terms that are not dummies. For instance, the conclusion might be "α is the beginning". We can then construct

[20] Kitcher 1981, pp. 527–528
[21] Gijsbers 2007, p. 485
[22] Gijsbers 2007, p. 484
[23] Gijsbers 2007, p. 484
[24] It is not immediately obvious that this means all patterns whose conclusion is a dummy letter must be excluded, but what *is* certainly true is that all patterns whose conclusions contain more than a dummy letter are not excluded.

A' such that

- A' is of the form $\frac{\alpha \textit{ is the beginning}}{\alpha \textit{ is the beginning}}$
- α is only substituted by sentences such that "α is the beginning" $\in C(A)$

A' is not excluded by the constraint, contains only self-explanations, and is as unifying as A, so the constraint is too weak to avoid the kind of set of argument patterns that motivated it in the first place.[25]

The constraint thus does not work and self-explanations remain problematic for unificationism. Which is to say that unification is an insufficient condition for explanation, because there are derivations that unify, but do not explain.

This is not the only way in which unification is insufficient. The second argument is made by Woodward in *Making Things Happen.*[26]

Kitcher claims to solve the issue of explanatory asymmetries. To explain how he uses the paradigmatic example of the flagpole and the shadow. Say we call our current practices in which we can not explain the dimensions of the pole by way of the shadow P, and compare this with an alternative P' where shadows *do* explain those dimensions. According to Kitcher, P includes a so-called "origin and development" (OD) pattern, which says that the dimensions of our object are traced back to how the object came to be (originated) and how it has changed since (developed).[27] In P' we have to add pattern S, which explains the dimensions of an object from its shadow. That means, however, that P' has more patterns, without deriving more conclusions, so P' is less unifying than P thus S is not an explanation. Alternatively, say we replace OD with S in P instead of adding S to it, and call this P''. Since objects do not always have a shadow, but would still have dimensions, we can derive fewer conclusions from P'' although it has the same amount of patterns as P, thus P'' is less unifying than P, and thus S is still no explanation.[28] It then seems Kitcher solves explanatory asymmetries.

[25]Gijsbers 2007, p. 486
[26]Woodward 2005
[27]Kitcher 1989, p. 485
[28]Kitcher 1989, p. 487

Kitcher's argument crucially depends on the fact there are fewer backwards derivations than forward derivations.[29] So, say we look at a scenario in which there are as many derivations going forward as backward. Take Newtonian mechanics applied to the solar system. We can derive the state of motion of the planets at time t_1 from information about their positions, masses, velocities and the forces working on them at some prior time t_0. However, the reversed derivation is equally possible.[30] Call the derivation of future states from past information forwards derivations, and the reverse backwards derivations. If we are comparing an $E(K)$ and $E'(K)$ where one includes the forwards-pattern and another the backwards-pattern, then one is not more unifying than the other, thus in terms of explanation they are equal for unificationism. However, if we were to ask "Why is planet x at location y at t_0?" and someone were to tell us "Well, in the future at t_1 the planets are in state of motion z, so we can deduce that planet x is at y at t_0," they are answering a different question. They are telling us how we might know the status of planet x, which can coincide with an explanation of planet x' status but is not the same.

So, unification is not a sufficient condition for explanation, because it does not solve explanatory asymmetry and thus there are derivations that unify, yet do not explain.

1.3 Unification as Epistemic Justification

The relation between unification and explanation seems apparent, so if unification is not explanation, what is it? I argue that unification is an epistemic justification for adopting one set of explanations over another. In order to make sense of this argument against unificationism we need to define what epistemic justification is and, more specifically, what it might be for science in specific.

Epistemic justification is the justification we use for adopting beliefs. Exactly what kind of justification is permissible when we are considering what to believe is subject to dis-

[29] Woodward 2005, p. 360
[30] Woodward 2005, p. 362

cussion.[31] In general, we can differentiate between two approaches to epistemic justification: deontological and non-deontological.[32] The argument in favour of the non-deontological approach is that many of our beliefs are not active choices, but rather conditions we are in.[33] Whether or not this is correct and whether it would necessarily undermine the deontological account can be left aside because what seems clear is that at least within science, the degree to which epistemic justification is an active process is much higher than in the case of day-to-day beliefs, which allows us to proceed deontologically.

When we focus on deontological epistemic justification, we can again see that we can split theories into two kinds: those that say that we need to believe what is supported by evidence, and those that say our beliefs need to be pointed towards a set of epistemic norms.[34] That leaves the questions what, exactly, should count as evidence, and what said epistemic norms ought to be. Forsaking these questions for the moment, I want to point towards a reconciliatory approach to the two positions. The straight-forward end-aim of beliefs seems to be the truth, so saying that direct evidence, for instance experience, pointing towards the truth or falsity of some belief ought to influence said belief seems natural. However, it is not uncommon for beliefs to be underdetermined in terms of evidence, in which case it makes sense that we refer to other epistemic norms or values to supply our beliefs with justification.

What I have just outlined is some variant of what Kuhn argue happens in science: our theories are underdetermined by empirical evidence, thus we use other epistemic values, such as simplicity, accuracy and scope, to determine which theories we adopt.[35] So, a deontological approach that holds the middle between evidence and epistemic values is appropriate if we want to talk about epistemic justification for science. Epistemic justification is then the set of reasons we give for adopting or rejecting certain scientific beliefs, where those reasons are evaluated in terms of a set of epistemic values. The difference between explanation and justification can then be put thus: explanation aims to answer why something happened

[31]Steup and Neta 2020
[32]Steup and Neta 2020
[33]Steup and Neta 2020
[34]Steup and Neta 2020
[35]Kuhn 1988, p. 357

in the first place, whereas epistemic justification aims to answer why we *believe* something happened.

It is still true that having good explanations can function as an epistemic justification for belief, just as we can use epistemic justification to argue for believing in one explanation over another, but the point is to show they are different and in which way.

I have argued that unificationism is an insufficient condition for explanation. What, then, *is* the relation between the two? I argue that unification serves as epistemic justification for choosing between two sets of explanations. A set of explanations that is more unifying does better in terms of two epistemic values we typically adopt in science: simplicity and scope. When a theory has fewer explanatory patterns, it is fair to say it is simpler. When it is able to explain more phenomena, its scope is bigger. These are things science strives for, and they are intuitively connected to explanation: to have a wide scope is to explain many different phenomena, and to have a simple theory means being able to explain phenomena with a smaller set of assumptions. Better performance in terms of these factors is a valid epistemic justification, but that does not mean something *is* an explanation in virtue of being part of a more unifying set.

The point of this chapter was to motivate focusing on causal theories of explanation, by arguing that the main alternative, unificationism, fails as a theory of explanation. In showing that unification is an insufficient condition for explanation and that unification is better understood as an epistemic justification, I motivate investigating whether there are causal theories of explanation that are compatible with Bell phenomena. If there was nothing wrong with unificationism, there would have been no need to try to solve causal theories' quantum problems, but as it stands it seems unificationism is not untroubled, so the effort is justified. As noted in the beginning of this chapter, however, there are many different causal theories of explanation. The next chapter introduces two such theories.

Chapter 2

Causal Explanation

The problem that quantum phenomena pose to causal accounts of scientific explanation is dependent on what causality is taken to be. That is, which account of causal explanation we are talking about because every theory of causal explanation, needs a theory of causality itself. This chapter will look specifically at two causal theories of scientific explanation: interventionism and the conserved quantities theory of causation (CQTC). First up is the interventionist theory of causal explanation as advocated by James Woodward.[1] This is one of the most prominent accounts of causal explanation, and it thus makes sense to start here. Second is Phil Dowe's CQTC.[2] Dowe's account is influential within causal mechanical accounts of causation, and, more importantly, it is one of the few theories that pays explicit attention to quantum mechanics.

The chapter can thus be divided into two: one part about Interventionism and one part about CQTC. In doing so, I set the stage for the next chapter, in which I will argue that the description of Bell phenomena given by the standard interpretation of Quantum Mechanics, also called the Copenhagen interpretation, does not count as explanation according to these theories.

[1] Woodward 2005
[2] Dowe 2000

2.1 Interventionism

2.1.1 The Theory

Simply put, Woodward's interventionism says A is a cause of B if there is a possible intervention on A such that the value of B changes. If there is no way to intervene in the process between A and B, not even principally, then that process is not causal.[3] The motivation for this approach is pragmatic: Woodward says we want to know the causes of the world so we can effectively intervene in the world, so the notion of intervention should be central to the notion of cause.[4]

Now, what does Woodward mean by intervention? He argues that the kind of intervention that, if possible, warrants talk of causality satisfies three conditions. Where I stands for the intervention in question, X for the would-be cause and Y for the would-be effect:

1. I is the only cause of X. Whereas there might have been other factors determining the value of X, once made the intervention completely determines X.

2. I causes Y through X, and only through X. There is no other route through which I causes Y.

3. There is no common cause of I and Y.[5][6]

Now we know what an intervention is, but Woodward's definition of a cause does not ask for an intervention, but a *possible* intervention, so we need to get a clear sense of the word possible here. Woodward is not talking about possible in the sense associated "with what humans can do", because that would exclude a large class of relations that we would arguably want to exclude.[7] We could not, for instance, talk about a causal relation between two events

[3] Woodward 2005, p. 26
[4] Woodward 2005
[5] Woodward 2005, pp. 96–98
[6] At face value, this is blatantly circular. Woodward however denies that. He argues that his is not a reductive theory, i.e. he is not trying to reduce causality to intervention, thus these definitions are not viciously circular in the sense that x causes is dependent on what is caused by x. Much discussion abound, but for the purposes of this thesis I will let this question rest.
[7] Woodward 2005, p. 128

in the past, because we are not able to intervene on the past. Neither is it physical possibility Woodward is after, because that too rules out cases we want to include.[8] For instance, we want to say that the moon's gravitation field causes the tides to be the way they are, yet it is not clear that there is a physically possible (meant in the weak sense of: being in congruence with laws of nature) intervention on the moon that satisfies the conditions set out above.[9] Rather, Woodward means that "an intervention on X with respect to Y will be "possible" as long as it is logically or conceptually possible for a process meeting the conditions for an intervention on X with respect to Y to occur."[10]

What does it mean that an intervention is logically/conceptually (im)possible? For one, it needs to be well-defined, in the sense that we can think of X as a variable that is subject to change, i.e. X needs to be a variable with some value, and we need to be able to conceptualize a change in said variable.[11] Moreover, change in X must be a notion that is clear enough to disentangle the effect of X's change on Y from the effects on Y caused by I through some other path.[12] Which is to say, an intervention needs to be well-defined in the sense that it needs be clear *that* and *how* it satisfies the conditions constraining intervention. So, where X is the would-be cause and Y the would-be effect, an intervention I is possible if:[13]

- A change in X is logically possible.
- We can differentiate between a change in Y due to a change in X and a change in Y due to other variables.

Perhaps strangely, then, whether an intervention I is possible has little to do with I itself, but rather with the situation we are researching. To make the connection between an intervention and its possibility stronger, we can take a look at a footnote in the paper "Independence, invariance and the causal Markov condition", where Woodward and Hausman briefly discuss

[8]Woodward 2005, p. 132
[9]Woodward 2005, p. 129
[10]Woodward 2005, p. 132
[11]Woodward 2005, p. 128
[12]Woodward 2005, p. 131
[13]Woodward 2005, p. 131

what this kind of possiblity means to them. We can glean the following interpretation: even if humans can not execute the interventions, and even if they are not physically possible, interventions can be logically/conceptually possible in the sense that there is a theory which "provides the resources for representing them and predicting their consequences."[14] So, for an intervention to be possible, not only do cause and effect need to satisfy the two conditions mentioned above, there also needs to be a theory that can represent the intervention and its consequences, whether it could actually happen or not.

Now that there is an idea of what an intervention is, we know what a cause is by remembering the characterisation I gave at the beginning of this chapter: A is a cause of B if and only if there is an intervention I on A such that the change in the value of A brings about a change in the value of B.[15] Knowing what a cause is, however, is not the same as knowing what a causal explanation is.

Woodward suggests that a causal explanation is "a matter of exhibiting systematic patterns of counterfactual dependence."[16] Generalizations allow us to show in conjunction with initial and boundary conditions that the explananda were to be expected, *and* how if some of these conditions were to change, the explananda would change.[17] That is to say, we explain things by use of counterfactuals: by asking "*what-if-things-had-been-different*".[18] Woodward's concept of causation is related to his concept of causal explanation through the counterfactual, because the counterfactuals he is concerned with are those that invoke interventions.[19] So, to be precise, we are asking *what-if-there-had-been-an-intervention*. An explanation E of an explanandum M such that M is the statement that a variable B takes value b consists of:

a) A generalization that establishes a relation between a variable A and B.

b) A statement that A has the particular value a.[20]

[14]Hausman and Woodward 1999, p. 537
[15]Woodward 2005
[16]Woodward 2005, p. 191
[17]Woodward 2005, p. 191
[18]Woodward 2005, p. 191
[19]Woodward 2005, p. 196
[20]Woodward 2005, p. 203

E then has to satisfy three conditions to be considered a proper explanation for M:

1. M and E are (approximately) true.
2. The generalization predicts that *if* A takes value a, *then* B takes value b.
3. There is an intervention such that a changes to a' (where $a \neq a'$) where through the generalization we can accurately predict the b' for B such that $b \neq b'$ under this intervention.[21]

We now have the notion of an intervention, a cause and a causal explanation, thus we have Interventionism as advocated by James Woodward.

2.2 Conserved Quantity Theory of Causation

Another theory of causal explanation takes causation to fundamentally be the exchange of conserved quantities. From this perspective, causation is a physical process. The theory is proposed by Phil Dowe, in response to two earlier theories: Wesley Salmon's mark transmission theory of causation on the one hand and Jerrold Arronson and David Fair's transference theory of causation on the other.[22] The mark transmission theory casts causation in terms of processes, where only those processes which transmit 'marks' are considered causal.[23] Transference theory argues that causation is the transference of certain quantities from cause to effect, where which quantities count depend on whether you are asking Aronson or Fair.[24][25] Phil Dowe distinguishes three questions about causation which, taken together and answered, give his theory of causality. The questions are:

1. What are causal processes and interactions?
2. What is the connection between causes and effects?
3. What distinguishes a cause from its effect?[26]

[21] Woodward 2005, p. 203
[22] Illari, Russo, and Williamson 2011, p. 2
[23] Salmon 1984
[24] Aronson 1971
[25] Fair 1979
[26] Dowe 2000, p. 89

I will work through these questions and his answers consecutively and in doing so outline the theory.

2.2.1 Causal Processes

What, for Dowe, distinguishes a causal process from other processes is that causal processes possess conserved quantities.[27] He offers the two following definitions for causal process and causal interaction respectively:

1. CQ1. A *causal process* is a world line of an object that possesses a conserved quantity.
2. CQ2. A *causal interaction* is an intersection of world lines that involves exchange of a conserved quantity.[28]

To understand these definitions, we need to understand the concepts used in them.

Dowe calls an object anything which we find in either the ontology of science or common sense. So, a particle is an object because we find it in the scientific ontology, and a chair is an object because we find it in our common sense ontology.[29] Not all objects, then, can be part of a causal process. For instance, a shadow is something we find in the common sense ontology and is thus an object, but it can not possess a conserved quantity and thus can not be an object in a causal process.

A process, for Dowe, is the world line of an object and a world line is "the collection of points on a space-time (Minkowski) diagram that represents the history of an object."[30] A space-time diagram is simply a graphical illustration of space and time and, more specifically, a Minkowski diagram is a graphical illustration of space and time which uses two dimensions, one for space and one for time, with time on the vertical axis and space on the horizontal axis.[31] The total collection of points on this diagram that belong to an object thus forms

[27]Dowe 2000, p. 89
[28]Dowe 2000, p. 90
[29]Dowe 2000, p. 91
[30]Dowe 2000, p. 90
[31]Mermin 1997

a line which represents the total history of that object in space and time, i.e. that object's word line.

Any quantity that is subject to a conservation law is a conserved quantity.[32] Conservation laws state that a quantity is universally conserved, which is to say that a quantity whose total value in a closed system does not change is a conserved quantity.[33] At this point circularity looms, because what is a closed system if not a causally isolated system. To evade this problem Dowe seems to define closed systems as systems where there is no net flow of quantities in or out of said system. It is, however, questionable whether this solves the issue, since what are flows of such quantities, if they are conserved, if not causal processes, so that the definition in terms of such flows comes down to a definition barring external causal influences.

Moreover, looking ahead slightly, if we interpret this definition as meaning that we can only describe causation in scenarios where we have a closed system, accounting for quantum mechanics becomes highly problematic, because a significant part of quantum mechanics issues have to do with the interaction of quantum level closed systems with the wider world the moment they are measured, i.e. if we can only talk about closed systems we can not talk about decoherence, which would severely limit any causal account of quantum mechanics. However, we need not take this definition to mean this, because we can also take it to mean that, whether we are talking about a closed system or not, the conserved quantities are those whose total value, *had* they been in a closed system, would not have changed.

Examples that Dowe gives are momentum, energy and charge.[34] If we say an object *possesses* a conserved quantity, we mean that according to science we can truthfully attribute that quantity to that object.[35] At this point it is worth looking ahead, by noting that here we have given an account of conserved quantities in classical physics, but in quantum mechanics this notion of conserved quantities does not easily translate. So, when we do discuss quantum

[32]Dowe 2000, p. 91
[33]Dowe 2000, p. 95
[34]Dowe 2000, p. 94
[35]Dowe 2000, p. 92

mechanics and CQTC, we will have to revisit what it means for a quantity to be conserved.

We have now defined all terms needed to understand CQ1, and in order to understand CQ2 we only need to elucidate the meaning of *intersection* and *exchange*. An intersection between two processes consists of all space-time points the world lines have in common; if there are no such points, there is no intersection.[36] An exchange is an event where at least two processes have a change in the value of their conserved quantities after an intersection.

2.2.2 The Connection Between a Cause and its Effect

Having given CQ1 and CQ2 and defined all the terms within them, I have given Dowe's answer to the first question: *What are causal processes and interactions?* Now, it is time to turn to his second question: *What is the connection between causes and effects?* Dowe sets out to answer this question in terms of causal processes and interactions, rephrasing the question as "If two events are related as cause and effect, what causal processes and interactions must there be?"[37]

A straightforward answer to that question would be that two events A and B are causally connected if and only if there is a succession of causal processes and interactions from A to B.[38] Dowe argues that the problem with that account, call it the naive picture of causal process theories, is that it has a problem with 'misconnections'.[39] We can find counter-examples that undermine this naive picture as a sufficient condition for causality. Take the following example due to Nancy Cartwright.[40] Say I have an unwanted plant in my garden and decide to spray it with a defoliant, which kills 90% of the plants it is sprayed on. Intuitively, the spraying of the defoliant explains the 90% that dies but does not explain the 10% that survives. Now say that the plant in my garden is part of the 10% that survives.[41] Just as when the plant had died, there is a succession of causal processes and interactions

[36]Dowe 2000, p. 91
[37]Dowe 2000, p. 146
[38]Dowe 2000, p. 147
[39]Dowe 2000, p. 148
[40]Cartwright 1983, p. 28
[41]Cartwright 1983, p. 28

from the spraying of the defoliant to the survival of the plant, so if we were to hold on to a naive picture of causal process theory we would say that the spraying of the defoliant causes, and explains, the survival of the plant.[42] The point is that you can have, and often do have, a schema of causal processes and intersections which involves events or states of affairs that cannot be said to cause or to be caused by any causal process or intersection in the schema.[43]

Dowe thus argues that the naive picture can not be the answer to our present question. His suggestion is instead that we can relate cause and effect by a process theory that draws on David Armstrong's metaphysics of states of affairs.[44] Our first question would then be, what kind of metaphysics is that?

