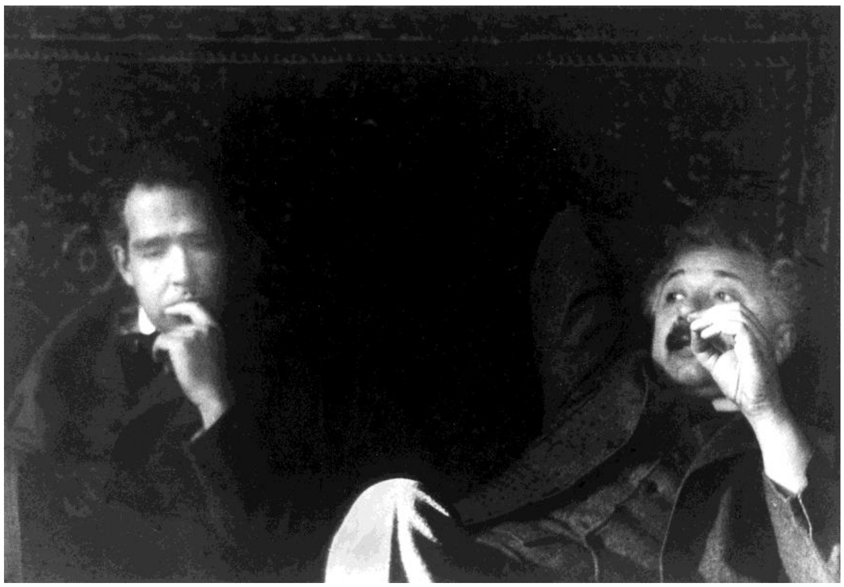
A Comprehensive Evaluation of the Viability of Superdeterminism



Thomas Vissers

s4555724

Supervisor: Prof. dr. H.W. de Regt

Second corrector: Prof. dr. N.P. Landsman

70,607 words, excl. bibliography

28-08-2023

Master Filosofie van de Natuur- en levenswetenschappen

Scriptie ter verkrijging van de graad “Master of arts” in de filosofie

Radboud Universiteit Nijmegen

Hierbij verklaar en verzeker ik, Thomas Vissers, dat deze scriptie zelfstandig door mij is opgesteld, dat geen andere bronnen en hulpmiddelen zijn gebruikt dan die door mij zijn vermeld en dat de passages in het werk waarvan de woordelijke inhoud of betekenis uit andere werken – ook elektronische media – is genomen door bronvermelding als ontlening kenbaar gemaakt worden. Plaats: Nijmegen. Datum: 28-08-2023

"I remember discussions with Bohr which went through many hours till very late at night and ended almost in despair; and when at the end of the discussion I went alone for a walk in the neighboring park I repeated to myself again and again the question: Can nature possibly be so absurd as it seemed to us in these atomic experiments?" (Werner Heisenberg, Physics and Philosophy, 1958)

"This problem of getting the interpretation proved to be rather more difficult than just working out the equations." (Paul Dirac, quoted through A. Pais, 2000, p.55)

Contents

Introduction

Quantum theory is an area of physics describing nature at the smallest and the most fundamental level we know of. The theory has achieved enormous empirical success in describing the microscopic world[1], its mathematical formalism and application being well-understood. The physical interpretation of the theory is, however, far from straightforward. Many quantum phenomena do not have a clear classical analogue, nor do they seem to be fully consistent with prior to its introduction widely held (meta)physical beliefs in physics such as determinism, realism and locality. Now, almost 100 years after the formulation of quantum mechanics, this state of affairs has given rise to many different interpretations. Nevertheless, none are free of criticism, and nothing close to a consensus on the issue has been reached (Schlosshauer, Kofler, & Zeilinger, 2013).

Superdeterminism

Amidst the forest of interpretations, one framework that has recently received increasing attention is called 'superdeterminism' ('t Hooft, The Cellular Automaton Interpretation of Quantum Mechanics, 2016) (Palmer, 2016) (Hossenfelder S. , 2020) (Andreoletti & Vervoort, 2022), including through a recent international conference on the topic (Hossenfelder, Palmer, & Price, Superdeterminism and Retrocausality, 2022). Superdeterminism is a local, deterministic hidden-variable theory whose main point of departure is that it relinquishes statistical independence. This means that in a superdeterministic framework, an underlying but as of yet unknown structure beneath quantum theory determines the outcomes when measuring observable properties of a quantum system like a particle. Moreover, variables introduced with this structure are *correlated with the measurement settings*. While the meaning and implications of this will be investigated extensively later, proponents claim this could answer key foundational questions plaguing quantum theory. It has even been stated that superdeterminism might be the alley through which progress in many important problems in physics outside of quantum foundations will be enabled (Hossenfelder & Palmer, 2020, p. 21) It should be noted that superdeterminism, while offering a clear stance on many of the aforementioned issues within the philosophy of physics, can be argued not to be an interpretation as such. Its empirical