The idea is to give an account of what causal relata are.[45] For Dowe, causal relata are either events or facts. An event is the change in a property of an object at some time. A fact is the having of a property by an object at some time or period of time.[46] Dowe focuses on facts, and so shall we. Causal facts are neither negative nor disjunctive.

Now that we can give an account of the causal relata, we can give Dowe's account of the kind of connection between such relata counts as a causal connection. Say we have two causal interactions (I_1 , I_2) which are connected by a causal process p. Then Dowe says that:

> Interactions (I_1 I_2) are linked by a causal connection by virtue of causal process p only if some conserved quantity exchanged in I_1 is also exchanged in I_2 and possessed by p.[47]

Why was the account of causal relata necessary for us to come to this definition? In order for us to come to a more general definition of causal connection. Say that we write a cause as $q(a)$ and an effect as $q'(b)$, where a and b are objects and q and q' are conserved quantities

[42]Dowe 2000, p. 148
[43]Dowe 2000, p. 149
[44]Dowe 2000, p. 147
[45]Dowe 2000, p. 168
[46]Dowe 2000, p. 170
[47]Dowe 2000, p. 171

possesed by those objects.[48] Then Dowe gives a general definition of a causal relationship as:

> There is a causal connection between q(a) and q'(b) if and only if there is a set of causal processes and interactions between q(a) and q'(b) such that:
>
> 1. Any change of object from a to b and any change of conserved quantity from q to q' occur at a causal interaction involving the following changes: $\Delta q(a)$,$\Delta q(b)$,$\Delta q'(a)$, and $\Delta q'(a)$.
> 2. For any exchange in (1) involving more than one conserved quantity, the changes in quantities are governed by a single law of nature.[49]

Dowe does not offer a deep explanation of what a law of nature is supposed to be. According to him, simple covariance is sufficient for his purposes. [50]

Let us return to Cartwright's splayed plant example to see how Dowe's account solves the naive picture's problems. The key is to see that what is essential is whether there is a thread of causal processes that has the effect to be described. So, say the plant had died. Then the spray killed certain cells of the plant that were essential for keeping the plant alive, and so there is a continuation of processes that kills the plant. If the plant survives, the defoliant might kill some cells of the plant, but although we can continue tracing causal processes and causal interactions from there, those interactions do not satisfy the constraints outlined here that force the conserved quantities involved to be related, and the plant survives because of some other unaffected cells.[51]

2.2.3 Distinguishing Cause and Effect

The final question, *what distinguishes a cause from its effect?*, concerns the asymmetry of cause and effect, i.e. causes and effects ought not to go in circles, so there must be

[48] I.e. *q(a)* is the *fact* that a posses a certain value of q, and similarly for *q'(b)*
[49] Dowe 2000, p. 172
[50] Dowe 2000, p. 172
[51] Dowe 2000, p. 173

some way to distinguish the two.[52] Dowe's goal is to give an account of the asymmetry of cause and effect, *without* positing a time directional demand. He explicitly aims to account for backwards-in-time causation (sometimes called retro-causality), on the assumption that quantum mechanics shows us that such causation exists.[53] That immediately shows us that there is at least one way in which CQTC intends to deal with Bell phenomena: through retro-causality.

Before we turn to quantum theory, however, we should explain that if it is not through temporal asymmetry, how Dowe *does* intend to distinguish causes from effects. He does so through a version of the fork asymmetry theory.[54] In a sentence, Dowe argues that the direction of a causal process is equal to the direction of an open conjunctive fork partially constituted by that process, unless there is no such fork, in which case the direction is the direction of the majority of open forks in the same net as the process.[55] That sentence needs some elucidating.

First, what are conjunctive forks? A conjunctive causal fork is a set of three events: two effects and one common cause (the first articulation of fork asymmetry theory is due to Reichenbach).[56] That means that a conjunctive causal fork is as set of three events A, B and C such that:

- $p(A \mid B) > p(A)$ and $p(B \mid A) > p(B)$
- $p(A \mid C \wedge B) = p(A \mid C)$ and $p(B \mid C \wedge A) = p(B \mid C)$

Where $p(X)$ stands for the probability of X happening, and $p(X \mid Y)$ stands for the probability of X happening if/assuming Y happens.

What does it mean for a conjunctive fork to be closed or open? A closed conjunctive fork has a screening off event in the past of the two effects *and* in its future. Conversely, an open fork does not have a screening off event in both its past and future.[57]

[52] Dowe 2000, p. 176
[53] Dowe 2000, p. 177
[54] Dowe 2000, p. 192
[55] Dowe 2000, p. 204
[56] Dowe 2000, p. 192
[57] Dowe 2000, p. 194

The second part of Dowe's version of the fork asymmetry theory talks about nets. The basic idea is that causal processes are linked together in a network, which contain at least some open forks. It is then direction of the majority of the open forks in the network in which a particular causal process exists that determine the direction of the said causal process.[58]

We have by now fully fleshed out CQTC, with our only missing piece being how we then give *an explanation* in terms of this kind of causation. Dowe argues that to give a causal explanation of a phenomenon, is to show how it fits into the causal structure of the world.[59] So, if I want to explain why the window broke, I can say that it is because I threw a ball at it, which would be fitting it into the causal structure of the world on a 'manifest' level, which then in turn is reducible to a physical level in which we describe causal processes and interactions from my throwing to the glass shattering such that all changes in objects and conserved quantities occur at causal interactions that involve $\Delta q(a)$,$\Delta q(b)$,$\Delta q'(a)$, and $\Delta q'(a)$, and if the interaction involves more than one conserved quantity, that interaction is governed by a natural law.[60]

2.3 Causation: concept and empiricity

Before concluding this chapter, I want to focus on the question whether these two theories aim to answer the same question. Dowe argues that when we ask 'what is causation?' there are, at least, two possible interpretations of that question.[61] For the first interpretation, the question is conceptual and the answer will be an explication of our common sense concept of causation. For the second interpretation, which Dowe claims as the one he is concerned with, the question is empirical, and thus the answer ought to tell us what causation is in the objective world.[62] It seems Dowe wants to categorize theories like interventionism in the

[58]Dowe 2000, pp. 198, 205
[59]Dowe 1992, p. 128
[60]Dowe 2000, p. 172
[61]Dowe 2000, p. 1
[62]Dowe 2000, p. 1

first camp, which can be explained by referring to the motivation for interventionism. The focus is on the use of knowledge about causal relations , which, according to Woodward, is the ability to effectively intervene on the world, so we might change it to our advantage. If we take this pragmatist approach, defining cause using its common sense function, it looks like we are interested in what the proper common sense definition of causation is. However, Woodward himself objects to this characterisation. According to him, although interventionism is different from Dowe's empirical analysis of causation, it is not squarely conceptual either.[63] There is much attention on how we use concepts like causation and explanation, but Woodward looks to these concepts and their use in practice to find the goal of such practices, and then proceeds to formulate a *normative* theory of the concept of causation.[64] In this sense he does not merely give a conceptual analysis of the way we use the notion of causality, but argues for a revision of that notion. On the basis of what does Woodward make such recommendations? Generally speaking, Woodward looks to the purpose of the concept and makes recommendations so that the concept can better fulfill its function.[65][66] However, one way to understand whether the concept does fulfill its function, is empirical, so the distinction is not so easily made.

Whether or not you believe we can or cannot make such clear distinctions, what should be clear is that these are not two theories that, using the same method, come upon different results, but instead their approach is different from the start, which might inform the decision the pursue one or the other. For this thesis, however, that choice will predominantly be determined by how they fare in dealing with Bell phenomena. To that end, the next chapter will introduce Bell phenomena as conceived of in the Copenhagen interpretation, and evaluate both theories in light of their ability to explain such phenomena.

[63]Woodward 2005, p. 7
[64]Woodward 2005, p. 7
[65]Woodward 2005, p. 8
[66]It's not the only metric, for instance he also mentions epistemological constraint.

Chapter 3

Explaining Bell Phenomena in the Copenhagen Interpretation

This chapter marks the first step of this thesis in connecting causal theories of scientific explanation to interpretations of quantum mechanics. To do so, I will first give a brief introduction to quantum mechanics, specifically I will describe the quantum formalism for making predictions and highlight some contentious issues in the interpretation of quantum mechanics, after which I can properly describe the phenomenon that is the focus of this thesis. Once I have done so, I will show how Bell inequalities exclude explanations of Bell phenomena in terms of common causes and, finally, argue that neither interventionism nor CQTC characterizes the description of Bell phenomena given by the Copenhagen interpretation as explanation.

3.1 Quantum Mechanics

In order to understand both Bell phenomena and the issues that come up when connecting causal theories of scientific explanation to interpretations of quantum mechanics, we need to get some understanding of quantum mechanics. To that end, this section starts by describing the 'recipe' with which we make predictions in quantum mechanics, if we were to forego

interpreting the tools we use to do so. Then I will introduce some central issues of dispute (wave-particle duality, the measurement problem, the uncertainty principle, Bell inequalities) between different interpretations of quantum mechanics that will turn out to be important in matching theories of explanation to quantum mechanics. In discussing these I will also introduce the concepts (spin, entanglement) needed to describe the phenomenon through which I try to find said match, i.e. entangled anti-correlated 1/2 spin particles.

3.1.1 The Predictive Recipe

Quantum mechanics without interpretation does not present us with a picture of the world, but it does allow us to very accurately predict certain empirical outcomes. It gives us a set of steps, or recipe, to make predictions. To get a grip on the challenges any actual interpretation of quantum mechanics must meet, it is good to understand this recipe, so this section will set out to describe it. However, a preliminary with a slight head start is needed here, because there are many complicated quantum systems, for which the quantum recipe also becomes more complicated. Thus, I discuss the simplest version of the recipe, the non-relativistic version dealing with 1/2 spin particles. What 1/2 spin particles are is not essential at this point, and will be described in detail later on in this chapter.

The first step in the quantum recipe is to associate a wavefunction (sometimes denoted by ψ or $\psi(x)$) with each particle in the experiment at hand.[1] A function is a mapping from one set of (mathematical) objects to another, so too the wavefunction, but exactly *what* in the physical world it is mapping, will depend on the interpretation you ascribe to.[2]

In the case of single 1/2 spin particle, ψ will take the form of a "*complex square-integrable function over space, as a function of time.*"[3] Even at its simplest, this sentence is dazzling, so let's break it down. First, complex functions over space are functions that map complex numbers to each spatial location, so if we coordinatize space in the usual manner using three coordinates (x, y, z), then each combination of values of x, y an z gets mapped

[1] Maudlin 2019, p. 37
[2] Maudlin 2019, p. 37
[3] Maudlin 2019, p. 37

to a complex number.

A complex function is square-integrable when we can take the product of it and its complex conjugate and measure the total volume under the function and end up with a finite number. But what's a complex conjugate! First, a complex number can be represented as $A + Bi$, where A and B are real numbers and i the imaginary square root of -1. For any complex number, there is a complex conjugate which we get by turning the imaginary part of the complex number either positive or negative, i.e. $A + Bi$'s complex conjugate is $A - Bi$.[4]

The part we have not detailed yet, is how ψ is a function of time. I.e. how does the wavefunction change as time progresses. It does so in the recipe in accordance with Schrödinger equation. In the nonrelativistic context we are considering, the equation is:

$$i\hbar\frac{\delta\psi(x,t)}{\delta t} = \hat{H}\psi(x,t)$$

where $\hbar$ is Planck's constant divided by 2π, $\frac{\delta\psi(x,t)}{\delta t}$ specifies the change of $\psi(x,t)$ over an infinitesimal amount of time and $\hat{H}$ stands for the Hamiltonian operator.[5] The Hamiltonian is a function that maps one wavefunction to another, and needs to be specified itself. For our present purposes, however, it is enough to note that Schrödinger's equation is a *linear* equation, such that, for example, if ψ turns into $\psi(t)$ after some time t and ϕ turns into $\phi(t)$ after some time t (ψ and ϕ both being wavefunctions), then $\alpha\psi + \beta\phi$ turns into $\alpha\psi(t) + \beta\phi(t)$.[6]

What rests is a way to go from these functions to real world observations.[7] The recipe makes this translation using the Born rule, which essentially says to use ψ to define a probability measure.[8] If we a have a number of possible outcomes, a probability measure is a function that assigns a real number between 0 and 1 to each such outcome or combination of such outcomes. Moreover, if we take the set of all possible outcomes, the measure must assign it 1 and if the measure assigns a number of disjoint (meaning they have no elements

[4]Maudlin 2019, p. 40
[5]Maudlin 2019, p. 42
[6]Wallace 2013, p. 4
[7]Maudlin 2019, p. 46
[8]Bohr et al. 1928

in common) subsets of our outcome set some range of values, then it must assign the union of these subsets the sum of those values.[9] So, for instance, if I have a probability measure for the location of a ping pong ball under three red cups, and that measure assigns 0.25 to cup A and 0.35 to cup B, then it must assign 0.60 to the possibility that the ball is under either A or B.

We use said probability measure to predict which particular result will obtain, e.g. if we were firing particles at a screen we would use the probability measure to predict the likeliness of observing a flash at certain locations. Born's rule defines a probability measure over ψ by taking the squared amplitude of the wavefunction and integrating it. The rule is set up like this because ψ only assigns complex numbers to locations in space, and we need to get real numbers between 0 and 1.[10] Taking the squared amplitude means we only get non-negative real numbers, and integrating it means any set of outcomes (if we were predicting locations on a screen, any region on the screen) gets a real number between 0 and 1 while the total set of outcomes gets 1.[11] Thus we have a way to connect the wavefunction and Schrödinger's equation with experimental outcomes, and thus the recipe is finished.

Any interpretation of quantum genetics must give an account of these three aspects of quantum mechanics. Some say the wavefunction does not tell us all there is, some say Schrödinger's equation is wrong and some disavow Born's rule, but they must contend with them nonetheless, and if they critique them they must give an alternative that makes sense of the success of the predictive recipe.

Now I will turn to some major points of contention in the interpretation of quantum mechanics.

3.1.2 Wave-particle Duality

As it will turn out, matters of ontology are very important when trying to fit causal theories of explanation to interpretations of quantum mechanics. As such, I will introduce one issue

[9] Maudlin 2019, p. 46
[10] Maudlin 2019, p. 46
[11] Maudlin 2019, p. 47

of ontology that often takes centre stage when it comes to quantum mechanics: wave-particle duality. The question is whether the stuff of quantum mechanics consists of waves or particles, or both or neither, and the question arises because we have empirical data pointing towards both waves and particles.[12]

First, let us take a look at why we might think in terms of particles. To do so we need to know what a cathode ray tube is. Take the positive electrode of a battery, i.e. the anode, and the negative electrode, i.e. the cathode, and wire them to a high voltage generator. Make a hole in the anode and place a phosphor coated screen beyond it. Finally, connect a heating element to the cathode and enclose the apparatus in an evacuated tube, and you have a cathode ray tube.[13] Now, if the heating element is quite hot, we see a shape appear on the screen that corresponds roughly to the hole in the anode. However, when we lower the heat, after a while we will not see a continuous spot but instead individual flashes that over time trace out the same area as the original shape.[14] What that seems to suggest is that the rays coming from a cathode ray consist out of many individual particles (what we now call electrons).[15]

Let me introduce two other experiments: the single and double slit experiment. Starting with the single slit experiment, imagine we shoot a ray at a screen, but place a barrier with a slit in between the cathode ray and the screen. At first, what happens is what we might expect. A shape in the form of the slit appears on the screen and when we make the slit thinner, the shape also becomes thinner.[16] However, as we continue to make the slit thinner, at some point the shape will become wider, as illustrated in figure 3.1.

Before searching for an explanation here, let us look at the double slit experiment, where instead of one we put two slits in the barrier. When we fire a ray at the barrier/screen now, we see "alternating light and dark bands on the screen," as seen in figure 3.2.[17] If our ray

[12]Maudlin 2019
[13]Maudlin 2019, p. 6
[14]Maudlin 2019, p. 7
[15]Maudlin 2019, p. 8
[16]Maudlin 2019, p. 8
[17]Maudlin 2019, p. 11

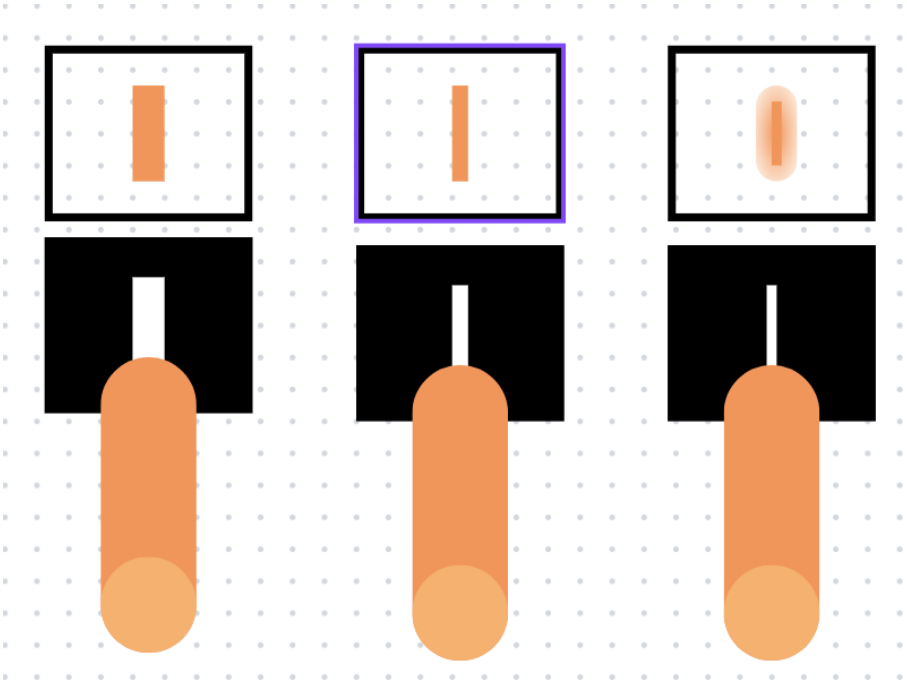


Figure 3.1: Diffraction

was simply a collection of particles, we would classically expect to just see two slit shapes on the screen.

There is no ready explanation for either phenomenon if we take the stuff at hand to consist of particles.

So, this is all quite strange. It becomes less strange when we take the rays to be waves instead of particles because waves exhibit analogous phenomena to what we see in these experiments: diffraction and interference.[18] When waves hit a gap in a barrier that is small relative to their wavelength, it creates a circular wave pattern after that barrier.[19] We call this diffraction. If there are two such gaps in the barrier, the two circular waves behind the barrier interact with each other.[20] When one wave's crest meets the other's trough, they cancel out, and if crest meets crest or trough meets trough, they amplify each other, which would create the kind of band pattern we saw in the double slit experiment.

Then, why not be done with it and conclude the stuff of quantum mechanics consists

[18]Maudlin 2019, pp. 8–12
[19]Maudlin 2019, p. 9
[20]Maudlin 2019, p. 9

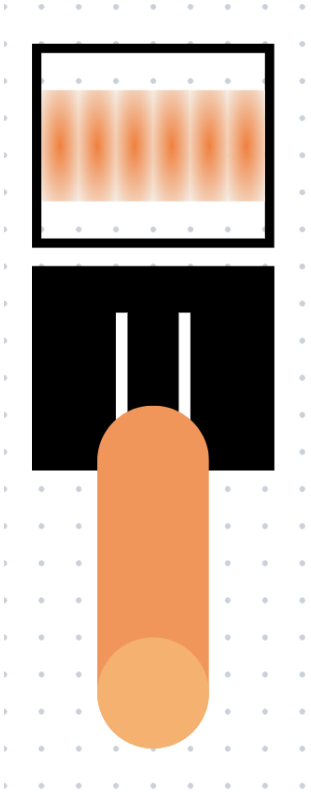


Figure 3.2: Interference

of waves? If we again lower the heating of the device we use to send the rays towards the barrier/screen, we again see individual flashes instead of a continuous shape, thus these phenomena still suggest particles too. Those individual flashes, crucially, over time still trace out the same shape in both the single and double slit experiment. That is, they still exhibit diffraction and interference.[21]

Thus the data that we have is not obviously decided on the ontological stuff of quantum mechanics, and as we shall see the different ways in which interpretations of quantum mechanics handle this question is important in the question of fitting causal theories of scientific explanation to them.

One way people have attempted to deal with wave particle duality, is to say that it is a wave right up until we measure, and then the wave 'collapses'. But it is not entirely clear what a measurement is, and how it collapses the wavefunction. This is called the measurement problem, and how it works I will describe next.

[21] Maudlin 2019, p. 9

3.1.3 The Measurement Problem

Before explaining the measurement problem, I will describe some phenomena that will hopefully make it easier to understand the problem itself.

Spin

Classically, a charged particle that is spinning is magnetically polarized, and thus changes its trajectory when it moves through an in-homogeneous magnetic field.[22][23] If a particle that is magnetically polarized would be moving through a field with a stronger north pole relative to its south pole, it would be pushed down if that particle's north pole is orientated up, because the two north poles would repel each other more than the two south poles. Conversely, if the particle's north pole is oriented down, it would be pushed up, because the attraction between the field's north and the particle's south pole would be stronger than the attraction between their south and north poles respectively, as seen in figure 3.3.

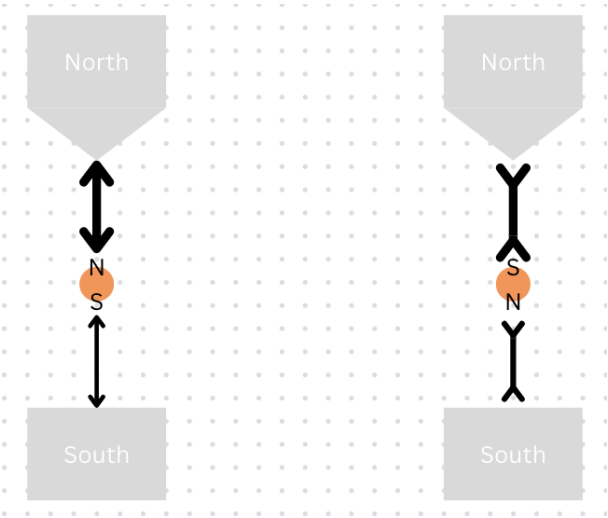


Figure 3.3: *Poles*

What we would expect classically if we were to point cathode rays at an in-homogeneous magnetic field, is that some electrons would go straight through and that we'd see a continu-

[22] Maudlin 2019, p. 18

[23] Spin is possibly somewhat of a misleading term: we know electrons have an electric charge and magnetic field. In classical physics spinning electrical charges create magnetic fields, thus we say that our electrons have spin, but we do not know whether there is actually something spinning.

ity of different deflections dependent on the orientations of the electrons we shoot through. That, however, is not what happens. Instead, we find that *no* electrons go straight through and that all electrons are either pushed down or up to the same degree. Instead of a continuity of results, our results are discrete.[24] This phenomenon, the change from continuous to discrete outcomes, we call quantization, and spin is the simplest phenomenon that exhibits quantization.[25]

One device that creates such an inhomogeneous magnetic field is called a Stern-Gerlach (SG) apparatus. If we were to send a bunch of electrons through a SG apparatus that is oriented along some axis x, about half of them will be deflected upwards. If we then send all those deflected upwards through another SG apparatus that is oriented along the x axis, we see that *all* electrons get deflected upwards. However, if we send them through an apparatus that is oriented along a perpendicular y axis, we see that again half is deflected upwards and half downwards, *and* if we send them through a x oriented apparatus after that, again half get deflected up and half down. It turns out that the electrons are either spin-up or spin-down in any direction, and if it has a spin-property in one direction, it is principally impossible to have a spin-property in another direction.[26] Moreover, because of this indetermineteness, the Copenhagen interpretation says particles have *no spin* until measured.

The Measurement Problem in Two Ways

The measurement problem in the manner defined just now is still somewhat vague. To be a little more precise, I will introduce the issue in two out of the three guises that Tim Maudlin formulates in "Three Measurement Problems".[27][28] The first problem Maudlin calls the problem of outcomes, and the second the problem of statistics.

[24]P. J. Lewis 2016, p. 7
[25]Maudlin 2019, p. 18
[26]P. J. Lewis 2016, p. 8
[27]Maudlin 1995
[28]The third version, termed the effect problem, is no issue for the interpretations discussed in this thesis and so I leave it be.

Take the following three claims:

1A The wavefunction is physically complete.

1B The wavefunction evolves in accord with a linear equation (e.g. the Schrödinger equation).

1C Measurements of, e.g., spin have determinate outcomes.[29]

These three statements mirror the three steps in the recipe for making quantum predictions. However, they cannot all three be true. If both 1A and 1B are true, 1C must be false. Imagine we have a particle whose spin we measure in the z-direction and turns out to have spin up, and then subsequently we measure it in the x-direction. Said particle, before measurement, is then in the superposition $\frac{1}{\sqrt{2}}(|\uparrow_x\rangle + \frac{1}{\sqrt{2}}\ |\downarrow_x\rangle$. The initial wavefunction of the particle $\otimes$ SG apparatus is then $\frac{1}{\sqrt{2}}(|\uparrow_x\rangle + \frac{1}{\sqrt{2}}\ |\downarrow_x\rangle \otimes \mid READY\rangle_{sg}$, and if 1B is true it must evolve into $\frac{1}{\sqrt{2}}(|\uparrow_x\rangle \otimes \mid UP\rangle_{sg} + \frac{1}{\sqrt{2}}\ |\downarrow_x\rangle \otimes \mid DOWN\rangle_{sg}$. And that is all there is to the story, if 1A is true, but then there is no determinate outcome to the measurement, so 1C is false.[30] To avoid this outcome, i.e. make 1C true, we need to either say 1A is false and say *what* else there is over and above the wavefunction, or say 1B is false and say *how* the wavefunction does evolve. Different interpretations make different choices here, and we shall explore those differences in depth in the next chapter, but for now it suffices to say that the Copenhagen interpretation must deny either 1A or 1B, because it postulates that upon measurement the wavefunction collapses and said collapse is not captured by the wavefunction itself nor by the Schrödinger equation.[31] The problem with the Copenhagen interpretation is that it simply does not tell us *what* more there is, and instead stipulates that upon measurement ψ collapses. As we shall see, precisely this lack of detail makes the Copenhagen interpretation unfit for either an interventionist or a CQTC reading.

The second problem centers on these three claims:

2A The wavefunction is physically complete.

[29] Maudlin 1995, p. 7
[30] Maudlin 1995, p. 8
[31] Maudlin 1995, p. 9

2B The wavefunction evolves in accord with a deterministic dynamical equation (e.g. the Schrödinger equation).

2C Measurements which are described by identical wavefunctions can have different outcomes.[32]

These three statements are also incompatible with one another. If 2B is correct then two systems with identical wavefunctions will always keep having identical wavefunctions since they evolve deterministically, and if 2A is correct then that is all there is to say about these systems, which means 2C can not be true, since they then can not have different outcomes.[33]

Thus the measurement problem can be understood in two ways, and in each iteration you must give up on one out of three claims that intuitively you would not want to.

The next issue is called the uncertainty principle, and it concerns the extent to which different quantities can have definite values simultaneously. As we have seen, CQTC sees causation essentially as the exchange of conserved quantities, thus it could prove problematic if such quantities are not always definite.

3.1.4 The Uncertainty Principle

In simple terms, the uncertainty principle states that there are pairs of values such that when you know one value *more* precisely, the preciseness with which which you can know the other value becomes *less*.[34] The principle is introduced by Heisenberg in relation to position and momentum.[35] Another example is the spin in different directions we just discussed: if I know a particle has spin-up in the x-direction, I am completely oblivious to its spin in the z-direction.[36] I describe it here as an epistemic issue, but it depends on which interpretation of quantum mechanics we are talking about whether that limitation is epistemic or ontological.[37] Simultaneously, to call it an epistemic issue has often been misinterpreted as claiming that it is a matter of having better measuring equipment that *can* allow us to know both, but

[32]Maudlin 1995, p. 11
[33]Maudlin 1995, p. 11
[34]Hilgevoord and Uffink 2023
[35]Heisenberg 1927
[36]Maudlin 2019, p. 21
[37]Hilgevoord and Uffink 2023

that is not strong enough.[38] Rather, the epistemic position claims that it is in principle impossible to know both, without claiming that they cannot both exist simultaneously.[39]

In the Copenhagen Interpretation it is generally taken to be an ontological limitation. So, say we are talking about the uncertainty principle for the momentum and position of a particle, then the Copenhagen Interpretation says it is not just that if we know a particle's position more precisely, we know its particle momentum less precisely, but rather that if the value of a particle's position *is* more precise, the value of its momentum *is* less precise.[40] Thus, if a particle has a completely precise position, it simply does not have a momentum value.[41]

3.1.5 Bell Inequalities

Finally, we turn to Bell inequalities. In order to understand Bell inequalities, we need to first have an idea of what "entanglement" is.

Entanglement

Suppose we have a spinless particle that devolves into two anticorrelated 1/2 spin particles, such that their respective spins are opposite, so that they cancel each other out. We say particles are entangled when their spin (or another property) are not independent of each other but (anti-)correlated.[42][43] The state these electrons are in is called the 'singlet state' and is mathematically represented as:[44]

$$\Psi = \frac{1}{\sqrt{2}}(|\uparrow_{1\theta}\rangle\ |\downarrow_{2\theta}\rangle - |\downarrow_{1\theta}\rangle\ |\uparrow_{2\theta}\rangle)$$

[38] Hilgevoord and Uffink 2023
[39] Hilgevoord and Uffink 2023
[40] Hilgevoord and Uffink 2023
[41] Note that this is different form its momentum being 0, that being a momentum value still.
[42] P. J. Lewis 2016, p. 8
[43] In this specific example we are talking about anti-correlated particles, but they can just as well be positively correlated. In fact, for Bell phenomena there are four Bell states, two of which imply correlation and two of which imply anti-correlation.
[44] Suárez 2007, p. 4

This expression needs some explanation. Call the particles '1' and '2' and represent the angle of measurement by θ. $|\rangle$ is the Dirac notation, which represents the state of a quantum system. So $|\uparrow_{1\theta}\rangle$ refers to the spin state of particle 1 measured at angle θ being up.

From the singlet state, we can derive information about the outcomes of measurements on each separate particle.[45] First, we apply the axiom of reduction to get the state of the individual particles:

$$W_1 = \tfrac{1}{2}\,|\uparrow_{1\theta}\rangle\langle\uparrow_{1\theta}| + \tfrac{1}{2}\,|\downarrow_{1\theta}\rangle\langle\downarrow_{1\theta}|$$

$$W_2 = \tfrac{1}{2}\,|\uparrow_{2\theta}\rangle\langle\uparrow_{2\theta}| + \tfrac{1}{2}\,|\downarrow_{2\theta}\rangle\langle\downarrow_{2\theta}|$$

Then, by applying the quantum statistical algorithm to W_1 and W_2 we can calculate the probabilities of the different outcomes in all measurement directions θ for particles 1 and 2:

$$prob(|\uparrow_{1\theta}\rangle) = \langle\Psi\,|\uparrow_{1\theta}\rangle\langle\uparrow_{1\theta}|\,\Psi\rangle = (\tfrac{1}{\sqrt{2}})^2 = \tfrac{1}{2}$$

Replacing 1θ for 2θ gives the probabilities for particle 2.

EPR

The fact that some particles are entangled, leads to strange conclusions. Einstein, Podolsky and Rosen (EPR) argue in their paper "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" that the fact of entanglement must imply that quantum mechanics is incomplete.[46] Imagine the following set-up, due to David Bohm: We have a spinless particle devolving into two 1/2-spin particles, which we again call particles '1' and '2'.[47] The 1/2-spin particles are sent in opposite directions, for arbitrary distances, after which both are measured by θ-oriented Stern-Gerlach apparatus. Now, we know if particle 1 has up-spin, then particle 2 must have down-spin. I.e., if we measure particle 1, we do not need to measure particle 2 in order to know its spin.

Not so surprising if we assume that the particles had their respective spin states all along. But, crucially, that's not the case in the Copenhagen interpretation of quantum

[45] Suárez 2007, p. 4
[46] Einstein, Podolsky, and Rosen 1935
[47] Bohm 1989



Figure 3.4: [48]

mechanics. Remember, faced with the measurement problem the Copenhagen interpretation says the wavefunction is true up until the moment we measure, at which point it collapses. The idea is that, taking spin as an example, a particle has no definite spin until it is measured, and that measurement brings about the result it obtains.

This is the part of the story that makes what happens in EPR-like experiments weird. If particles 1 and 2 do not have their respective spin states ever since the no-spin particle devolved, then how do we explain the correlation between them? Beyond the mysterious influence measuring particle 1 has on the spin state of particle 1, we now have to explain the influence measuring particle 1 has on the spin state of particle 2. Furthermore, the distance between the particles when we measure them is arbitrary, which is to say that they could be light-years apart from each other, yet instantaneously have an effect on one another. It implies what Einstein famously called "spooky action-at-a-distance". That is to say, if it is the case that there is a direct causal relation between the two events, then something must be travelling faster than light, which is in contradiction with relativity theory. For Einstein, Podolsky and Rosen action-at-a-distance could obviously not be true, thus they suggested that quantum theory must be incomplete. However, Bell showed that the possible incompleteness of quantum theory is no evasion to the problem posed by EPR-like experiments.[49]

[49]Bell 1966

Bell Phenomena

Bell's theorem purports to show that no theory that is compatible with quantum phenomena can be directly local.[50] So no theory that is compatible with quantum phenomena can avoid some form of action-at-a-distance. Just like Bohm did for the EPR experiments, Daniel Greenberger, Michael Horn and Anton Zeilinger made a conceptually simpler version of Bell's experiment by casting it in terms of spin.[51][52] Imagine that you have three particles, each going into a different direction, and that we measure each one at arbitrary distances from each other. Each particle goes through either a z-oriented or an x-oriented magnet, so we have eight possible experimental configurations.[53] And now out of those eight possible experimental configurations we are interested in four of them. Only the ones where either two of the magnets are z-oriented and the third is x-oriented or those where all of them are x-oriented. If we have those kind of configurations we would notice the following unbroken regularities. Whenever all magnets are x-oriented, an odd number of the particles are deflected in the up direction. However, in all the other situations, so when two of the magnets are z-oriented, an even number of particles are deflected up.[54] These are the frequencies we see experimentally. The first thing to note is that knowing two outcomes means we can deduce the third. So, either the third outcome is dependent on the first two, which would imply superluminal causation, or there must be something in the interaction of the particles with the measurement apparatus that causes these correlations. We can prove, however, that the second option can not be right.[55]

Imagine our three particles en route to their respective Stern-Gerlach apparatus. If the second option is to be right, we need to be able to stipulate some configuration of how each electron responds to each possible orientation of their Stern-Gerlach apparatus, so that we get our experimentally confirmed frequencies. Figure 3.5 maps all possible experimental

[50]Bell 1964
[51]Maudlin 2019, p. 29
[52]Greenberger, Horne, and Zeilinger 1989
[53]Maudlin 2019, p. 30
[54]Maudlin 2019, p. 31
[55]Maudlin 2019, p. 31

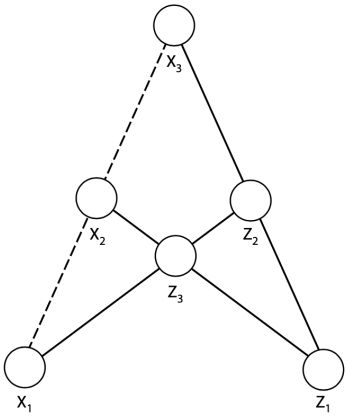


Figure 3.5: Possible Experimental Setups[56]

set-ups. X1X2X3 needs to give either one or three spin-up states. If all three are spin-up, then either there is at least one set-up involving two z-oriented Stern-Gerlach apparatus where all three particles are spin-up, or there is at least one such set-up where there is only one particle which is spin-up, both possibilities do not agree with our measured frequencies. On the other hand, if in the X1X2X3 set-up only one particle is spin-up, then we get a similar contradictory result: there is either a set-up involving two z-oriented Stern-Gerlach apparatus that has three particles spin-up, or there is such a set-up where there is only one particle that is spin-up, or there is a set-up where there are no particles spin-up whatsoever. All three possibilities contradict our measured frequencies, i.e. contradict quantum phenomena. That is to say that we can not stipulate some configuration that agrees with our quantum phenomena, thus it can not be that there is something in the interaction of the particles with their respective measurement apparatus that causes the correlations.[57]

This is what we call Bell phenomena: EPR-like experiments that seem to show that we can not have local theories that account for quantum phenomena. Our next question is, do causal theories of scientific explanation characterize the Copenhagen descriptions of such phenomena as explanations?

[57]Maudlin 2019, p. 31

3.2 Common Causes and Bell phenomena

One way to try to solve spooky action-at-a-distance, is to say that what happens at the two measurements is simply determined by a common cause to both. But, as we shall see, we can cast Bell's theorem in causal terms to show that the import of the theorem is exactly that this is impossible. To do so, I will first explain a formal account by Hans Reichenbach of the intuition of common causes, after which I will give a critique of applying Reichenbach's theory to Bell phenomena as given by Bas van Fraassen in the article "The charybdis of realism: Epistemological implications of Bell's inequality". [58][59]

Reichenbach argued for the principle of the common cause: whenever two events A and B do not stand in a causal relation to each other yet are positively correlated ($p(A \mid B) > p(A)$ and $p(B \mid A) > p(B)$), there must be a third event C which is the cause to them both, i.e. their common cause, which screens off the correlation.[60] Which is to say an event C such that $p(A \mid C \wedge B) = p(A \mid C)$ and $p(B \mid C \wedge A) = p(B \mid C)$. If we apply this principle to the EPR experiment, we can say that the spin state of particle 2 is not caused by the spin state of particle 1, but has a common cause.[61] We need to be careful, however, because although C screening off two correlated events is a necessary condition for C to be a common cause, it is not a sufficient condition.[62] This is easy enough to see if we realize that an effect E of C, otherwise unrelated to A and B yet prior to them, will screen off both events.[63] We would not call such an event a common cause, yet it screens off the correlation, so there must be more to it. I will not go into more detail however, because what *is* established is that screening off is a *necessary* condition for being a common cause and, as we shall see, that is what gives the trouble.