[1] For the sake of completeness, it should also be mentioned that there also exist numerous examples of quantum phenomena manifesting themselves on the macroscopic scale, such as superconductivity and the quantum hall effect. These are also accurately described by quantum theory. Most quantum phenomena, however, are restricted to the (sub)atomic realm, reducing to classical phenomena in the macroscopic realm.

predictions need not be equivalent to those of quantum theory, which may allow for experimental tests against quantum theory. This issue of the conceptual status of superdeterminism shall be clarified later, but due to its frequent association with the subject of 'the interpretations of quantum theory' it has for now been introduced as a member of this class.

As a relatively little researched approach in the interpretation of quantum theory, the substantial promises made by it might be worthy of attention. If it were successful, it could be one of the, if not *the*, biggest development in physics of this century so far. At the same time, superdeterminism is certainly not devoid of critiques, many of which are highly philosophical in nature. Among other things, it has been accused of being heavily fine-tuned, invalidating the scientific method and being incompatible with free will. If such critiques are judged to be legitimate, the attention of physicists and philosophers may be better required elsewhere, while leaving a smaller amount of viable 'interpretations' on the market.

Research question

The stakes for both physics and philosophy in this ongoing debate are thus quite high. Combining my interest and training in both these fields, I therefore aim to critically evaluate superdeterminism to judge whether it holds up to its promises. This leads me to the follow research question for this thesis:

Is superdeterminism a viable theoretical framework for addressing foundational philosophical questions in quantum theory?

A more detailed description of both superdeterminism and the philosophical problems in quantum theory will be provided in the coming chapters. More importantly for now is how the word 'viable' is to be understood, as the answer to the research question strongly hinges upon it. Firstly, the viability of any theoretical framework ought not be confused with its truth-value. It is, more so, about whether the framework is deserving of serious recognition, with it at least conceivably turning out correct. We may consider whether it holds up when compared to its many competitors in the field. Secondly, the viability question need not be a binary one. There may be reasons to suggest one framework is 'more viable' than another. In fact, the very criteria for a viable theory are subject to intense debate in the Philosophy of Science, meaning one would be hard-pressed to find a single, uncontroversial binary question with which to determine this. Therefore, I will evaluate superdeterminism's viability to the end stated in the research question using the following five criteria:

- Can superdeterminism provide clear answers to the philosophical and foundational questions and problems in quantum theory?
- Is superdeterminism self-consistent?
- Is superdeterminism consistent with well-established physical theories in all measurable regimes?
- Does superdeterminism have the prospect of possibly being testable now or in the future?
- Are the metaphysical consequences of superdeterminism acceptable?

I will briefly attempt to motivate the use of these criteria in particular. The first question must be answered affirmatively, because as we will see, the supposed solutions to these problems are what lead some to superdeterminism in the first place. Secondly, out of logical necessity, superdeterminism should not be found to be self-contradictory. Thirdly, if superdeterminism were to be in conflict with well-established physical theory, this would be a clear argument in favor of choosing an alternative way of viewing quantum theory that is not. A viable framework can be reasonably expected not to contradict empirical reality, after all. One exception to this rule is the case where a superdeterministic theory predicts diverging measurement results from those of quantum theory in regimes we cannot (yet) probe. For example, if the difference between a numerical value predicted by the two theories is too small to be measurable with current experimental capabilities, then the proposed model need not conflict with reality. Instead, it would have the (positive) quality of being testable, which brings us to the fourth criterium. An interpretation or theory that allows for the prospect of, at least one day, providing us with a means to empirically determine if it is correct or not, may be argued to be more viable than an alternative that does not. Such a prospect introduces the commonly held epistemic virtue of experimental falsifiability and would prevent the risk of an eternal standoff of, ultimately, metaphysical preferences that may be associated with some interpretations. The last question is explicitly philosophical, but no less important. To a large degree, the debate on the interpretations of quantum theory lies outside the domain of empirical science. Therefore, as will be seen in practice when going over them, viability judgments on the frameworks cannot be fully separated from philosophical commitments. For example, if an interpretation can technically answer foundational problems, but comes at the cost of much controversial ontological baggage, then this may be a reason to judge it as less viable. An example could be that even though the interpretation sometimes referred to as ‘Bohmian mechanics’ is generally argued to solve the soon to be introduced ‘measurement problem’, many reject it on the basis that it adds an ontology perceived as overly complex (Sivasundaram & Nielsen, 2016, p. 13).