Bell's theory shows us why an explanation via a common cause can not be the kind of explanation we are looking for, as argued by Bas van Fraassen. Let us return to the EPR-

[58] Reichenbach 1971
[59] Van Fraassen 1982
[60] Reichenbach 1971, pp. 157–159
[61] Van Fraassen 1982, pp. 27–28
[62] Suárez 2007, p. 8
[63] Suárez 2007, p. 8

Bohm experiment. Assume that the two Stern-Gerlach apparatus are not necessarily oriented in the same direction, and refer to their orientation setting by 'a' and 'b' for respectively particle 1 and 2. Call the actual measurement events s_1 (for particle 1) and s_2 (for particle 2). Van Fraassen starts by analyzing Bell's idea of "local causality" in terms of three distinct assumptions: "Causality", "Hidden Locality" and "Hidden Autonomy".[64]

1. *Causality*: $p(s_1 \mid s_2 \wedge a \wedge b \wedge \Psi) = p(s_1 \mid a \wedge b \wedge \Psi)$
2. *Hidden Locality*: $p(s_1 \mid a \wedge b \wedge \Psi) = p(s_1 \mid a \wedge \Psi)$
3. *Hidden Autonomy*: $p(\Psi \mid a \wedge b) = p(\Psi)$

And vice versa for particle 2. Causality then just means that the result of s_1 is screened off from s_2 by a, b and Ψ. Similarly, Hidden Locality would mean that s_1 is screened off from b by a and Ψ. Finally, hidden autonomy means that the Ψ is independent from a and b. I will not discuss *Hidden Autonomy* in depth here, because *Causality* and *Hidden Locality* are enough to establish the Bell Theorem and Van Fraassen's argument against a common cause explanation of EPR-like experiments.[65]

Now, remember that screening off is a necessary condition for something to be a common cause, which is to say that if Causality is false then $a \wedge b \wedge \Psi$ cannot be a common cause for s_1 and s_2. For Hidden Locality something similar is going on, but we have to keep in mind that the version for particle 2 would be $p(s_2 \mid a \wedge b \wedge \Psi) = p(s_2 \mid b \wedge \Psi)$, so that Hidden Locality's falsity would bar $a \wedge \Psi$ from being a common cause for s_1 and b, and $b \wedge \Psi$ from being a common cause for s_2 and a.

Let us state the so called "factorizability" condition at the heart of the Bell Inequalities as follows:

$$p(s_1 \wedge s_2 \mid a \wedge b \wedge \Psi) = p(s_1 \mid a \wedge \Psi)\ X\ p(s_2 \mid b \wedge \Psi)$$

Van Fraassen say this condition has been experimentally proven to be false.[66] The crux of the argument is then the following: if both Causality and Hidden Locality hold, then the

[64] Van Fraassen 1982, p. 31
[65] Suárez 2007
[66] Van Fraassen 1982, pp. 34–35

factorizability condition holds as well. This can not be true, so one of the two conditions must be false. Causality then must be the condition that is false according to Van Fraassen, because it contradicts the EPR correlations.[67]

First, we need to prove that the conjoined truth of Causality and Hidden Locality indeed implies factorizability. Both the proof and this interpretation of Van Fraassen's argument is due to Mauricio Suàrez.[68] First, note that:

$$p(s_1 \wedge s_2 \mid a \wedge b \wedge \Psi) = p(s_1 \mid s_2 \wedge b \wedge a \wedge \Psi)\ X\ p(s_2 \mid a \wedge b \wedge \Psi)$$

Then, from the assumed truth of Causality, it follows that:

$$p(s_1 \wedge s_2 \mid a \wedge b \wedge \Psi) = p(s_1 \mid b \wedge a \wedge \Psi)\ X\ p(s_2 \mid a \wedge b \wedge \Psi)$$

From Hidden locality it follows that we have factorizability:

$$p(s_1 \wedge s_2 \mid a \wedge b \wedge \Psi) = p(s_1 \mid a \wedge \Psi)\ X\ p(s_2 \mid b \wedge \Psi)$$

So, either Causality or Hidden Locality must be false and Van Fraassen argues it is Causality.[69] If we imagine that both Stern-Gerlach are orientated in the same direction, i.e. $a = b = \theta$, then Causality reduces to:

$$p(s_1 \mid s_2 \wedge \Psi) = p(s_1 \mid \Psi)$$

$$p(s_2 \mid s_1 \wedge \Psi) = p(s_2 \mid \Psi)$$

However, quantum mechanics gives us:

$$p(s_1 \mid s_2 \wedge \Psi) = 1$$

$$p(s_2 \mid s_1 \wedge \Psi) = 1$$

$$p(s_1 \mid \Psi) = \tfrac{1}{2}$$

$$p(s_2 \mid \Psi) = \tfrac{1}{2}$$

[67] Suárez 2007, p. 17
[68] Suárez 2007
[69] Van Fraassen 1982, p. 35

so:

$$p(s_1 \mid s_2 \wedge \Psi) \neq p(s_1 \mid \Psi)$$

$$p(s_2 \mid s_1 \wedge \Psi) \neq p(s_2 \mid \Psi)$$

And similarly for s_2. So, Van Fraassen has shown that Bell's theorem shows us that we can not give a common cause explanation for the correlation between s_1 and s_2.

If we cannot explain the correlation through a common cause, why not just bite the bullet and say that the two events *are* causally related? In the coming, sections I will discuss how the Copenhagen Interpretation fares in dealing with Bell phenomena according to Interventionism and CQTC.

3.3 Copenhagen Interpretation and Interventionism

3.3.1 Impossible Interventions

Woodward has said relatively little about quantum mechanics and how it relates to causal explanations.[70] Yet little is not nothing; he pays some attention to it in his joint paper with Daniel M. Hausman.[71] There, they say that according to quantum mechanics it is *impossible* to intervene in Bell phenomena. This in turn means that, since there is no possible intervention, Bell phenomena cannot be causally explained in the interventionist account of causal explanations.[72]

Hausman and Woodward give two arguments for the impossibility of intervention in Bell phenomena.[73] The first argument I will quote extensively because it seems contradictory with their account of possibility thus far.

> First, as a number of writers have observed, once the orientation of the measurement apparatus is determined, there is no physically possible way to alter or fix

[70] Shrapnel 2014
[71] Hausman and Woodward 1999
[72] Hausman and Woodward 1999, p. 537
[73] Hausman and Woodward 1999, p. 565

> the value of either measurement result. Indeed it is not even possible to intervene to change separately the probability distribution of either measurement. The experimenter can manipulate the orientation of the measurement apparatus, but neither she (nor any possible causal process) can, for a given orientation of the apparatus, manipulate the measurement outcome or its probability distribution. So it is impossible to produce a change in one measurement outcome by literally setting the other.[74]

The only sense of possibility that seems to be invoked here are whether a human could execute the intervention and whether the intervention is physically possible. That is to say, precisely those kinds of possibility which hitherto have explicitly been excluded from having bearing on whether an intervention is possible in the sense relevant to causation. Why, then, these kinds of impossibility should convince us Bell phenomena as taken by the Copenhagen interpretation can not be explained by interventionism is unclear.

Their second argument is that there are no interventions on Bell phenomena that satisfy the second condition Woodward's account of causal explanations posits:

> I causes B through A, and only through A. There is no other route through which I causes B.

Hausman and Woodward claim that any intervention on a Bell phenomenon is not well defined in this sense, because there is no distinction to be made between intervening on the measurement of particle 1 and acting directly on both particle 1 and 2.[75] They say that the particles should not be thought of as separate objects. Rather, the two spatially separated particles ought to be thought of as a single object, and the measurement results should not be thought of as two separate events, but rather as a single, albeit non-local, event.[76] The reason they are taken to be one single object, is that, since they are entangled, they are described by a single ψ-function and if we take, as we do in the Copenhagen interpretation, the ψ-function

[74]Hausman and Woodward 1999, p. 565
[75]Hausman and Woodward 1999, p. 566
[76]Hausman and Woodward 1999, p. 566

to give complete information, we should thus view the particles as one object.[77] Einstein summed up this position, which he attributed to Bohr, as follows:

> If the partial systems A and B form a total system which is described by its ψ-function $\psi/(\mathrm{AB})$, there is no reason why any mutually independent existence (state of reality) should be ascribed to the partial systems A and B viewed separately, not even if the partial systems are spatially separated from each other at the particular time under consideration. The assertion that, in this latter case, the real situation of B could not be (directly) influenced by any measurement taken on A is, therefore, within the framework of quantum theory, unfounded and (as the paradox shows) unacceptable.[78]

It seems that this representation of Bohr's ideas comes from a certain interpretation of his Quantum Postulate, which "attributes to any atomic process an essential discontinuity, or rather individuality, completely foreign to the classical theories and symbolised by Planck's quantum of action."[79] We can read this as saying that any quantized property that comes in multiples of $\hbar$ should be taken as an object, no matter whether we can localize it.[80] So, in this reading, because of the "the impossibility of subdividing the individual quantum effects" we ought to view such inseparable effects as single objects.[81] It must, at this point, be said that interpreting Bohr as such is a stretch, and, as far as I know, there is no place where it is explicitly stated that we ought to view entangled particles as one. Not even Einstein's rendition of the argument necessitates us taking the particles as one single object, but even if we were to accept that interpretation, it seems this position does not necessarily imply that the measurement results should also be thought of as a single event. I imagine the argument for that conclusion would focus on the idea that if we can not view the particles as different objects, then "the distinction between intervening with respect to X and acting directly on

[77] Skyrms 1984, p. 247
[78] Schilpp 1951, p. 681
[79] Bohr et al. 1928, p. 580
[80] Falkenburg 2007, p. 219
[81] Bohr 1948

both X and Y cannot be drawn," where X and Y represent our two particles.[82]

However, it is still unclear how the fact that you must view the two particles as one object, means that you can not see an intervention on one part of that object as having an effect, through that specific part of the object, on the other part of the object. It is unclear why their being one object implies that the measurements must be taken as one event. Let me give a common sense example: if I were to give a sufficiently hard shake to one end of a long chain, the links at the end of the chain would move, and arguable a chain can be viewed as a single composite object. It is, however, perfectly intuitive to say that the movement of each link causes the movement of the next, despite them being parts of one object. Call one link in the chain A and the next B, me stepping on link A would seem to be a proper intervention, as it becomes the only cause of A not moving, only stops B from moving through A and has no common cause of stepping on A and B. So, even if we concede much that is doubtful, it is still not clear why this argument invalidates the idea of an intervention on particle 1 having an effect on particle 2 through 1.

Nonetheless, looking more closely at the kind of relation that the Copenhagen interpretation claims binds the particles in objecthood might give support for Hausman and Woodward's position. The key phrase here is *mutually independent existence*, which the particles are said to lack.[83] On one reading of this term, you can say that to have mutually dependent existence, means that one can only exist if the other does, and thus we are talking about a significantly stronger relation between parts that make up a whole when compared to links in a chain. In all other real objects that constitute a whole, we can break their connection, without destroying the objects themselves.[84] Their existence is connected, but not conditional on each other. Still, it does not follow that any events on one part of the

[82]Hausman and Woodward 1999, p. 566

[83]Schilpp 1951, p. 681

[84]Without delving to deep into an ontology of objects, it is worth noting that if you have a broad definition of what an object is, you do have examples where their existence is conditional on each other. So, say you are a Platonist about mathematical objects, then those objects are what they are in virtue of the place they take in a structure, so the destruction of their relationships, i.e. their place in the structure, would also destroy the objects themselves. The word *real* here is meant to denote physical objects. So for instance, if you have an employer and employee, separating their relationship will mean they are no longer employer and employee, but the physical objects, the people, that were playing those roles persist.

object must necessarily be taken as an event happening to the whole. Indeed, the sentence "the assertion that, in this latter case, the real situation of B could not be (directly) influenced by any measurement taken on A is, therefore, within the framework of quantum theory, unfounded and (as the paradox shows) unacceptable," points directly against this, as it makes no sense to speak of the influence of a measurement taken of A on the real situation of B if any event involving A is one and the same event involving the whole of A and B.

Moreover, we can interpret mutual dependent existence differently. We might also simply take it to mean that something happening to A *always* has an influence on B and vice versa. This interpretation is supported by the idea that existence here is meant to denote *state of reality*, which, although again suspect to disagreement, suggests that the point is not that *whether* the objects exist is dependent on each other, but that the *kind* of existence, the state of existence, they are in is. If this is the right way to speak about this position concerning the objecthood of entangled particles, then that does not mean that an event on A is equal to an event on B.

So, the idea that a measurement on a particle in a Bell state can not be seen as a separate event from the measurement result of the other particle does not follow from the particles constituting one single object. The idea comes from Skyrms' 1984 paper, but he offers it as a way for counterfactualists about causality to avoid being committed to a causal relation between particles in Bell states. In a footnote, he says "the result could be avoided if the underlying theory of events does not recognize the measurement results down on left; up on right as distinct events but regards them rather as aspects of a single, indivisible nonlocal event - the collapse of the wave packet," but it does not say that this is actually the case.[85]

Let us go back to the core of the assertion made by Hausmann and Woodward. Namely, with respect to particles X and Y in Bell states "the distinction between intervening with respect to X and acting directly on both X and Y cannot be drawn."[86] The minimal interpretation of the particles joint objecthood is biconditional influence. If something hap-

[85]Skyrms 1984, p. 255
[86]Hausman and Woodward 1999, p. 566

pens to X, something happens to Y, and vice versa. The question that needs answering then seems to be when we can distinguish between intervening on one entity in a state of biconditional influence with another and intervening on both entities. I argue that we can only do so by having a clear idea of *how* an intervention on one side of the biconditional affects the other side. If there is no explanation of the effects an event on the one entity has on the other, there is no reason to assume they are distinct at all. This comes down to the final sense of possible intervention mentioned in the previous chapter: interventions can be said to be possible only if there is a theory that "provides the resources for representing them and predicting their consequences."[87]

The question is whether the Copenhagen interpretation provides such resources for Bell phenomena, and the answer is that it does not seem to. If we take measurement to be the intervention, then the Copenhagen interpretation is mum on the consequence of a measurement of X on Y, it only tells us that Y's spin state becomes definite.

However, perhaps a measurement is not the correct kind of intervention to focus on. A minimum sense of intervention suggested by Woodward is Lewis' small miracle.[88] Simply put, a small miracle is the minimal change, not bound by physical laws, of some value which we are considering as cause.[89] If such a miracle satisfies the constraints on intervention put forward in Chapter 2, then it counts as a possible intervention.[90]

So, say our two entangled particles are on their way, and one of them gets measured in some direction and has spin up, giving us certainty the other particle will have spin down if measured in that same direction. But now, imagine a small miracle occurs and instead of spin up, the first particle suddenly has spin down. Does the Copenhagen interpretation tell us what the consequences of this would be? It is tempting to say yes, because since the particles are entangled we now know that upon measurement our second particle will have spin up, so the Copenhagen interpretation tells us what the consequences of the miracle are.

[87]Hausman and Woodward 1999, p. 537
[88]Woodward 2005, p. 135
[89]D. Lewis 1979, p. 468
[90]Woodward 2005, p. 142

I argue, however, that this does not take the miracle as an intervention, but rather as a case where things happen to have been different all along. To take the miracle seriously, we need to interpret it as a change in the world we are considering, and precisely because the Copenhagen interpretation tells us nothing of *how* the correlation between our particles come to be, it lacks the resources to predict the consequences of the small miracle, if we take it seriously *as a change*.

So, it seems an intervention on Bell phenomena can not be well-formed in the Copenhagen interpretation because the Copenhagen interpretation lacks the theory to describe the effects of intervention.

3.4 The Copenhagen Interpretation and CQTC

Dowe developed CQTC specifically with quantum mechanics in mind, but he was thinking of retro-causal interpretations rather than the Copenhagen interpretation, so it is not surprising that there are incongruities between CQTC and the Copenhagen interpretation.[91] However, seeing where things chafe will give us a map to the issues in matching CQTC to quantum interpretaions..

I will argue there are four issues that problematize characterizing the descriptions of Bell phenomena in the Copenhagen interpretation as explanations according to CQTC. First, in the Copenhagen interpretation, there is no causal connection between the measurement of a particle on one side of a Bell phenomenon and the particle on the other side, so applying CQTC to the Copenhagen interpretation of quantum mechanics does not explain Bell phenomena. Second, Dowe defines a causal process using the notion of objecthood, but there is a tension between what Dowe sees as objects and, as we have seen in the previous section, what counts as an object in the Copenhagen interpretation. Third, we need to defend how to use Dowe's concept of causal interaction in a context where we do not always have a defined position value.[92] Fourth, the notion of a conserved quantity and the role it

[91] Dowe 2000, p. 176
[92] Sprickerhof 2013, p. 64

plays in CQ does not translate well into quantum mechanics.

We can reasonably expect some of these issues to disappear once we shit to a different interpretation of quantum mechanics. For instance, if we look at a retro-causal interpretation of quantum mechanics there will be a causal process between the measurement of A and particle 2. However, it is not obvious that all problems disappear.

3.4.1 Causal Processes

Let us start with the most obvious issue: the lack of a causal connection between particle 1 and 2. Remember, for Dowe there is a causal connection between $q(a)$ and $q'(b)$, where q and q' are conserved quantities possessed by objects a and b, if there is a set of causal processes and interactions satisfying some constraints.[93] Those constraints are irrelevant for the moment; the point is that at minimum there must be *some* set of causal processes/interactions. If there is no such set between two objects, there is no causal connection and thus we can not explain what happens to b in virtue of a.

According to the Copenhagen interpretation, in Bell phenomena we have a number of entangled particles, two for instance, moving in different directions, each without spin (ignoring for the moment whether we see these particles as separate entities). For the moment we can count the trajectory of both particles as causal processes, with open forks pointed forward in time, which at some point come upon SG apparatus, where their spin is measured, thus collapsing the wave-function and giving them spin properties that are correlated according to their Bell state. We do not have a set of processes or interactions pointing back towards the source of the particles, nor is there such a set connecting the particles directly. Thus, if we accept both the Copenhagen interpretation and CQ, then we have *no* explanation for Bell phenomena.

[93] Dowe 2000, p. 172

3.4.2 Objects

However, we might critique this representation because it takes the two particles as distinct objects in CQTC, but, as we have seen, we can argue that the Copenhagen interpretation takes entangled particles as one object. Say, for the sake of argument, that it does. Then the question is which ontology we defer to. Dowe says that an object is "anything found in the ontology of science (such as particles, waves and fields), or common sense (such as chairs, buildings and people)."[94] *The ontology of science* presumes a shared ontology among scientific theories, while different theories might disagree on what counts as an object. When it comes to those disagreements, it is not entirely clear what Dowe expects us to do. If we assume we are meant to defer to the specific theory we are discussing, we should say that entangled particles form a single object, but then the concept of a world line does not make sense, since entangled objects can be spatially separated, which means that they do not have a single location at all times. To see this clearly, let's take a look at a specific example of an entangled pair, and how it was prepared. Say we have a spinless particle that decays into two 1/2 spin particles.[95] Those two 1/2 spin particles are entangled in an anti-correlated way after said decay, but let's say they do still constitute a single object. What does a representation of this object's history on a space-time diagram look like? Up until the moment the particle decays, it's a straightforward world line, but the moment it decays said line would have to fork because now the spatial location of the 'object' is spread out over the two particles. If we take this seriously, we would thus need to amend the concept of a world line to allow it to split. However, there is no immediate reason to think this forking is problematic for the rest of CQTC.

Alternatively, we can interpret "anything in the ontology of science" more literally, and say that anything in the ontology of any scientific theory counts as an object, no matter the theory we are working in. That would preserve particles as objects in a CQTC treatment of the Copenhagen interpretation but would leave us in a strange situation where we have a

[94]Dowe 2000, p. 91
[95]P. J. Lewis 2016, p. 8

direct contradiction between the ontology of our theory of causality and the ontology of our scientific theory.

3.4.3 World Lines and the Uncertainty Principle

Recall that a world line is a collection of points on a space-time diagram that represents the history of an object, by giving the location of said point for each consecutive moment in time. We can question whether we can represent the history of a particle in quantum mechanics using a world line.[96]

Also remember that the uncertainty principle as understood in the Copenhagen interpretation states that there are pairs of values such that when one value is *more* precise, the other is necessarily *less* precise.[97] For example, momentum and position have this relation to each other, and so do spin in one direction and its opposite.

If a world line is a collection of points in a space-time diagram, a causal process is a world line possessing a conserved quantity, and we subscribe to the Copenhagen Interpretation, then we seemingly severely limit the sense in which we can have a causal process involving momentum as a conserved quantity. Since a world line consists of position points, some argue the object involved simply cannot *also* posses momentum according to the uncertainty principle.[98] We could try to escape this problem by decreasing the certainty we expect of the position values that make up a world line, thus then simultaneously say that we can then know the momentum the object possesses up to a particular degree of accuracy, but that would constitute a significant departure from the definition of a causal process as given by Dowe.

Moreover, it would not completely solve the problems of CQTC in relation the the uncertainty principle. Lupher criticises the applicability of CQTC to quantum mechanics by noting the following.[99] In order for CQTC to identify causal interactions, we need to be

[96]Sprickerhof 2013, p. 65
[97]Hilgevoord and Uffink 2023
[98]Lupher 2009
[99]Lupher 2009, p. 77

able to do three things:[100]

1. Identify an object before and after a causal interaction.
2. Verify the intersection between the world lines of two causal processes.
3. Know the values of the conserved quantity before and after an intersection.

Now, say we are interested in the (possible) intersection between two particles, then we are interested in their linear momentum. To establish that a causal interaction has taken place, then Lupher says "the particles' linear momentum must be known just before the collision, the collision must be verified, and the linear momentum must be known for all outgoing particles."[101]

However, this contradicts the uncertainty principle according to Lupher. To know the world lines intersect is to know the two particles occupy the exact same position at some time. But then there is no linear momentum at that moment since there being an exact position means there is no exact momentum.[102]

So, Lupher says "Roughly speaking, the more precisely the positions are measured for incoming particles, which is necessary to verify their intersection (2), the less precisely their linear momenta can be measured, which is necessary to determine if a causal interaction (3) occurred."[103] However, is it true that we must measure the positions of incoming particles in order to verify their intersection? Intuitively, if we adopt an ontological interpretation of the uncertainty principle, then it seems straightforward that we cannot know both momentum and position, since a particle cannot have both simultaneously. However, Heisenberg himself argues this ontological position does not apply to particles in the past. He says "the uncertainty relation does not hold for the past."[104] This is because "if the velocity of the electron is at first known, and the position then exactly measured, the position of the electron for times previous to the position measurement may be calculated. For these past times, $\delta p \delta q$

[100]Lupher 2009, p. 77
[101]Lupher 2009, p. 77
[102]Lupher 2009, pp. 77–78
[103]Lupher 2009, p. 7
[104]W. 1930, p. 15

is smaller than the usual bound."[105] So, in the Copenhagen interpretation the uncertainties in question, are uncertainties about things in the future and present; we cannot know these quantities in tandem *beforehand*.[106]

Now, if this is true, then we can verify the fact that world lines have intersected without measuring the position of two particles in that instant, by retroactively calculating the positions of the particles. So, although on the surface the uncertainty principle seems highly problematic for CQTC, closer inspection shows that this is, on the one hand, dependent on our interpretation of the uncertainty principle and, on the other, no knock-down argument even in ontological interpretations of the uncertainty principle.

3.4.4 Quantum Conserved Quantities

Dowe defines a conserved quantity as any quantity subject to a conservation law, with the intention to create a theory that can change in response to scientific discovery. In classical physics, a conservation law states that the total value of some quantity in a closed system does not change over time. There is a distinction between two kinds of conservation laws, universal and local, where local conservation laws state that not only is a quantity conserved but its concentration only changes by *flowing* from one place to another. Things are, however, different in quantum mechanics. There a conservation law generally expresses that "the probability distribution of the conserved variable over the entire ensemble should not change."[107]

According to Dowe, the claim of CQTC is not about quantities subject to local conservation laws, although he says that it will follow that the conserved quantities are also conserved locally: "The claim is not that certain quantities are locally conserved in an interaction or by the process in the absence of interactions, although that will follow."[108] Rather, the point is to focus on universally conserved quantities that have theoretical underpinning

[105] W. 1930, p. 15
[106] Hilgevoord and Uffink 2023
[107] Aharonov, Popescu, and Rohrlich 2021, p. 3
[108] Dowe 2000, p. 95

for their conservation, to avoid accidentally conserved quantities.

I will argue, however, that it *is* the fact that quantities are locally conserved which makes CQTC an intuitive candidate for a theory of causation, and that thus the fact that conservation laws in quantum mechanics are not conserved, locally or globally, in the classical sense is highly problematic for CQTC.

If the conserved quantities we are talking about are locally conserved, then when the value of such a quantity changes, it must be that it has gone to or come from somewhere local. It is this fact that makes conserved quantities a classically intuitive candidate for the construction of a causal theory because it is this continuation that allows us to connect a change in the value of a conserved quantity possessed by one object to the change in another.

It is this continuity which allows Dowe to make sense of his notion of 'exchange'. Recall that a causal intersection is defined as the intersection of two world lines that involves an 'exchange' of a conserved quantity, and that 'exchange' in turn is defined as an event where at least two processes have a change in the value of their conserved quantities after an intersection. Dowe says

> An *exchange* occurs when at least one incoming, and at least one outgoing process undergoes a change in the value of the conserved quantity, where 'outgoing' and 'incoming' are delineated on the spacetime diagram by the forward and backward light cones, but are essentially interchangeable. The exchange is governed by the conservation law, which guarantees that it is a genuine causal interaction.[109]

If we understand conservation laws globally, what sense does it make to say that *an exchange* is governed by a conservation law? If it is merely global conservation we stipulate, then two causal processes intersecting and changing the values of their conserved quantities might change those values to different degrees, and there is no constraint forcing the change in a conserved value to be exchanged with its immediate surroundings. So, the fact that a quantity is universally conserved does not guarantee causality (unless we take that to mean,

[109]Dowe 2000, p. 90

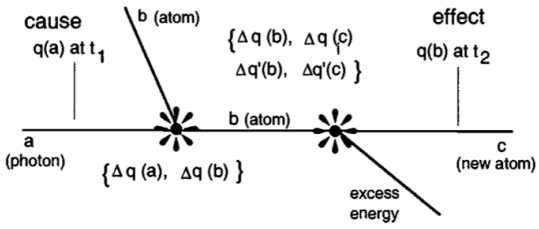


Figure 3.6: Dowe 2000, p. 173

it is guaranteed because it is stipulated). If changes in conserved quantities can be completely unrelated to their immediate surroundings, there is no reason to assume that a change in the value of such a quantity *guarantees* anything about its immediate surroundings. That something is *exchanged* can be guaranteed by the idea that a de/increase in the value of a conserved quantity has to flow to its local surroundings. If not, there is no reason to assume these value changes are necessarily related.

Worse still, if we are considering a quantum mechanical context, an increase in the value of a conserved quantity at one point does not guarantee a decrease of that value at any other point, thus the connection between conserved values and causation becomes even harder to establish. At this point it is worth taking a look at two examples Dowe gives of conserved quantities in a quantum context, and subsequently seeing how such quantities are not necessarily conserved in the sense that Dowe purports to have in mind.

First, Dowe uses a quantum example to illustrate the concept of a causal connection.[110] Say we have an unstable atom that is hit by a photon so that said atom decays into a new atom (see figure 3.4).[111] The exact mechanism of how these processes and interactions form a causal connection is not relevant at this moment. Rather, the point is to not what kind of quantities are taken to be conserved here, which are energy and charge, as Dowe tells us.[112]

The second example requires us to look ahead because Dowe gives it in the context

[110]Dowe 2000, p. 172
[111]Dowe 2000, p. 173
[112]Dowe 2000, p. 173

of Cramer's retro-causal interpretation of Quantum Mechanics.[113] The example in question is the Freedman-Clauser experiment, designed to confirm Bell phenomena, in which two entangled photons are emitted in a cascade of Calcium atoms and whose linear polarization is subsequently measured in a number of directions.[114][115] Cramer's theory suggests that the source of the photons generates two *offer* waves $F(x,t)$ and $G(x,t)$ (waves proliferating into the future), which are constrained *by conservation laws* to be in the same state of polarisation.[116] The conservation law that is constraining these waves here is the law of conservation of momentum. I will explain how this example continues and how it is supposed to explain Bell phenomena in the next chapter. For now, however, this example serves to point out that momentum is another quantity that Dowe takes to be conserved in a quantum context.

The overall point of these two examples is to show that Dowe takes quantum conserved quantities to be the same quantities as those that are classically conserved, and the point is to question if these are actually conserved in the same manner in the Copenhagen interpretation of quantum mechanics.

Before arguing so, it is worth noting that in many situations in quantum mechanics conserved quantities *do* behave the way we would expect them to from a classical perspective. That is, if we have prepared the system in a certain way and measure in a way that is congruent with that preparation. For example, say I have a particle and I measure it in the x-direction and say it comes out as spin-up. If left undisturbed, I can measure it in the x-direction over and over again, and it will turn out spin-up every time.[117] Thus, here spin is conserved in the manner in line with classical mechanics.

Remember that in quantum mechanics a conservation law is taken to say that "the probability distribution of the conserved variable over the entire ensemble should not change,"

[113]Cramer 1986
[114]Dowe 1997, p. 234
[115]Freedman and Clauser 1972
[116]Dowe 1997, p. 236
[117]Susskind and Friedman 2014, p. 7

and that is still the case.[118] The reason for this conceptual change is that "in quantum mechanics where the initial state of a system may not have a well-defined value for the conserved quantity, but some superposition of it, and the outcome of the final measurement cannot, in general, be predicted from the knowledge of the initial state, one cannot directly apply the same concept of "conservation" as in classical physics." What remains to be argued is whether this change is problematic.

I argue it is, because although looking at things statistically is valid, it is equally legitimate to ask what happens in a particular scenario, instead of that totality of possibilities. Especially because the conception of conservation laws that Dowe focuses on does not mention this difference, and in fact treats conservation in classical and quantum physics as equivalent, which they turn out not to be. Causation is not a statistical matter, in the sense that when we talk about it we wonder about particular instances, so if we want to understand causality through conserved quantities, we need to know what happens to those quantities in particular instances. For CQTC an essential aspect of conserved quantities is that whenever there is a change in the value of a conserved quantity in a closed system, that quantity must go somewhere, but as we have seen this is precisely what is not the case for quantum mechanics, because the quantity in question might be in a superposition of values.

Thus we see that conserved quantities like energy are not globally conserved in the Copenhagen interpretation of quantum mechanics, thus CQTC's explanation of Bell Phenomena is inadequate for said interpretation. The concept Dowe uses to define a causal connection does not easily translate to a quantum mechanical context, and especially the description of Bell phenomena offered by the Copenhagen interpretation of quantum mechanics does not qualify as an explanation according to CQTC.

[118] Aharonov, Popescu, and Rohrlich 2021, p. 3

3.5 Conclusion

As we have seen, we have reasons to believe neither theory of causal explanation characterizes the Copenhagen Interpretation's descriptions of Bell phenomena as explanations. What rests is the decision which theory to focus on in light of other interpretations of quantum mechanics. On the one hand, CQTC is confronted with a myriad of conceptual obstacles on its path toward an adequate explanation Bell phenomena; more problems than interventionism. However, there are good reasons to believe that those issues might be resolved in other interpretations. There is a clear sense in which there *is* a causal connection between the particles in retro-causal interpretations of quantum mechanics, other interpretations generally do not take entangled particles as single objects and conserved quantities behave differently in the other interpretations.

That last point is worth lingering on. Let us consider *why* conserved quantities are conceptualized statistically in the Copenhagen interpreting of quantum mechanics, because it will show why CQTC might fare better in alternative interpretations. Say we have a specific quantum state. From that state we can calculate the total energy, and other conserved quantities, of that specific quantum system and the evolution of the wavefunction according to the Schrödinger equation precisely preserves that value.[119] However, upon wavefunction collapse that total quantity collapses into one definite value which might not be the same as the total value of the quantity according to the wavefunction prior to collapse, which suggests that these quantities are not actually conserved at the quantum level. Only if we look at the averages of wavefunction collapse, which thus motivates conceiving of conserved quantities statistically, can we retrieve something of the conserved nature of these quantities. That does mean, as argued before, that in the Copenhagen interpretation conserved quantities can not fulfil the role CQTC needs them to. None of the three alternative interpretations I will discuss in the next chapter postulate wavefunction collapse, thus we have good reason to be hopeful that this trouble is avoided.

[119]Carroll 2020, p. 173

Meanwhile, although I only focused on two issues for interventionism, I am less optimistic that we can evade said issues in other interpretations. Other interpretations might not claim joint objecthood for entangled particles, but it will most likely still be the case that they are in a state of biconditional influence, because that influence does not arise out of an interpretation of the formalism, but out of the formalism itself, which holds equally in all interpretations.

Thus I believe that it is less likely to find a solution to interventionism's qualms in other interpretations, than to find a solution to CQTC's issues using a similar approach, and as such the next and final chapter will pursue CQTC and alternative interpretations of quantum mechanics. Which is not to say that this is the final word on interventionism and quantum mechanics. In fact, there is much substantial work being done on quantum graphs and quantum causality which bases itself on Judea Pearl's work, which is closely related to Woodward's interventionism.[120][121] It merely means that for the reasons mentioned, I estimate our chances higher with CQTC than interventionism, and thus make the choice to focus on CQTC.

[120] Wood and Spekkens 2015

[121] Woodward himself says to have drawn inspiration from Pearl's work, but claims that while Pearl's work is focused on finding interventionist causal relations, his work attempts to give a qualitative definiton of said relations.

Chapter 4

CQTC and Alternative Quantum Interpretations

This final chapter will see whether the descriptions of Bell phenomena in other interpretations of quantum mechanics do count as explanations according to CQTC. In specific, I will look at three interpretations: Cramer's retro-causal interpretation, Bohm's interpretation (also called pilot-wave theory) and the many worlds interpretation.[1][2][3] Cramer's retro-causal interpretation because it is this interpretation that Dowe has in mind when he constructs CQTC. Bohm's interpretation is known under many names, amongst which the Causal Theory of Quantum Mechanics, thus we have good motivation to hope for a resolution in this interpretation.[4] Finally, I look at the Many Worlds interpretation because there is good reason, to be given in the relevant section, to believe it is compatible with Dowe's conception of conserved quantities.

For each interpretation, I will use the following structure. First, introduce the interpretation generally. Then, show how the interpretation presents Bell phenomena. Finally, ask whether this description counts as explanation according to CQTC. Having done that for Retro-causal, Bohm's and Many Worlds interpretation (in that order), we should be able

[1] Cramer 1986
[2] Bohm 1989
[3] Carroll 2020
[4] Riggs 2009, p. 10

to conclude whether or not CQTC is still a promising candidate to explain Bell phenomena.

4.1 Retro-causal Interpretations

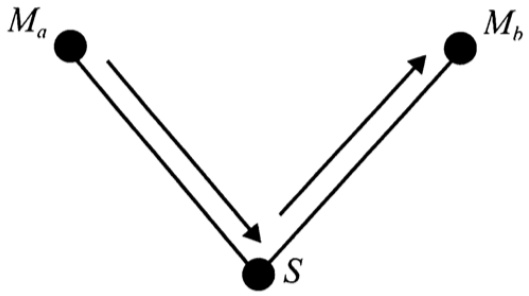


Figure 4.1: Dowe 1996, p. 229

One of the explicit goals of CQTC is to formulate a causal theory of explanation that is mindful of quantum mechanics. That is, to offer an account of causation by which we can offer a causal explanation of quantum mechanics, and of Bell phenomena in particular.[5] Last chapter I argued that CQTC does not do so for the Copenhagen Interpretation. However, the Copenhagen Interpretation is not the interpretation that CQTC attempts to offer a sense of explanation for. Rather it points its arrows at retro-causal interpretations, i.e. interpretations that posit the existence of causal relations that work backwards in time. For Bell phenomena that means, briefly put, that the measurement of particle 1's spin has a backwards in time effect on the origin of the two particles, where it causes (partially) characteristics of both particles, which in turn have an effect on the outcome of the measurement of particle 2's spin.[6] The details of this story are dependent on which specific backwards causality model of quantum mechanics we are looking at, but in its most generic form a retro-causal model of Bell phenomena looks like figure 4.1, where S denotes the source and M_a and M_b the measurements of particles a and b respectively.

[5]Dowe 2000, p. 176
[6]Dowe 2000, p. 185

The specific version of the retro-causal interpretation that I will discuss in this chapter is the 'transactional' interpretation due to Cramer, which Dowe explicitly endorses.[7]

4.1.1 The Transactional Interpretation

Cramer typifies each quantum event as a 'transaction' between waves moving forwards and backwards in time. According to him, there are retarded waves moving into the future, and advanced waves moving into the past.[8] Cramer describes what happens in a quantum event in four stages: 1) emission, 2) response, 3) stochastic choice and 4) repetition to completion.[9]

1. Emission: A source, which Cramer calls "the object supplying the quantities to be transferred", emits an offer wave. The offer wave is referred to in the formalism as ψ, and is a retarded wave, i.e. it is a wave traveling forward in time. It spreads out through space-time until it comes upon an absorber (often more than one), which is characterized as the object receiving the quantities.[10]

2. Response: When the offer wave comes upon an absorber, the absorber responds by producing a confirmation wave. Said wave is an advanced wave, meaning it travels backwards in time and is denoted in the formalism by $\psi*$, i.e. the complex conjugate of ψ.[11] The confirmation wave travels back to the source, arriving with an amplitude of $\psi^*\psi$.[12]

3. Stochastic choice: The source makes a stochastic choice between all the different confirmation waves, selecting one of them. The strength of a confirmation wave amplitude affects the chance that it is selected.[13]

4. Repetition to completion: the source and the absorber that generated the selected

[7]Dowe 2000, p. 186
[8]Dowe 2000, p. 186
[9]Cramer 2015, p. 9
[10]Cramer 2015, p. 9
[11]Cramer 2015, p. 9
[12]Elitzur and Dolev 2006
[13]Elitzur and Dolev 2006

confirmation wave then repeat this process until all the quantities are transferred, at which point the potential quantum event becomes actual.[14]

Now, we can look at what this means for our example of a Bell phenomena, i.e. for two entangled (spin anti-correlated) particles that originate in the decay of another particle. There seems to be no direct treatment of this situation in the literature, but there is discussion of the previously introduced Freedman-Clauser experiment, from which we can understand how Cramer would say our own situation would unfold.[15] The source is the no-spin particle that decays into two 1/2-spin particles. The decay particle, wanting to send out the entangled particles, sends out two offer waves, which propagate forward in time and space until they come upon the two Stern-Gerlach apparatus, the absorbers in this situation.[16] The apparatus then send back confirmation waves, which are the complex conjugates of the offer waves sent out by the source. The transaction between the source and the two absorbers can only be completed if the two chosen confirmation waves respect the conservation of spin, which is what makes it so that the observed anti-correlation is produced as it is.[17] This process is illustrated in figure 4.2.