A final question to confront considering the criteria of viability is in what way they combine to coherently answer the central research question. Conceivably, some questions may yield a 'yes' while others a 'no', and the acceptability of this for overall viability may vary per criterium. Self-consistency ought to be demanded in any case, but testability is generally accepted to be absent in interpretations like the later to be discussed Everett interpretation. Moreover, there is likely to be plenty of nuance behind these binary answers, particularly in the case of the final criterium which is hard to answer in any objective way. One could, e.g., reasonably ask how much 'baggage' is too much. This should, however, not entirely come as a surprise. If universal agreement could be reached about what set of (possibly hierarchical) criteria should be used to judge the validity of an interpretation of quantum theory, let alone on the metaphysical questions underlying physics, the debate on interpretations would probably not have been raging on for a century. That is why I aim to be explicit and complete with regards to the defined criteria, so that at the end readers can make their own informed call on the issue rather than just having mine.

I will use and evaluate the direct writings of both proponents and opponents of superdeterminism. In addition, more general background literature, primarily on the philosophy of quantum theory, will be used.

Thesis structure

In the remainder of this thesis, I will start with a general and necessary introduction of quantum theory, its development and its philosophical issues in chapter 1 and 2. Chapter 3 will introduce John S. Bell's famous no-go theorem, which restricts a large class of hidden-variable theories and is key to understanding where superdeterminism came from. Next, chapter 4 and 5 will go into what an interpretation of quantum theory is in the first place, as well as lay out some of the well-known competitors in the field. Chapter 6 will then introduce and clarify superdeterminism. The core of the thesis can be found in chapter 7 and especially chapter 8, where the arguments in favor of and in opposition to superdeterminism, will make their appearance, respectively. These will be critically evaluated in order to accurately judge the viability of superdeterminism at the end. Chapter 9 discusses the possible future of superdeterministic approaches, including that of empirical testability. A conclusion with an answer to the research question will follow, to finally end with a discussion including an outlook.

Finally, it should be noted that even though the subject at hand is embedded in an extensive physical and mathematical framework, the text is written with the goal of being accessible to all Master's students of philosophy. Understandability is a major guideline to the overall thesis. To this end, explicit care is taken to explain the necessary quantum physics and to avoid its usual heavy reliance on mathematics, such that no significant background in these fields is required. In the benefit of keeping the overall, rather lengthy story clear, explicit references to the underlying structure and brief reminders on key ideas are occasionally employed.

Chapter 1: A brief history of quantum theory

The debate surrounding the interpretations of quantum theory, and by extension superdeterminism, cannot be understood without a basic grasp of (the development of) quantum theory. In this section, I will aim to provide this in three phases. I will also introduce some physical concepts in so far as these are necessary for the understanding of the debate surrounding philosophical interpretations thereof.

The beginnings of quantum theory can be traced back to the turn of the twentieth century. In the early twentieth century, physics as a discipline underwent two more or less simultaneous revolutions through the formulation of quantum theory and the theory of relativity. The physics before this time is commonly referred to as 'classical physics'. Classical physics accurately describes the behavior of nature in areas such as everyday mechanics, electromagnetism and thermodynamics, but it was found not to be able to do so when dealing with, for example, situations at very small scales, at high speeds or with strong gravity. It is primarily the discrepancies between microscopic phenomena and the predictions of classical physics that sparked the creation of quantum theory.

The old quantum theory

The development of quantum theory can be described in three phases. The first of these is often referred to as the 'old quantum theory', corresponding to the first quarter of the twentieth century. This was not so much a single theory as it was a collection of ideas merging quantum and classical concepts to solve a number of problems physicists at the time were faced with. This started with Max Planck (1858-1947) who was faced with explaining the intensity and wavelength of electromagnetic radiation given a material with a certain temperature. Think, for example, of the yellow glow of a hot iron poker. Planck was ultimately able to derive a mathematical formula relating these three quantities, but he had to do so through the introduction of the 'quantum'. Planck had derived his formula on the condition that energy could be emitted and absorbed only in indivisible discrete packets called 'quanta'. Hence, *quantum* theory. This was in sharp contrast to classical physics, where energy was viewed as a continuous quantity that could be traded in any arbitrary amount. Planck himself did not initially believe these quanta really existed. Rather, he saw them as mathematical instruments. In his view, they would ultimately make place for continuous matter once again (Kumar, 2008, pp. 3-31).

Two more important steps in the formulation of the old quantum theory came from two men whose philosophical ideas will later be discussed extensively, as they were arguably the most important voices in the early debates on the interpretation of quantum theory. They are Albert Einstein (1879-1955) and Niels Bohr (1885-1962).

In 1905, Albert Einstein used Planck's quantum to explain the photoelectric effect. This is the effect whereby if one shines light of a certain frequency onwards on a metal surface, it emits electrons. Where most viewed light as being a wave, Einstein viewed it as being composed of quanta: discrete particle-beams that could knock electrons out of their atomic orbits. While the wave model was unable to explain the photoelectric effect, Einstein's light-quanta could. At the same time, there were also plenty of experiments in which light did seem to be a wave, and where the particle ontology would not hold up. An example of this would be diffraction and interference effects in slit experiments. This demonstrates what would come to be known as the dual nature of matter, a point that shall be returned to later (Kumar, 2008, pp. 31-67).

Niels Bohr was the first to create a model for the atom that combined classical with quantum properties. This atomic model came to be known as the 'Bohr model'. It was a response to various difficulties with earlier atomic models such as those of Joseph Thompson (1856-1940) and Bohr's own mentor Ernest Rutherford (1871-1937), especially concerning issues of radiation. Radiation entails the transmission of energy over large distances, of which there exist several variants such as electromagnetic radiation or alpha particles (helium-nuclei). At the time it was known atoms could emit several kinds of radiation, but which and through what mechanism was not well-understood. By proposing that electrons, orbiting atomic nuclei in circles, could only occupy specific quantized 'orbitals' with specific associated energies, Bohr was able to explain many atomic radiation phenomena. Electrons could 'jump' between orbitals, emitting photons when jumping to a more inward atomic orbital. He could also explain other issues such as the structure of the periodic table of the elements, even accurately predicting the existence and properties of new ones. This model, however, assumed that electrons could only orbit nuclei at certain distances, with others being forbidden. The allowed orbitals could be labelled through the so-called 'principal quantum number', denoted by n. For $n = 1$, an electron was in its lowest energy orbit, closest to the nucleus. The energy and the radius of the orbit then increases with this principal quantum number. In addition, when an electron emitted a photon (light-quantum) and dropped to an orbital closer to the atomic nucleus, this was to happen instantaneously. Such processes, and the quantized nature of the atom, were in direct contrast with classical physics, even though it was still used in much of the derivation of the Bohr atom.

Nevertheless, the explanatory power and predictive successes of the Bohr model made the quantum increasingly popular (Kumar, 2008, pp. 93-117).

Quantum mechanics

Planck, Einstein, Bohr, and others introduced new models that merged classical physics with the new concept of quanta. While these models were able to accurately describe many up until then unexplainable phenomena, there were still others that resisted explanation. Another problem was that the solutions were often rather ad-hoc in nature, resulting from an awkward union of classical and quantum ideas. It was not possible to derive them from a clear and consistent set of first principles. Furthermore, there was no underlying framework connecting all these distinct quanta-utilizing models.

This led to the second phase in the development of quantum theory in around the mid-20s: the matrix mechanics of Werner Heisenberg (1901-1976) and the wave mechanics of Erwin Schrödinger (1887-1961), now both known as quantum mechanics.

Heisenberg published his paper on matrix mechanics in 1925. His aim was to find a theoretical basis for a new 'quantum mechanics' (the mechanics of quantum systems). He succeeded and was able to produce a consistent theory that could make correct empirical predictions on a wide array of microscopic phenomena. One key strategy of Heisenberg was to only allow for observable quantities in the equations of his theory, such as positions, momenta or energy levels. Quantities that were not observable through experiment, such as Bohr's quantized atomic orbitals, were gotten rid of. In his quantum mechanics, as the name suggests, Heisenberg adopted a relatively new type of mathematics in the form of matrices. He represented his observable quantities mathematically as matrices. These are rectangular arrays of numbers with the important property that a matrix A multiplied by a matrix B might yield a different result than B multiplied by A. This 'noncommutativity', as opposed to the commutativity we are used to for real numbers (e.g., 3*2=2*3), is related to Heisenberg's most well-known discovery: the uncertainty principle. This essential characteristic of quantum mechanics will be returned to later, since the concept of the 'wave function' must be introduced for that first (Kumar, 2008, pp. 177-201).