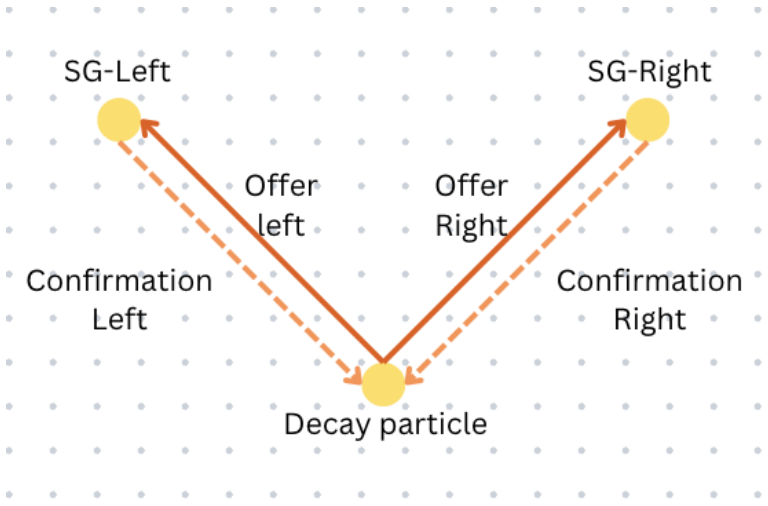


Figure 4.2: Bell phenomenon in the Transactional Interpretation

[14]Cramer 2015, p. 9
[15]Freedman and Clauser 1972
[16]Cramer 2023, p. 378
[17]Cramer 2023, p. 378

This picture departs from the Copenhagen interpretation in a number of significant ways, amongst which ontologically, i.e. which things we consider to be the objects in our theory, and in the way we interpret the uncertainty principle. Moreover, in the motivation for this theory, Cramer discusses conserved quantities, and takes them to behave the same in quantum and classical contexts. The objective of the next section is to see whether these differences enable CQTC to deal with Bell phenomena in the Transactional interpretation of quantum mechanics.

4.1.2 Explaining the Transactional Interpretation of Bell Phenomena

We are asking whether CQTC can establish an explanation of Bell phenomena as represented in the Transactional interpretation. This means asking whether the picture that the Transactional interpretation presents of Bell phenomena forms a causal connection. The answer to this question must decide the final verdict. However, to be able to give this answer, we need to know what counts as a causal process in the Transactional interpretation, which means getting clarity on how concepts that Dowe uses to define a causal process carry over to this interpretation. Recall:

> CQ1. A *causal process* is a world line of an object that possesses a conserved quantity.

The first matter to address is what constitutes an object in TI. We will assume that what is an object for Dowe in a specific scientific context is anything included in the ontology of the theory at hand. There is some ambiguity as to what belongs to the ontology of TI. Specifically, we can question whether particles belong to said ontology or not.

In the first step describing our Bell phenomenon, I said the decay particle sends out offer waves, because it wants to send out the 1/2-spin particles, which would suggest that said particles are considered part of the ontology. In reference to the Freedman-Clauser experiment, Cramer says "the source of the polarization-entangled photons seeks to emit

the photon pair by sending out offer waves ψ_L and ψ_R to the left and right detectors."[18] If photons are things which are meant to be sent out, it seems plausible to say they are part of the theory's ontology.

However, Cramer also argues that waves, which are definitely taken to be part of TI's ontology, are the only real objects in his interpretation.[19] He says:

> The transaction that forms after the emitter-absorber offer-confirmation exchange process goes to completion is the real object, what we would call the "particle" that has been transferred from emitter to absorber. In that sense, the real objects in our universe are waves, while particles are an illusion created by the boundary conditions that must be observed at the vertices of the wave-exchange transactions.[20]

Now, say we adopt the first option and say that both particles and waves are part of TI's ontology. I argue doing so spells problems for a CQTC conception of Bell phenomena because the particles would have to be bereft of conserved quantities. At the end of a quantum event, all conserved quantities must be exchanged according to TI. That is why we see the correlations in entangled 'particles'. However, that would mean that none of these quantities have been exchanged with the particles, because those are not the objects figuring in the exchange, thus the particles would be causally inapt according to CQTC. They are there, but do not appear to instantiate any conserved quantity, and do not actually figure in the causal connection. It is not that including both there is some principle that bars exchanging a conserved quantity between wave and particle, it is just that that exchange does not happen according to the interpretation. The particle has no influence on the story, so in terms of simplicity, it is preferable to leave it out of the story. Having said that, Dowe himself seems to stick to terminology involving particles.[21] In arguing that Bell phenomena in a retro-causal model fit well with process theories of causality, he says "the world line of

[18]Cramer 2023, p. 378
[19]Cramer 2015, p. 8
[20]Cramer 2015, p. 8
[21]Dowe 2000, p. 186

the particle as it moves from the source to the measurement is a process; the world line of the other particle is another process; while the separation at the source is an interaction."[22] However, in light of how Bell phenomena are described in TI, this characterisation seems to be inaccurate.

We thus take waves to be the central objects in the ontology of TI (and whatever objects fullfill the role of source and absorber). The other two concepts from CQ1 that we can question are *world lines* and *conserved quantities*, the first of which is rather straightforward. Waves have world lines, which is simply the collection of points on a space-time diagram that describes their trajectory from source to absorber and vice versa.

The nature of conserved quantities is taken to be the same in TI as in classical physics. In fact, in motivating TI Cramer says the following:

> Nonlocality comes from two seemingly conflicting aspects of the quantum formalism: (1) energy, momentum, and angular momentum, important properties of light and matter, are conserved in all quantum systems, in the sense that, in the absence of external forces and torques, their net values must remain unchanged as the system evolves, while (2) in the wave functions describing quantum systems, as required by Heisenberg's uncertainty principle, the conserved quantities may be indefinite and unspecified and typically can span a large range of possible values.[23]

In TI, conserved quantities thus behave the same way as in classical physics, i.e. they are quantities whose total value does not change over time in a closed system. The quote also hints at two issues that plagued CQTC in the Copenhagen interpretation: the uncertainty principle and indefinite values. Having settled the conceptual states of a causal process in TI, we can now turn towards these issues.

Cramer argues that Heisenberg's uncertainty principle is "a consequence of the fact that a transaction in going to completion is able to project out and localize only one of a

[22] Dowe 2000, p. 186
[23] Cramer 2023, p. 372

pair of conjugate variables (e.g., position or momentum) from the offer wave, and in the process it delocalizes the other member of the pair."[24] Thus a wave can simultaneously instantiate momentum and position, although it can not simultaneously project them.[25] It would thus solve the problem that the Uncertainty Principle poses to CQTC because now there can be an exchange of momentum since we can have a momentum value right before an intersection and right after. We can not, however, know these values, and as such the uncertainty principle remains problematic if we wish to discover causal relations.

Similarly, it seems there is no object possessing an indefinite value of a conserved quantity in TI.[26] To be sure, in the time between the decay of our no spin particle and the measurement of the 1/2-spin particle, said "particles" do not have a definite spin value. However, it is not these particles that are the objects of TI, they merely emerge as a consequence of the behaviour of the offer and confirmation waves.[27] Those waves, in turn, do have a certain value for the conserved quantities they possess, but are in the process of exchanging these between the source and the absorber.[28] As such, just like in the case of the uncertainty principle, we can not ask what the value of an offer wave is exactly, because then we interrupt the process and in fact start another one, but that does not mean that it does not actually have a definite value. Just like in the case of the uncertainty principle, the issue is turned epistemic rather than ontological.

By now we have all the tools to answer whether TI establishes a causal connection between the results of the measurements on the two wings of the experiment. Recall that there is a causal connection between q(a) and q'(b), where q and q' are conserved quantities and a and b objects instantiating them, if and only if:

> There is a set of causal processes and interactions between q(a) and q'(b) such that:

[24]Cramer 2023, p. 373
[25]Cramer 1986, p. 657
[26]Cramer 2015, p. 10
[27]Cramer 2015, p. 8
[28]Dowe 1997, p. 237

1. Any change of object from a to b and any change of conserved quantity from q to q' occur at a causal interaction involving the following changes: $\Delta q(a)$,$\Delta q(b)$,$\Delta q'(a)$, and $\Delta q'(a)$.
2. For any exchange in (1) involving more than one conserved quantity, the changes in quantities are governed by a single law of nature.[29]

This is then the story according to TI. From the source, i.e. the particle decaying, two offer waves propagate upon the decay event. Those offer waves carry the conserved quantity spin towards SGR (the Stern Gerlach apparatus on the right side) and SGL (the Stern Gerlach apparatus on the right side). A wave is an object and possesses a conserved quantity, thus the waves' world lines from the source to SGR or SGL constitute causal processes. The decay is an interaction, because there the source exchanges spin with two other objects (the waves). Those waves then come upon SGL/SGR which send back confirmation waves to the source. This again counts as an interaction, or perhaps two even. We have three objects here: the offer wave, the confirmation wave and SGL or SGR. The offer wave carries spin and exchanges it with SGL/SGR that in turn exchanges it with the confirmation wave. Said confirmation wave then travels back in time towards the source, where again an interaction happens, and the loop starts again. This continues until the quantum boundary conditions are satisfied. In this case, that means that because spin is a conserved quantity, whatever is exchanged with SGL, the opposite must be exchanged with SGR. Thus we get our anti-correlation.

It is worthwhile to make some notes with regard to Dowe's definition of a causal connection and our treatment of Bell phenomena in TI here. First, there is no need to invoke laws of nature here, because these exchanges only involve one conserved quantity. Second, because they only involve one conserved quantity, the only changes any causal interaction requires is $\Delta q(a)$ and $\Delta q(b)$. That is to say, the only changes we need are changes in the spin of two objects.

[29]Dowe 2000, p. 172

If nothing is at fault with this story, we have our answer: according to CQTC, the transactional interpretation explains Bell phenomena.

4.2 Pilot-Wave Theory

The second alternative interpretation we will discuss for CQTC is Pilot-Wave mechanics. Pilot-Wave mechanics is also know as the Causal interpretation of quantum mechanics, thus there is reason to be hopeful that it explains Bell phenomena according to CQTC. Just like the previous section, I will give a brief overview of what Pilot-Wave mechanics says differently about quantum mechanics and how it describes Bell phenomena involving 1/2-spin particles, after which I will argue whether this description indeed counts as an explanation according to CQTC.

4.2.1 Pilot-Wave explained

The starting point of Pilot-Wave Theory (PWT) is the reality of particles.[30] Those particles *always* have definite positions and continuously move through space, i.e. change configuration.[31]

The wavefunction for PWT is real in the sense that there is a real objective quantum space which is represented by ψ.[32] We track the motion of the particles by representing them as points in a configuration space.[33] The configuration space here is ${}^{N}R^{3}$, which is to say the configuration in a three dimensional space of indistinguishable particles. The wavefunction assigns a complex number to each point in that space, thus we can represent the actual configuration using the wavefunction.[34]

We also want to know how those particles move however, and we can find that out as follows. Knowing how the particles move is knowing how their configuration changes,

[30]Maudlin 2019, p. 139
[31]Maudlin 2019, p. 139
[32]Maudlin 2019, p. 140
[33]Maudlin 2019, p. 140
[34]Maudlin 2019, p. 140

which in turn we know if we know the motion of a single point in the configuration space. We get to this motion by looking at velocity. In PWT, we get a velocity field from the wavefunction on our configuration space.[35] A velocity field gives a velocity for each point in our configuration space, thus knowing the velocity field means knowing the motion of a point in the configuration space.[36]

PWT gets this velocity field by taking the gradient of the field represented by the wavefunction ψ, which gives us a vector field where the vectors represent velocity.[37] From here PWT derives a 'guidance' equation which given a wavefunction and a particle with a certain position and mass, gives us the velocity of said particle:

$$\frac{dq_k}{dt} = \frac{\hbar}{m_k} Im(\frac{\nabla_k \psi(x,t)}{\psi(x,t)})$$

q_k represents the position of the kth particle, and as such the left-hand side of the equation represents this particle's velocity. On the right-hand side, $\hbar$ is planc's constant, m_k the particle's mass, $\nabla_k \psi(x,t)$ the gradient of the wavefunction, and $\psi(x,t)$ the wavefunction itself. Im() is meant to communicate that we only take the imaginary part of the result between brackets, so as to ensure our result is not complex. Alternatively, I could have also taken the real part.

In explaining this guidance equation I took some shortcuts. Namely, I took the wavefunction to be a scalar field on the configuration space, i.e. assumed ψ to assign a *single* complex number to each point in the space.[38] Now, this can be the case, and then we call the wavefunction a scalar wavefunction. However, sometimes, namely whenever the wavefunction describes spin, the wavefunction assigns each point in the space *complex spinor* instead.[39] A spinor is "a matrix of two complex numbers $\frac{\alpha}{\beta}$ such that $\mid \alpha \mid^2 + \mid \beta \mid^2 = 1$."[40] This poses a problem, because to get a velocity field we took the gradients of the scalars,

[35]Maudlin 2019, p. 140
[36]Maudlin 2019, p. 140
[37]Maudlin 2019, p. 141
[38]Maudlin 2019, p. 140
[39]Maudlin 2019, p. 140
[40]Maudlin 2019, p. 59

but we can not do this for spinors since they are not scalars.[41]

However, by multiplying the spinor wave function with its complex conjugate we do get a scalar, thus we change the guidance equation to:

$$\frac{dq_k}{dt} = \frac{\hbar}{m_k} Im(\frac{\psi*(x,t)\nabla_k\psi(x,t)}{\psi*(x,t)\psi(x,t)})$$

Note that for any scalar wave function nothing changes, because all we did is multiply numerator and denominator by the same thing. Lastly, the way the wavefunction changes is given by Schrödinger's equation.[42]

All this yields some important differences with the Copenhagen Interpretation. First, there are the ontological commitments of the theory. PWT is staunchly committed to the real existence of both particles and the wavefunction as independent things in the world, whereas the Copenhagen interpretation is ambiguous on the status of ψ.

Second, given an initial wavefunction and configuration, the particles will move in a *definite* way. Where the Copenhagen Interpretation is indeterministic, predicting probabilities, PWT is deterministic.[43]

Third, there is no wavefunction collapse in PWT, where there is such a collapse for the Copenhagen interpretation. According to PWT, the entire physical universe is governed by a single wavefunction, i.e. the universal wavefunction.[44] Now, say that we want to talk about a specific part of the universe. We do not have a separate wavefunction for said part, but we can construct a *conditional* wavefunction that serves our purpose.[45] We do so by plugging in the *actual* locations of the particles we are not interested in.[46]

So, say there are 100 particles in our universe, and we are interested in the first 25, then, denoting actual location with Q, our conditional wavefunction is:

$$\phi_{conditional}(q_1, q_2, ..., q_{25}) = \psi(q_1, q_2, ..., q_{25}, Q_{26}, Q_{27}, ..., Q_{100})$$

[41]Maudlin 2019, p. 148
[42]Maudlin 2019, p. 142
[43]Maudlin 2019, pp. 142–143
[44]Maudlin 2019, p. 150
[45]Maudlin 2019, p. 150
[46]Maudlin 2019, p. 153

What is a collapse of the wavefunction in the Copenhagen interpretation is merely shifting from the universal wavefunction to the conditional wavefunction in PWT. That is not an ontological commitment, there nothing changes. The world is still wholly determined by the universal quantum state, the space-time structure and the particles, which in PWT always have a definite location.[47]

We now know enough about PWT to describe Bell phenomena.

4.2.2 PWT and 1/2 spin Bell phenomena

To see how Bell phenomena involving anti-correlated 1/2 spin particles work in PWT, it is instrumental to first look at a single particle going through a SG apparatus. The essential thing to realize is that for PWT whether an unentangled particle has spin up or down for some orientation, i.e. whether it *deflects* up or down, is wholly determined by its initial location.[48] If the particle enters the SG apparatus above the midline of the magnet, it is deflected up. Conversely, if it enters the SG apparatus below the midline, it is deflected down.[49] This is a consequence of the fact that in PWT trajectories in the configuration space can not cross each other; the ones that go up must have always been positioned above the ones that go down.[50]

So, how does it work for particles that *are* entangled? First, because we are talking about non-relativistic quantum mechanics, one of the two particles must alwasy arrive at the apparatus first.[51] Now, say we have two entangled particles seperated into a right and left direction that will both arrive at their respective SG apparatus above the midline. Call them particle a and b, where a gets measured before b. The spin part of the universal wavefunction for these particles in terms of z-orientation is:[52]

$$\Psi = \frac{1}{\sqrt{2}} \mid z\uparrow\rangle_a \mid right\rangle_a \mid z\downarrow\rangle_b \mid left\rangle_b - \frac{1}{\sqrt{2}} \mid z\downarrow\rangle_a \mid right\rangle_a \mid z\uparrow\rangle_b \mid left\rangle_b$$

[47] Maudlin 2019, p. 154
[48] Maudlin 2019, p. 158
[49] Maudlin 2019, p. 158
[50] Maudlin 2019, p. 158
[51] Maudlin 2019, p. 159
[52] Maudlin 2019, p. 156

Now, if a goes through a z-oriented SG apparatus and is deflected up, and we plug its location into the universal wavefunction, b gets the conditional wavefunction $\mid z\uparrow\rangle_b \mid left\rangle_b$. Similarly, if a is deflected down, b's conditional wavefunction becomes $\mid z\downarrow\rangle_b \mid left\rangle_b$.[53] That is to say, the kind of experiment performed on a and the outcome of said experiment, influences the result of possible experiments on b, even though they're spatiotemporally separated. The question is: how?

When particles are unentangled, the outcome is determined by their initial position, but this can not be the whole story for entangled particles.[54] For instance, say a and b both enter their SG apparatus above the midline. Now, if a does not get measured, is subject to a measurement in another orientation, etc. then b will be deflected up, since it enters the SG apparatus above the midline and is unconstrained by the universal wavefucntion. If a gets measured by a z-orientated SG apparatus, however, b will be deflected down despite entering its SG apparatus above the midline, because a will be deflected up which means b will be constrained by the wavefunction.[55] The point is that PWT is explicitely non-local. The non-locality of the universal wavefunction means that what happens on one wing of the experiment, can influence what happens on the other.

4.2.3 PWT and CQTC

Just like in the section on retro-casual interpretation of quantum mechanics, the central question is whether CQTC says there is a causal connection between our entangled particles such that we can explain Bell phenomena.

I will argue this is not the case for two reasons. First, the world line of the particles in PWT cannot be considered causal processes, because they do not possess conserved quantities, thus they cannot figure in a causal connection. Second, the explicitly non-local effect of measurement of one particle on the results of the measurement of the other is described through the wave function, but that effect of wave function does not satisfy the definition of

[53]Maudlin 2019, p. 157
[54]Albert 1994, p. 157
[55]Maudlin 2019, p. 159

a causal process either, so again it can not figure in a causal connection.