Early 1926, Schrödinger published his version of quantum mechanics called 'wave mechanics'. Here, he built on the work of Louis de Broglie (1892-1987) who had shown that rather than just the phenomenon of light, all matter could display wave characteristics. De Broglie developed a mathematical relation where he related momentum and wavelength: a typical particle and a typical wave property respectively. It seemed that whether we are talking about photons, electrons, neutrons, etc., all behave as if they are particles in some experimental settings, while behaving as if they are waves in other experimental settings. For example, electrons display interference in double slit experiments, a typical wave property. At the same time, they can scatter with photons in Compton scattering experiments, a typical particle property. This dual nature of all quantum entities is referred to as 'wave-particle duality'. This duality is very much foreign to classical physics, where it was assumed that something like light must be *either* a wave *or* a particle, with each position having its adherents. Wave-particle duality is another important quantum property to add to our philosophical arsenal for later (Kumar, 2008, pp. 143-155).

Thus, having been inspired by De Broglie, Schrödinger's goal was to find a 'wave equation' for these 'matter waves'. This ultimately resulted in his famous equation:

$$\left(-\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{x},t)\right)\Psi(\vec{x},t) = i\hbar\frac{\partial\Psi(\vec{x},t)}{\partial t}$$

Not all about this equation needs to be understood for the purposes of this thesis, but there are three things about it that should be.

The first is $\Psi(\vec{x},t)$, which is called the 'wave function' of a quantum system (be it a single particle or collection thereof). The wave function contains all information about the system under consideration. For this reason, we will sometimes also refer to it as the 'state' of the quantum system, as is customary in later formulations of quantum mechanics. Unlike quantities appearing in physics like position and mass in classical physics, the wave function does not have a straightforward interpretation. Schrödinger held that the correct interpretation of his wave function for, for example, an electron, was that its absolute value squared $|\Psi(\vec{x},t)|^2$ denoted the density of electric charge at a position $\vec{x}$ at a time t. Today, the wave function is interpreted very differently, and its relevance for the interpretation of quantum theory makes it so that we will discuss it shortly.

The second thing to understand about the Schrödinger equation is a general remark about the role it plays within quantum mechanics. This role is analogous to that of Newton's second law, $\vec{F} = m\vec{a}$, in classical mechanics. It tells you that if you know the forces acting upon a particle, you can predict the trajectory of this particle at all future times[2]. The Schrödinger equation, rather, tells you that if you know the potential of a single quantum particle, you can predict its wave function at all future times[3]. As we will get to later, this shows that some aspects of quantum mechanics are indeed deterministic.

Thirdly, the Schrödinger equation is a *linear* equation. This means that if one finds a solution Ψ_A to the equation (that is, a wave function Ψ for which the left-hand side and the right-hand side of the equation are equal), and then you find another distinct solution Ψ_B, the linear combination $(c_1\Psi_A + c_2\Psi_B)$ is also a solution to the equation. Here c_1 and c_2 are just numbers you can put in front of these wave functions[4]. This is called the superposition principle, which will be very important for understanding some of the interpretive issues that arise in quantum mechanics.

At first, the quantum mechanics of Heisenberg and Schrödinger seemed completely at odds, with each side arguing theirs was the correct theory[5]. However, Paul Dirac (1902-1984), and later Schrödinger himself, quickly showed that the two theories were, in fact, mathematically equivalent. Both 'pictures' of quantum mechanics could be used to make accurate calculations about atomic phenomena (Casado, 2008). There was now finally a clear quantum mechanical framework. But the work was not yet done. Quantum mechanics was, for example, not consistent with the special theory of relativity, and it could not describe some situations where the number of particles changed. Dirac is usually credited to have kickstarted the final phase of the development of quantum theory that would solve these problems: quantum field theory (Kuhlmann, 2020).

However, before jumping to this third and final phase of quantum theory, first three more important concepts of quantum mechanics need to be introduced. These are Max Born's (1882-1970) statistical interpretation of the wave function (the so-called 'Born rule'), Heisenberg's uncertainty principle and quantum spin. The conjunction of these does not tell one all there is to say about quantum mechanics, as many phenomena, such as quantum tunneling, are left out. Our goal is, however, to discuss only that which is required to ultimately enable understanding of superdeterminism

---

[2] Given two initial conditions, such as the position and velocity of a particle at a certain time.
[3] Again, given two initial conditions.
[4] In reality, c_1 and c_2 are not just ordinary numbers, but *complex* numbers. This is, however, not necessary to know in order to understand what we need from quantum mechanics for this thesis.
[5] This went so far that Heisenberg called wave mechanics "crap" while Schrödinger claimed to have felt "repelled" by matrix mechanics (Kumar, 2008, p. 212).

and the debates surrounding it. For this purpose, the following three concepts will do.