To see this plainly, let us look again at what causal processes, interactions and connections are supposed to be: causal processes are the world lines of objects that posses conserved quantities, causal interactions are intersections of two causal processes where conserved quantities are exchanged, and there is a causal connection between two causal interactions if there is a causal process between them such that there is a conserved quantity that is exchanged in the two interactions and possessed by the causal process.[56][57]

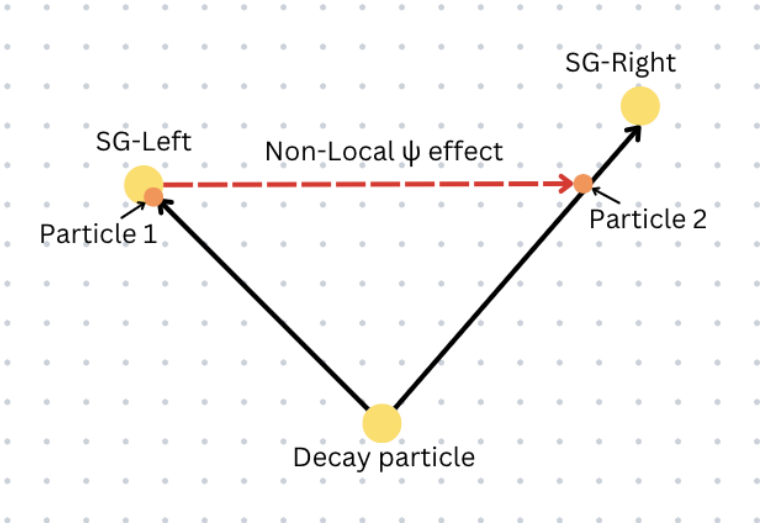


Figure 4.3: Model of PWT Bell Phenomenon

The anti-correlated 1/2 spin Bell phenomenon would have an explanation in PWT according to CQTC if in figure 4.3 the black and red arrows would all count as causal processes, because then there would be a causal connection between the measurement on the left and right sides.

Conserved Quantities

I will first argue that the black arrows in figure 4.3, the arrows representing the world lines of the particles between the source and the measurement, are not in fact causal processes. In order for something to be a causal process it needs to be: a world line (they are) of an object

[56]Dowe 2000, p. 171

[57]A causal connection involving different conserved quantities has a more complicated definition, but because in our Bell phenomenon we are concerned with only one, we can make do with this simpler definition.

(particles are *the* central part of PWT's ontology) that *possesses conserved quantities*. What exactly are conserved quantities in PWT however? The same argument made in Chapter 3 in the context of the Copenhagen Interpretation could possibly be made here, although there are also papers arguing that classically conserved quantities are also strictly (as opposed to statistically) conserved in PWT.[58] Simultaneously, however, there is a different problem connected to conserved quantities for PWT.

PWT lives by a realtor's motto: "Location, location, location!" The idea is that in PWT ψ and the position of a particle (or system of particles) tells you everything, and that other properties *cannot* be pinned down exactly, they are not considered properties intrinsic to the particle (or the wavefunction, or a combination of the two).[59] We can specify position and wavefunction exactly, but all other properties are dependent on these two *and* on how and under which circumstances they are measured.[60] David Albert says that "it turns out that color and gleb and scrad and momentum and energy and every one of the traditional quantum-mechanical observables of particles other than position are contextual properties on this theory," which thus includes many properties which are conserved quantities according to Dowe. The question then is whether we can say a particle *possesses* these quantities, if we accept this contextual characterisation.

Dowe argues that for an object to possess a conserved quantity "it must simply be that the quantity may be truly predicated of the object."[61] So, can we? In a sense, certainly. Just because something is not an intrinsic property and is context dependent, does not mean that it can not be truthfully predicated. That is to say, a property being context dependent means little more than that the property can only be truthfully predicated in certain contexts. The title of record sprinter of my 12th grade class is not intrinsic to me, and is dependent on Barry not running faster than I did, but for the time between me setting the record and Barry breaking it, the property of record sprinter could truthfully be predicated of me.

[58]Potvin 1999
[59]Albert 1994, p. 155
[60]Albert 1994, p. 155
[61]Dowe 2000, p. 92

Similarly, we can truthfully predicate spin as a property of the particles in our Bell phenomenon, even though spin is not intrinsic to the particles. That means that the world lines of these particles must qualify as causal processes, because they satisfy the conditions to possess conserved quantities.

It does, however, bring up further questions for CQTC. Namely, it brings back the trouble of indefinite values. If the only intrinsic property of a particle is its position and wavefunction, how can we then speak of the change in the value of its conserved quantities prior to the particle coming into the proper context? And if we can not speak of that change, how can we then speak of an exchange? The problem thus shifts so to say, because although we can speak of the particles' world lines as causal processes, the class of situations that we can can truthfully call causal interactions shrinks because that class is now limited to the situations in which the particle is in the proper context to speak of contextual properties for a time period extend long enough so that we can clearly say: "For this section the object possessed this value for conserved quantity x, and after its world line intersected with another object's world line, which also possessed a value for said conserved quantity, the value of x changed." Perhaps more troubling still, this return to conserved quantities as indefinite for large periods of time makes their strictly unconserved nature more likely, because what should we call it when a value shifts from definite to indefinite? It is not a loss, because indefinite is not the same as zero. Yet, it is also too easy to assume it is conserved, because the next time it becomes definite, depending on the context, its value may very well be changed.

Location, Location, Location

All this spells troublesome for the marriage between PWT and CQTC . It is not, however, so troublesome in the specific context of the Bell phenomenon we have been discussing. Not so, because of how spin emerges as a property out of position. So let us see if we can CQTC-explain that specific Bell phenomenon in terms of PWT, and if so try to extend that to other scenarios. Take a look at figure 4.4

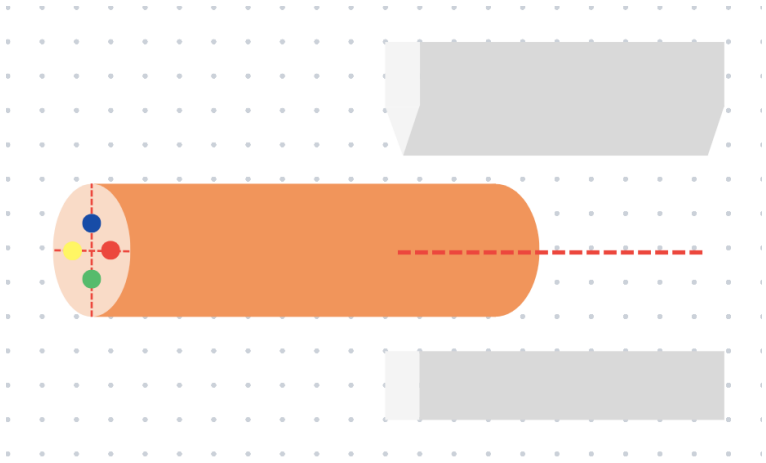


Figure 4.4:

As said before, any property in PWT is emergent from position, wavefunction and context. For a single, unentangled particle we can suffice with position and context. So, in figure 4.4. the blue particle has spin up in the direction of the SG apparatus in virtue of being above the beam's/apparatus' mid-line, and the green has spin down in virtue of being below it. Were we to rotate the apparatus 90 degrees clockwise, then the red particle would have spin up in that direction and the yellow spin down.

Now, when we look at the entangled case, things become slightly more complicated, because the wavefunction starts to play an explicit role and, as we have seen, a particle might be spin down despite being above the mid-line through the non-local interference of the wavefunction. The property of spin, at this moment, is still properly defined however, just not simply in terms of position and context, but in terms of the wavefunction too.

If the way the wavefunction affects the particle can be characterised as a causal process, then we have something approaching an explanation according to CQTC. Say a spinless particle decays into two 1/2 spin particles, both of which are above the mid-line. The particles, call them particle a and b, move in different directions. At this point the wavefunction is mum, because we do not know which particle will be measured first, and in which direction. We do know, however, for each orientation what the result will be for the first measured particle. Thus, vis a vis the orientation of which we said both particles are above the mid-line, both have spin up. Then, however, particle a gets measured and

deflected upwards. We plug its location into the wavefunction, and the wavefunction affects particle b, and from its context, its position and the wavefunction emerges that particle b has spin down. So when it comes to the SG apparatus it is deflected down, thus it seems we have a causal connection between particles a and b, and thus it seems we have an explanation for this Bell phenomenon.

But we do not, because the wavefunction's effect can not play the role it is supposed to. There are two ways to think of ψ here: as a causal process, or as an immediate intermediary between particle a and b.[62]

Say we take that first path. First, there is no sense in which the measurement of particle a constitutes a causal interaction between a and the wavefunction. A causal interaction involves the intersection of two world lines and the exchange of a conserved quantity. It is doubtful whether it will do to call the wavefunction an object at all, despite its real character in PWT. On top of that, even if we grant it objecthood, associating a world line with it is highly questionable. The wavefunction is the structure that makes up the configuration space; it is universal, meaning that it is everywhere so we can not associate a definite position with it. Even if we did, there is no reason to say that position coincides with particle a's position upon deflection.

At this point, describing the connection between particles a and b as a causal process represented by ψ seems lost as an option. However, let us make some wild conjectures and see if we can not salvage something.

Say that the wavefunction has a definite position, namely it has definite position everywhere, so that we immediately also know it intersects with the world line of particle a. It would not improve our position much. The second condition on a causal interaction is the exchange of a conserved quantity, i.e. one object needs to lose some of the quantity, and the other needs to gain some. For the moment, we can ignore that particle a, at least in the way I described it, does not gain spin-up upon measurement/deflection. The point

[62]Immediate intermediary sounds rather contradictory, but it is simply meant to say that ψ is the thing through which particle a's immediate, non-local, effect on b is accomplished.

here is that in no way can we predicate truly of the wavefunction that it has spin, so no exchange of spin can occur, thus no causal interaction. By extension, even if we would grant ψ objecthood, it would still not constitute a causal process simply because it can not posses conserved quantities.

So, we cannot explain Bell phenomena by conceiving of the wavefunction's effect as something akin to a causal process. However, that interpretation does not stay very true to the spirit of PWT anyway. By placing some object between particles a and b, we make the story sound less local because some object is traversing the distance between the two particles, whilst simultaneously grossly contradicting relativity theory because now it is an *object* going faster than the speed of light.

Let us thus attempt to interpret ψ in CQTC more congruent with the spirit of PWT. When particle a is deflected b's spin changes, and *instantly* because of the non-locality of ψ. Meaning that the causal interaction of interest is directly between a and b.

First of all, we have to return to my characterisation of spin in PWT. I said that as the particles approach their SG apparatus, they are both *already* in spin-up, but that means that when particle a *actually* gets deflected, the value of its spin does not change, thus we can not speak of an exchange.

So, say that both particles have no spin during the approach until the moment one of them gets deflected, and then they both do, which would immediately also be more in line with spin being a conserved quantity.

Can we now think of that moment as a causal interaction? No, because there is a locality built into CQTC: the condition on a causal interactions that world lines must intersect. Saying objects can only causally interact when their world lines intersect, i.e. when the objects are in the same location, is the same as saying objects only affect each other locally. It is tempting at this point to say: "Well, if this part of CQTC makes it principally incompatible with PWT, then let us forget about this part." I.e. say that world lines do not need to intersect for there to be a causal interaction. The issue is that Dowe has explicitly *not* given us a way to identify quantities over time, thus if we let go of the locality

condition we lose all ways of connecting the change in the conserved quantity in one object to that in another.[63]

Here we come to the end of the road for PWT and CQTC, with seemingly no more moves to make. We have to concede that according to CQTC, PWT does not explain Bell phenomena, not even in the case of spin Bell phenomena.

4.3 The Many Worlds Interpretation

The departing principle of the Many Worlds Interpretations (MWI) is to take the formalism seriously. That is, MWI is often presented as an effort to take quantum mechanics and interpret it as faithfully to the formalism as possible.[64]

In this section I will first lay out what the bare theoretical minimum *is* according to MWI, and how this implies the splitting of one world into many. Afterwards, I will explain what this means for anti-correlated 1/2 spin Bell phenomena and finally discuss whether this description constitutes an explanation according to CQTC.

MWI says the bare minimum of a quantum theory involves two things:

1. The wave function ψ (or something equivalent).

2. The Schrödinger equation which ψ obeys.[65]

MWI does not posit a collapse of the wavefunction, or particles, or waves moving back into time.[66] Moreover, the wavefunction is, like in PWT, taken to be universal. So, the core of the theory is a universal wavefunction evolving according to the Schrödinger equation without collapsing. How, then, does it explain that it *looks* like things do collapse?

Say we have a particle in spin up state traveling towards a SG apparatus. According to MWI, we need to view measuring devices, in this case the SG apparatus, as quantum

[63]Dowe 2000, p. 42
[64]Carroll 2020, p. 139
[65]Carroll 2020, p. 32
[66]Carroll 2020

systems in and of themselves. So, say the SG apparatus indicates whether it has measured spin-up, down or nothing yet (call this the ready state). Now, before our particle arrives at the apparatus, we can call the state of the particle+apparatus system "particle up, apparatus ready".[67] When the particle has gone through the apparatus, the total system is in the state "particle up, apparatus up". Similarly, if our particle had been spin-down the total system would have evolved from "particle down, apparatus ready" to "particle down, apparatus down".[68] To put it in formal terms, if we express "particle up/down" as $\psi_{\uparrow/\downarrow}$ and apparatus ready/up/down as $\phi_{ready/\uparrow/\downarrow}$, then we say:

$$\psi_{\uparrow} \otimes \phi_{ready} \rightarrow \psi_{\uparrow} \otimes \phi_{\uparrow}$$

$$\psi_{\downarrow} \otimes \phi_{ready} \rightarrow \psi_{\downarrow} \otimes \phi_{\downarrow}$$

Now, because the dynamics of the wavefunction are linear, we know how the combination of these states ought to evolve. We know that if our particle is not in some prepared state, but instead in a superposition of spin up and down, the wavefunction of the particle+apparatus systems evolves like this:

$$(\alpha\psi_{\uparrow} + \beta\psi_{\downarrow}) \otimes \phi_{ready} \rightarrow \alpha\psi_{\uparrow} \otimes \phi_{\uparrow} + \beta\psi_{\downarrow} \otimes \phi_{\downarrow}$$

I.e. the wavefunction evolves in such a way that at the end the measurement apparatus is in a superposition of showing up and down.[69] One face value interpretation of that statement, the one taken by the MWI, is that there are now two measurement apparatus: one that has measured spin up and one that has measured spin down. That seems problematic because we do not actually *see* two apparatus. The essential insight for MWI is that we, too, are quantum systems that evolve in the same way as the measurement apparatus, so that if we present our quantum state as χ, the total system evolves like this:[70]

$$(\alpha\psi_{\uparrow} + \beta\psi_{\downarrow}) \otimes \phi_{ready} \otimes \chi_{ready} \rightarrow \alpha\psi_{\uparrow} \otimes \phi_{\uparrow} \otimes \chi_{\uparrow} + \beta\psi_{\downarrow} \otimes \phi_{\downarrow} \otimes \chi_{\downarrow}$$

[67] Carroll 2020, p. 115
[68] Carroll 2020, p. 115
[69] Wallace 2013, p. 6
[70] Wallace 2013, p. 8

We, then, are in a superposition of having seen the measurement apparatus indicate spin-up and having seen the measurement apparatus indicate spin-down. There is one further thing to note. On a microscopic level we can arguably talk about unentangled particles, but any macroscopic object inevitably interacts and becomes entangled with the totality of its environment, and it is because of this that MWI speaks of many worlds. That is, representing the environment as $\{E_n\}$, the quantum system and its evolution we are concerned with is:

$$(\alpha\psi_\uparrow + \beta\psi_\downarrow) \otimes \phi_{ready} \otimes \chi_{ready} \otimes \{E_{ready}\} \rightarrow$$
$$\alpha(\psi_\uparrow \otimes \phi_\uparrow \otimes \chi_\uparrow \otimes \{E_\uparrow\}) + \beta(\psi_\downarrow \otimes \phi_\downarrow \otimes \chi_\downarrow \otimes \{E_\downarrow\})$$

The totality of the environment, i.e. the world, is thus also in a superposition of two states, and the reason we do not see the two different worlds is that we ourselves can not be separated from this superposition state.[71] Thus wavefunctions seem to collapse when worlds branch, because we are part of one particular branch.

4.3.1 MWI and Bell Phenomena

Now it should be relatively straightforward to describe Bell phenomena in MWI. Remember, the initial state of our entangled particles (call them a and b again) in the z direction is:

$$\tfrac{1}{\sqrt{2}} \mid z\uparrow\rangle_a \mid z\downarrow\rangle_b - \tfrac{1}{\sqrt{2}} \mid z\downarrow\rangle_a \mid z\uparrow\rangle_b$$

Focusing on the particles+apparatus system for clarity, the initial state of that system would thus be:

$$(\tfrac{1}{\sqrt{2}} \mid z\uparrow\rangle_a \mid z\downarrow\rangle_b - \tfrac{1}{\sqrt{2}} \mid z\downarrow\rangle_a \mid z\uparrow\rangle_b) \mid ready\rangle_{Apparatusa} \mid ready\rangle_{Apparatusb}$$

Now that we know how the entangled state evolves, we know by linearity how the particles+measurement system evolves.

[71] Wallace 2013, p. 9

$$(\tfrac{1}{\sqrt{2}} \mid z\uparrow\rangle_a \mid z\downarrow\rangle_b - \tfrac{1}{\sqrt{2}} \mid z\downarrow\rangle_a \mid z\uparrow\rangle_b) \mid ready\rangle_{Apparatusa} \mid ready\rangle_{Apparatusb}$$

$$\rightarrow$$

$$\tfrac{1}{\sqrt{2}}(\mid z\uparrow\rangle_a \mid z\downarrow\rangle_b \mid up\rangle_{Apparatusa} \mid down\rangle_{Apparatusb}) - \tfrac{1}{\sqrt{2}}(\mid z\downarrow\rangle_a \mid z\uparrow\rangle_b \mid down\rangle_{Apparatusa} \mid up\rangle_{Apparatusb})$$

This is to say the wavefunction of the particles+apparatus system evolves into a superposition of particle a having spin-up and particle b having spin-b and both being measured as such, and a having spin-down, b spin-up, and the corresponding measurements. According to MWI, the world branches into two here, so we actually do have one world in which particle a had spin-up and b spin-down and one world in which particle a had spin-down and particle b had spin-up.

The world splits the moment the entangled pair interacts with a macroscopic object, so in the case described here that would be when a particle is measured by a SG-apparatus. According to Sean Carrol, this description of Bell phenomena does away with Einstein's 'spooky action at a distance'.[72] The correlations we see are not caused by some non-local element of nature, rather we see them because of the "branching of the wave function into different worlds, in which correlated things happen."[73]

This description thus suggests the model depicted in figure 4.5. The entangled particles, represented by the green circles, are on their merry way to their respective SG-apparatus, and only upon arriving there the world branches in two, where in one branch particle a is spin-up and particle b spin-down and vice versa in the other branch.

That description, however, does not do away with non-locality, because it does not show why in one branch particle a is spin-up and b is spin-down. If we take figure 4.5 to accurately describe what happens in Bell phenomena according to MWI, the universal wavefunction plays a role which is comparable to its role in PWT, i.e. that of the medium through which non-local effects come to be.

[72] Carroll 2020, p. 105
[73] Carroll 2020, p. 105

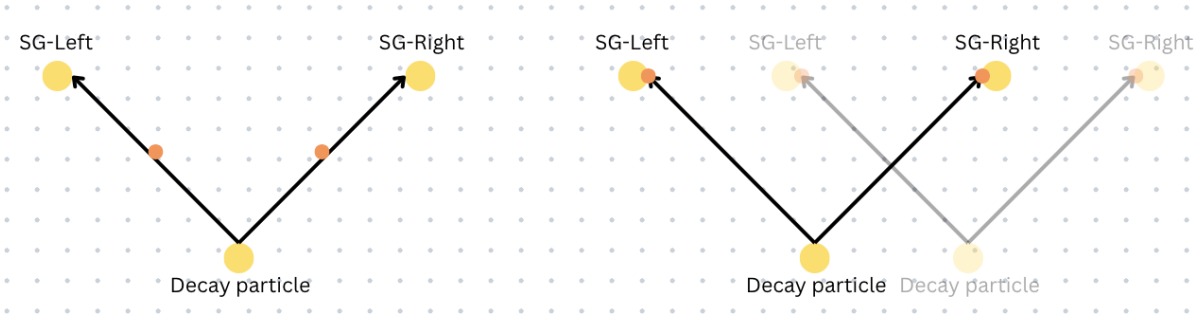


Figure 4.5:

The problem is that this description does not take seriously the superposition of the entangled pair prior to their entanglement with the world at large. The *world* branches once macroscopic objects evolve to be in a superposition, but how do we qualify microscopic objects in superposition? Sean Carrol argues that it too must be taken at face value, meaning that if a particle's wave function indicates it to be in a superposition of two spin-states, *it is* in those two spin-states.[74] The most parsimonious interpretation of superposition of not-with-world-entangled systems in MWI, is to grant them a kind of branching of their own. When a particle is in superposition, there are actually two particles, in effect pre-sorting for the branching of the world they will inhabit. That definition transfers to any quantum system in superposition but unentangled from the world, so that system of the entangled pair of particles in our Bell phenomenon prior to measurement should be said to be two pairs of particles from the very start, where in one pair particle a is spin-up and b spin-down and in one pair particle a is spin-down and b spin-up. Taken as such, we can take Bell phenomena in MWI to be represented by figure 4.6.

This description does seem to properly remove non-locality from the story because there is a kind of branching *prior* to the time when the particles are light years apart.

However, it is good to remember what MWI set out to do: take the formalism seriously. As stated before MWI seeks to be a bare-bones interpretation of the formalism, which at its core only assumes ψ and Schrödinger's equation. That then begs the question where, suddenly, particles have come from. In fact, there are no particles at all in MWI, just the

[74]Carroll 2020, p. 34

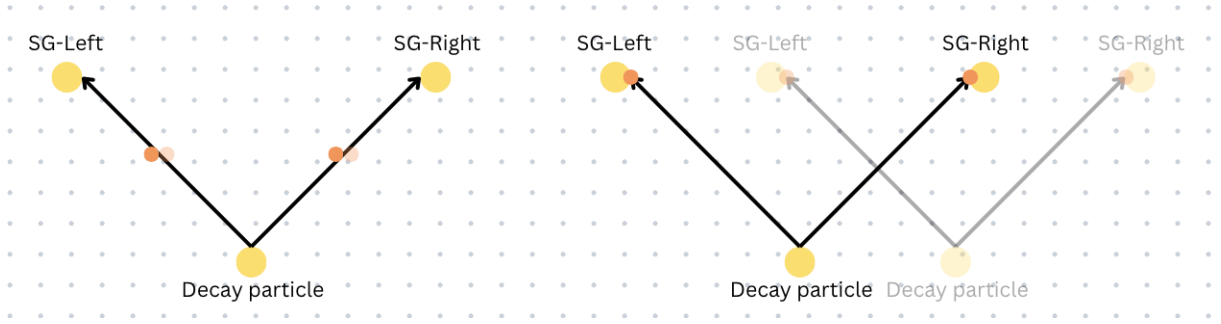


Figure 4.6:

wavefunction that behaves particle-like (sometimes).[75] What the green orbs thus represent are not actually particles, but waves that behave particle-like and can be unentangled from the rest of the wavefunction.[76]

4.3.2 MWI and CQTC

For the fourth and final time, I will now discuss whether this description of Bell phenomena counts as an explanation according to CQTC. I will argue that it does arguing that MWI gives us the tools to handle problems relating to world lines and conserved quantities, which in turn means that the description it gives can be accurately characterized as an explanation according to CQTC.

Objects and World lines

At the end of the last section, I discussed the ontology of MWI: it is all waves. Thus the objects that can figure in an explanation à la CQTC of Bell phenomena, must be waves.

That does not seem to be a problem, as we have seen in the case of Cramer's transactional interpretation, so let us now consider those waves' world lines. Here, things are trickier but still manageable. According to MWI, any object at present has a unique past, but multiple futures. The reason for this isthat the world can branch in the future at which point there is no saying which branch belongs to our present object, but if we take the object

[75] P. J. Lewis 2016
[76] Carroll 2020, p. 101

in one of those worlds, it always traces its path back to our present object.

On the face of it, that sounds problematic because it sounds like our world lines are now forking. We have to be aware, however, that once the world line forks, the objects belonging to each respective fork are not identical. The (practical) solution is to simply take an object's unique history as its world line. The forking brings us no conceptual difficulty since it does not mean *one* object is actually in different positions at the same time, thus the most perspicuous path should be the one we take.

Conserved Quantities

Since there is no wavefunction collapse, MWI avoids the problems the Copenhagen interpretation had with respect to conserved quantities. However, at the face of it MWI has its own conceptual qualms with conserved quantities.

Imagine a billiard ball with a certain mass and potential energy. Then imagine that a previously unentangled-with-the-world particle in superposition with regards to some property interacts with the billiard ball, branching the world in two. In both, branches, the ball maintains the same energy, but from the perspective of the universe at large it seems there are now two balls with said energy, i.e. twice our original energy, thus energy would appear not to be conserved.[77] What is going on?

MWI argues that the key is to consider the 'weight' of the worlds we are talking about.[78] To calculate the total energy (and similarly for other conserved energies) of the universe, MWI tells us to add the energy of ech world times the *weight* (amplitude squared) of that world. So, when a world branches into two, with each branch having an equal amplitude, those two worlds are said to be thinner than the original. To be precise, they are said to both have half the weight of the original world, whatever weight may mean here.[79]

Although this answer is not without its problems (for one, words like weight and thinner seem to be thrown around with quite a bit of hand waving), this story does func-

[77]Carroll 2020, p. 173
[78]Carroll 2020, p. 173
[79]Carroll 2020, p. 174

tion for CQTC. In the sense that it means that in one world, conserved quantities must go somewhere. Even if we are not convinced by the talk of weights of different worlds, that proliferation of energy would not be problematic for the role conserved quantities are supposed to play.

The Causal Connection

The story, then, is this: From the source, a pair of entangled particles depart in different directions. The pair is in superposition of particle a being spin-up and b being spin-down and vice versa, so in a sense there are two pairs departing, one for each combination. The world line of each particle in the pairs, or rather each wave, represents a causal process, because the waves are objects possessing conserved quantities (spin, in this case). Then, each particle arrives at its SG apparatus, at which point these apparatus are also in a superposition of two states, and with them, the world. In one branch we see the left SG apparatus measure spin-up for particle a and the right apparatus measure spin-down for particle b, and vice versa for the other branch. The particles themselves do not change, besides branching off to their respective worlds. The point, thus, is to realize that in the many worlds interpretation the correlations we see *are* caused by a common cause, which can be if we accept the formalism at face value. Thus CQTC is compatible with the many worlds interpretation.

Thus it is the transactional interpretation and the many worlds interpretation that turn out to be compatible with CQTC, whereas pilot wave theory is not. I argued that the description of Bell phenomena in Pilot Wave Theory does not count as an explanation according to CQTC, because the non-local effect between the particles in Bell phenomena does not satisfy the CQTC-defintion of a causal process. More than that, I argued that CQTC is essentially a local theory of causation, because locality is ingrained in the definition of a causal interaction, through the intersection of world lines, thus no theory which includes direct non-locality is likely to be compatible with CQTC. Then, we have discovered and elucidated two obstacles standing between CQTC and theories of quantum mechanics: in chapter 3 I argued that wavefunction collapse is problematic for CQTC, and in chapter 4 I

argued the same for direct non-locality.

Conclusion

The goal of this thesis was to find a match between a theory of scientific explanation and an interpretation of quantum mechanics.

First, the main alternative to causal theories of explanations, unificationism, was considered. However, unificationism was rejected because it was not a sufficient condition for explanation. Instead, it was argued that it functions as an epistemic norm.

Next, I introduced two theories of causal explanation: interventionism and the conserved quantities theory of causation. Interventionism says A is a cause of B if there is a possible intervention on A such that the value of B changes, where possible is understood as logically possible. One of the conditions on such possibility is that there must be a theory which specifies the effects of the intervention. The crucial question in terms of combining interventionism with the Copenhagen interpretation turned out to revolve around that last condition on possibility. I argued that the ontological motivation that Woodward and Hausman introduce to argue against applying interventionism to Bell phenomena was unfounded, but that they were right to say we cannot use interventionism to qualify the Copenhagen description of Bell phenomena as explanations, predominantly because the Copenhagen interpretation lacks the theory to spell out the consequences of intervention.

I ventured our chances with other interpretations were better for CQTC than interventionism and thus focused on CQTC. It is, however, worth emphasizing that this was in no way meant to be a final verdict on interventionism and quantum mechanics. As a matter of fact, there is substantial work being done on quantum causation which is based on Judea Pearl's theory of causation, which is closely related to Woodward's interventionism.

The conserved quantities theory of causation roughly says that causation is fundamentally the exchange of conserved quantities. I argued that on this account, the Copenhagen description of Bell phenomena is no explanation because there is no causal connection between the objects figuring in Bell phenomena in the Copenhagen interpretation and because conserved quantities are not conserved in the same manner in the Copenhagen interpretation.

In the final section of the thesis, I asked whether CQTC might work with other interpretations of quantum mechanics. I argued that the description of Bell phenomena in Pilot Wave Theory does not count as an explanation according to CQTC, because the non-local effect between the particles in Bell phenomena does not satisfy the CQTC-defintion of a causal process. Both the transactional interpretation and the many worlds interpretation descriptions were concluded to count as explanations according to CQTC; they describe causal connections leading the particles to have either spin up or down and none of the obstacles in translating the terms in CQTC to these respective interpretations proved insurmountable.

From these (in)compatibilities we can draw two further conclusions. First, any theory that posits direct non-local effects (i.e. the non-locality does not emerge from some other aspect of the theory like retro-causality or splitting of worlds) is unlikely to be compatible with CQTC because locality is built into the concept of a causal interaction. Second, any theory that posits wavefunction collapse is less likely to be compatible with CQTC, because in the wavefunction and Schrödinger's equation conserved quantities stay conserved. The problems come up because wavefunction collapse interrupts these equations. So, those theories have more difficulty with CQTC.

Non-locality and wavefunction collapse are two of the issues I paid specific attention to when introducing the quantum mechanics for this thesis, and this conclusion warrants their introduction. Furthermore, I introduced the steps to make predictions using quantum mechanics, wave-particle duality, the measurement problem and the uncertainty principle. The 'recipe' for quantum predictions showed us the steps any interpretation must deal with and the measurement problem showed us the logical space of dealing with them, and also showed us where the Copenhagen interpretation arguably lacks detail. Wave-particle duality

was an issue that came up for both interventionism and CQTC, as ontological issues were one of the primary problems at hand. The uncertainty principle has been raised as an objection to CQTC, which I argued it was not necessarily, and the way different interpretations handled indefinite values was an important factor in their fit with CQTC.

It must be said that the interpretations of quantum mechanics as presented here are but one way to present them. Each interpretation comes in many forms, differing in little or a lot, and it could very well be that what works here does not work in a different version of the same interpretation. It must also be said that even if everything said here is correct, there are other issues causal perspectives on these interpretations must deal with. One prominent current issue for instance is that retro-causal theories are subject to fine-tuning.[80] Moreover, there are some more foundational questions about scientific explanations which I have not discussed here, which could bear on this thesis. For instance, you could argue that we do not, in fact, want explanations from science, or even if you agree on this point, there are discussions about ontic versus epistemic explanation that might influence the way you read these arguments, that I have not treated here directly.[81]

Nonetheless, that there are arguably interpretations of quantum mechanics that are compatible with CQTC stands, and, perhaps more importantly, it has been shown that wavefunction collapse and non-locality are the main culprits standing between CQTC and quantum interpretations and the manner in which they do so has been elucidated. Thus the search for a match between causal theories of explanation and interpretations of quantum mechanics has given us deeper insight into quantum causation.

[80] Wood and Spekkens 2015
[81] Wright and Van Eck 2018

Bibliography

Aharonov, Yakir, Sandu Popescu, and Daniel Rohrlich (2021). "On conservation laws in quantum mechanics". In: *Proceedings of the National Academy of Sciences* 118.1, e1921529118.

Albert, David Z (1994). *Quantum mechanics and experience*. Harvard University Press.

Aronson, Jerrold L (1971). "On the Grammar of'Cause'". In: *Synthese*, pp. 414–430.

Bell, John S (1964). "On the einstein podolsky rosen paradox". In: *Physics Physique Fizika* 1.3, p. 195.

— (1966). "On the problem of hidden variables in quantum mechanics". In: *Reviews of Modern physics* 38.3, p. 447.

Bohm, David (1989). *Quantum Theory*. 1st Edition by this publisher. Dover Publications. ISBN: 9780486659695,0486659690. URL: `http://gen.lib.rus.ec/book/index.php?md5=c9c71dc5cb0440f32af2150b6e2776f7`.

Bohr, Niels et al. (1928). *The quantum postulate and the recent development of atomic theory*. Vol. 3. Printed in Great Britain by R. & R. Clarke, Limited.

Bohr, Niels (1948). "On the notions of causality and complementarity 1". In: *Dialectica* 2.3-4, pp. 312–319.

Carroll, Sean (2020). *Something deeply hidden: Quantum worlds and the emergence of spacetime*. Penguin.

Cartwright, N (1983). *How the Laws of Physics Lie Oxford University Press Oxford*.

Cramer, John G (1986). "The transactional interpretation of quantum mechanics". In: *Reviews of Modern Physics* 58.3, pp. 647–687.

Cramer, John G (2015). "The transactional interpretation of quantum mechanics and quantum nonlocality". In: *arXiv preprint arXiv:1503.00039*.

— (2023). "The Quantum handshake explored". In: *arXiv preprint arXiv:2302.06748*.

Dowe, Phil (1992). "An empiricist defence of the causal account of explanation". In: *International Studies in the Philosophy of Science* 6.2, pp. 123–128.

— (1996). "Backwards causation and the direction of causal processes". In: *Mind* 105.418, pp. 227–248.

— (1997). "A defense of backwards in time causation models in quantum mechanics". In: *Synthese* 112, pp. 233–246.

— (2000). "Physical causation". In.

Einstein, Albert, Boris Podolsky, and Nathan Rosen (1935). "Can quantum-mechanical description of physical reality be considered complete?" In: *Physical review* 47.10, p. 777.

Elitzur, Avshalom C and Shahar Dolev (2006). "Multiple Interaction-Free Measurement as a Challenge to the Transactional Interpretation of Quantum Mechanics". In: *AIP Conference Proceedings*. Vol. 863. 1. American Institute of Physics, pp. 27–43.

Fair, David (1979). "Causation and the Flow of Energy". In: *Erkenntnis* 14.3, pp. 219–250.

Falkenburg, Brigitte (2007). *Particle Metaphysics - A Critical Account of Subatomic Reality*. The Frontiers Collection. Springer. ISBN: 3540337318,978-3-540-33731-7.

Freedman, Stuart J and John F Clauser (1972). "Experimental test of local hidden-variable theories". In: *Physical Review Letters* 28.14, p. 938.

Gijsbers, Victor (2007). "Why unification is neither necessary nor sufficient for explanation". In: *Philosophy of Science* 74.4, pp. 481–500.

Greenberger, Daniel M, Michael A Horne, and Anton Zeilinger (1989). "Going beyond Bell's theorem". In: *Bell's theorem, quantum theory and conceptions of the universe*. Springer, pp. 69–72.

Hausman, Daniel M and James Woodward (1999). "Independence, invariance and the causal Markov condition". In: *The British journal for the philosophy of science* 50.4, pp. 521–583.

Heisenberg, Werner (1927). "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik". In: *Zeitschrift für Physik* 43.3-4, pp. 172–198.

Hempel, Carl G and Paul Oppenheim (1948). "Studies in the Logic of Explanation". In: *Philosophy of science* 15.2, pp. 135–175.

Hilgevoord, Jan and Jos Uffink (2023). "The Uncertainty Principle". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Summer 2023. Metaphysics Research Lab, Stanford University.

Illari, Phyllis McKay, Federica Russo, and Jon Williamson (2011). *Causality in the Sciences*. Oxford University Press.

Kitcher, Philip (1981). "Explanatory unification". In: *Philosophy of science* 48.4, pp. 507–531.

— (1989). "Explanatory unification and the causal structure of the world". In.

Kuhn, Thomas (1988). "Objectivity, Value Judgement and Theory". In: *Introductory Readings in the Philosophy of Science, revised ed., eds. ED Klemke, Robert Hollinger and A. David Kline*, pp. 277–291.

Lewis, David (1979). "Counterfactual dependence and time's arrow". In: *Noûs*, pp. 455–476.

Lewis, Peter J (2016). *Quantum ontology: A guide to the metaphysics of quantum mechanics*. Oxford University Press.

Lupher, Tracy (2009). "A physical critique of physical causation". In: *Synthese* 167, pp. 67–80.

Maudlin, Tim (1995). "Three measurement problems". In: *topoi* 14.1, pp. 7–15.

— (2019). *Philosophy of physics: Quantum theory*. Vol. 33. Princeton University Press.

Mermin, N David (1997). "An introduction to space–time diagrams". In: *American Journal of Physics* 65.6, pp. 476–486.

Potvin, Guy (1999). "Dualist interpretation of quantum mechanics". In: *arXiv preprint quant-ph/9908019*.

Reichenbach, H. (1971). *The Direction of Time*. Univ. of Calif. Press.

Reutlinger, Alexander (2017). "Explanation beyond causation? New directions in the philosophy of scientific explanation". In: *Philosophy Compass* 12.2, e12395.

Riggs, Peter J (2009). *Quantum causality: conceptual issues in the causal theory of quantum mechanics*. Springer.

Salmon, Wesley C (1984). *Scientific explanation and the causal structure of the world*. Princeton University Press.

— (2006). *Four decades of scientific explanation*. University of Pittsburgh press.

Schilpp, Paul Arthur (1951). "Albert Einstein; Philosopher, Scientist". In: *Philosophy* 26.99.

Shrapnel, Sally (2014). "Quantum causal explanation: or, why birds fly south". In: *European journal for philosophy of science* 4, pp. 409–423.

Skyrms, Brian (1984). "EPR: Lessons for metaphysics". In: *Midwest Studies in Philosophy* 9.1, pp. 245–255.

Sprickerhof, Jakob (2013). "The Conserved Quantity Theory of Causation and Entangled States". In.

Steup, Matthias and Ram Neta (2020). "Epistemology". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2020. Metaphysics Research Lab, Stanford University.

Suárez, Mauricio (2007). "Causal inference in quantum mechanics: a reassessment". In.

Susskind, Leonard and Art Friedman (2014). *Quantum mechanics: the theoretical minimum*. Basic Books.

Van Fraassen, Bas C (1982). "The charybdis of realism: Epistemological implications of Bell's inequality". In: *Synthese* 52.1, pp. 25–38.

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

Wallace, David (2013). "A prolegomenon to the ontology of the Everett interpretation". In: *The wave function: Essays on the metaphysics of quantum mechanics*, pp. 203–222.

Wood, Christopher J and Robert W Spekkens (2015). "The lesson of causal discovery algorithms for quantum correlations: causal explanations of Bell-inequality violations require fine-tuning". In: *New Journal of Physics* 17.3, p. 033002.

Woodward, James (2005). *Making things happen: A theory of causal explanation*. Oxford university press.

Woodward, James and Lauren Ross (2021). "Scientific Explanation". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2021. Metaphysics Research Lab, Stanford University.

Wright, Cory and Dingmar Van Eck (2018). "Ontic explanation is either ontic or explanatory, but not both". In.