The Born rule

Although Schrödinger introduced the wave function, his interpretation thereof was not universally agreed upon. One such detractor was Born, who rejected Schrödinger's view that the wave function of an electron ought to be interpreted as the density of electric charge smeared out over space. He introduced the statistical interpretation of the wave function, arguing that it just provides us with probabilities. Here, the wave function itself has no physical meaning, but one should look instead to the absolute value of the wave function squared, $|\Psi(\vec{x},t)|^2$. This expression, argued Born, allows you to determine the probability of finding a quantum particle at a position[6] $\vec{x}$ at time t upon measurement. This procedure is mainstream nowadays, and has come to be known as the 'Born rule'. An example with a graphical visualization of $|\Psi(\vec{x},t)|^2$ as well as the application of Born's rule is given with figure 1.

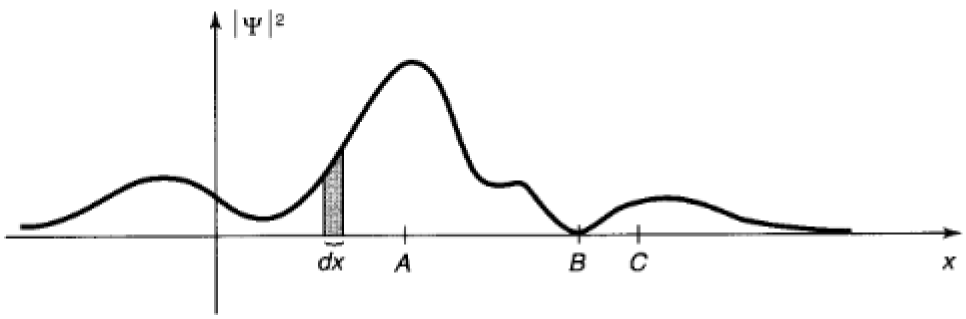


Figure 1: An example of the absolute value of the wave function squared, which according to Born is to be interpreted as the probability of finding the particle at a certain location upon measurement. The grey area tells you the probability of finding the particle in the interval dx. We require that the total area beneath the curve is equal to 100% since we are certain to find the particle somewhere. This is called 'normalization'. Given the height of the curve, it is likely that the particle will appear near A, less likely so near C and very unlikely near B (Griffiths, 2014, p. 3).

---

[6] At least in the case of the wave function in position space. Likewise, one can have a wave function in momentum space whose absolute value squared, e.g., $|\Phi(\vec{p},t)|^2$, expresses the probability density of finding a particle with a certain momentum.

The statistical interpretation of the wave function is very much at odds with the determinism of classical physics. Determinism is at its core the idea that all effects have sufficient causes. If the state of a system is known at a certain time and one is aware of all the (causal) laws of physics, it is in principle possible to predict precisely how the system will evolve in the future. Included in the idea is usually that this process is invertible, such that the state at all past times is likewise uniquely determined. Think of Newton's second law, where given the forces acting upon a particle and its initial position and velocity, it is possible to predict the path the particle will take at all future and has taken at all past times. Yet the wave function does not allow one to precisely predict where a quantum particle can be found upon measurement. If one assumes quantum mechanics to be complete, this opens the door to probabilism about nature: the idea that some processes in nature are fundamentally random, with all that is given being probabilities rather than deterministic certainty. We can then only say something about the probability of, for example, finding a particle at some location upon measurement. The notion of probability often adhered to in quantum mechanics, although further discussed in chapter 9, differs significantly from how the concept is used in classical physics. There, probability is usually viewed as a result of our ignorance[7]. In quantum mechanics, its source is often *understood*[8] as being the fundamental indeterminacy inherent in nature. In *general*[9], no amount of knowledge about the physical system would allow you to predict a property such as the position of a quantum particle precisely (Landsman, Randomness? What randomness?, 2019, pp. 3-10).

The above was not just Born's vision on the issue. The Born rule has been experimentally verified up to extreme precision. There is broad consensus of, at the very least, its *empirical* validity, even from many who would reject the *metaphysical* stance that the universe is ultimately probabilistic in nature. It is now one of the basic postulates of quantum mechanics. Nevertheless, we will see that some superdeterminists challenge the Born rule on both of these counts.

---

[7] Whether a coin flip results in heads or tails is, *in principle*, something that can be precisely determined, if one were to know all the forces acting on the coin, the coin's mass, its distance from the ground, etc. Our assignment of a probability to what side the coin will show is thus a result of our ignorance about all of these factors, rather than being an uncertainty that is inherent in nature.
[8] This view has its detractors however, as will become clear in the next couple of chapters.
[9] To be precise, a quantum one-particle system can be prepared such that $|\Psi(\vec{x},t)|^2$ peaks sharply around one point in space. In that case, relatively accurate predictions are possible. However, not only is this not the general case, in a moment will also be shown that this will come at great cost of how certain we can be about the momentum of this particle.

Heisenberg’s uncertainty principle

Knowing about Born’s rule enables us to introduce Heisenberg’s famous ‘uncertainty principle’. Heisenberg had shown that some observable properties of quantum systems, mathematically represented by matrices, did not always commute. He went on to show that every pair of non-commuting ‘observables’[10] gives rise to an ‘uncertainty relation’. The most well-known pair of observable quantities that do not commute are the position-matrix and the momentum-matrix of a particle. I will proceed using this pair as an example, but some other examples of non-commuting observables[11] are energy & time and angular momentum & angular position.

The uncertainty relation between position and momentum places a *fundamental* limit on how accurate predictions of these quantities of a quantum system, such as one particle, can be. Mathematically, it can be expressed as follows:

$$\Delta x \Delta p \ \geq \frac{h}{4\pi}$$

Here, Δx and Δp are the standard deviations in the position and momentum of the particle in question, respectively. Furthermore, $h = 6.63 \cdot 10^{-34} \frac{m^2 kg}{s}$ is Planck’s constant, which for our intents and purposes is just a ‘very small number’.

Why is it possible to predict positions and momenta in the first place, if from Born’s rule we concluded that quantum mechanics deals with inherent probabilities? This is the case because even if we cannot precisely know where a particle can be found, we do, as stated, have knowledge of the probability of it being found at different locations in the space under consideration. If the probability function $|\Psi(\vec{\mathrm{x}}, \mathrm{t})|^2$ is localized at a relatively small spatial interval, we can make a good prediction of where the particle can approximately be found upon measurement. In this case, the standard deviation of the position Δx would be very small.

The uncertainty principle, however, shows that such relatively precise predictive capability comes with a tradeoff. Since the product of the position and momentum standard deviations $\Delta x \Delta p$ must be larger than at least the number $h/4\pi$, a relatively small Δx must come with a relatively large Δp. This means that precise predictive

[10] In more technical mathematical language, observables in quantum mechanics are self-adjoint operators on a Hilbert space. I include it in this footnote for completeness, but the fine mathematical details of quantum mechanics, much of it owed to figures such as John von Neumann, are in principle not necessary to understand for the goals of this thesis.
[11] In the literature these are called canonically conjugate variables.

capability of where we are likely to find a quantum particle upon measurement, comes with very imprecise predictive power for what the momentum of that quantum particle is likely to be. Rather than being relatively localized, the probability function $|\Phi(\vec{p},t)|^2$ for the momentum of the particle will be very much smeared out, such that it is practically impossible to point out a small and highly probable range of momenta the particle will 'pick from'. A visual example is provided in figure 2. However, since $h/4\pi$ is such a small number, of the order $\sim 10^{-35}\frac{m^2 kg}{s}$, the uncertainty principle is completely negligible for everyday situations. It explains why this fundamental limit had not been found earlier. This observation is one example of what Bohr called the 'correspondence principle', the idea that quantum behavior disappears in favor of that of familiar classical physics when applying the equations of quantum mechanics to large everyday systems[12]. The uncertainty principle thus shows that simultaneous knowledge of both non-commuting observables in a pair is fundamentally limited, irrespective of how much one knows about the system under consideration.

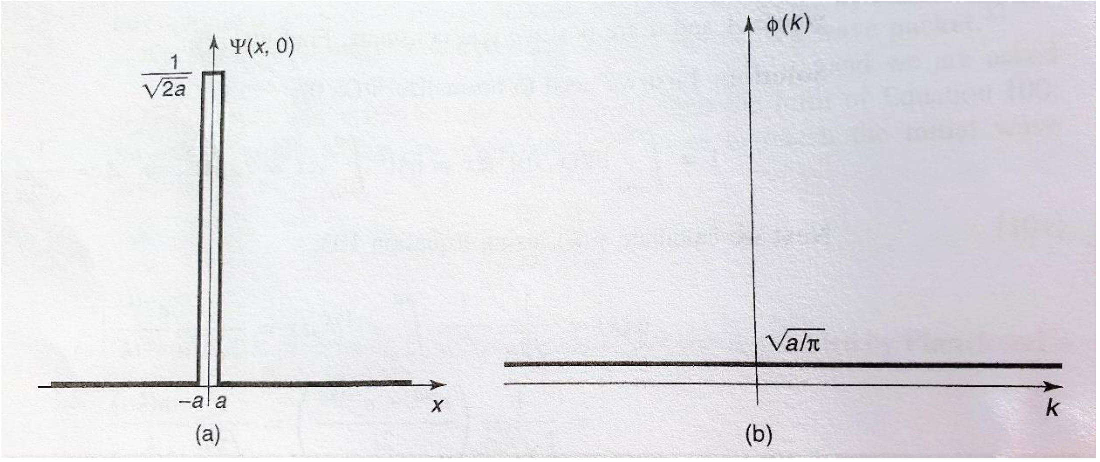


Figure 2: In these graphs we see an example of the wave function of a quantum particle in position and momentum space. The left graph shows the wave function for the position. It peaks in a very small interval while being zero elsewhere, so we can predict quite precisely where the particle will turn up upon measurement. The right graph shows the wave function for the momentum. Since this curve is equal in height across the entire interval, we cannot make a precise prediction of the momentum at all. The curve is much lower because the total area beneath it must sum up to 100%

[12] A bit more formal and precise: quantum physics recovers classical physics in classical limits such as the so-called thermodynamic limit $N \to \infty$, with N being the number of particles in a system.

for both graphs, so a smaller interval must mean a higher curve, as is the case with the left graph. The situation could also be reversed, but the point is that a very small uncertainty in one of the two quantities will result in a very large uncertainty in the other, visualizing the uncertainty relation between position and momentum (Griffiths, 2014, p. 64).

Quantum spin

Lastly, quantum spin must be introduced. To show what this concept entails it is helpful to briefly discuss the Stern-Gerlach experiment, conceived by Otto Stern (1888-1969) in 1921 and carried out a year later by Walther Gerlach (1889-1979). While coming back to this intuition in a moment, I ask the reader unfamiliar with the concept to, for the moment, resist thinking of ordinarily spinning objects, such as a basketball on the finger of a professional player. Imagine quantum spin had been referred to using any arbitrary unused word in our language, befitting of an entirely new concept. Wolfgang Pauli (1900-1958) had, in fact, introduced the property as a quantum number called *Zweideutigkeit*.

The Stern-Gerlach experiment can, somewhat simplified, be described as follows. A beam of silver atoms are sent through a magnetic field. The atoms are then deflected upwards or downwards and thus reach one of two spots. In figure 3 this is shown graphically.

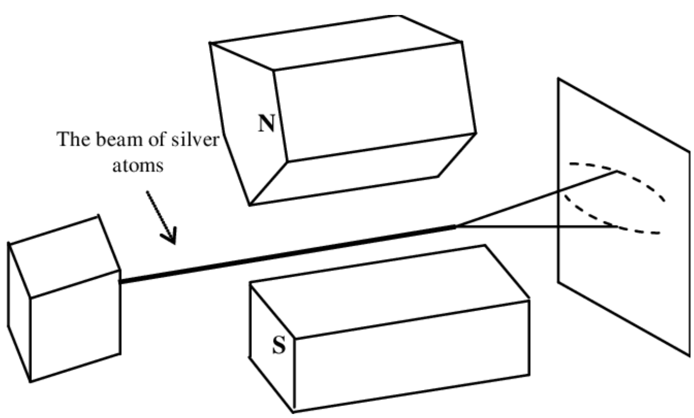


Figure 3: A simplified portrayal of the Stern-Gerlach experiment. Silver atoms in the beam are, when passing through the magnetic field, either deflected upward a certain amount or downward by the same amount. This is the influence of spin (Erkoç, 2009).

Classically, one would expect to find the atoms between these two spots as well. The Stern-Gerlach experiment is, therefore, an example of the quantized nature of the atom. Stern and Gerlach had intended to test some predictions of the Bohr model with their experiment, but only later it was realized that they had discovered a new quantum property referred to as 'spin'. All elementary particles, i.e., all particles that are not composed of other particles, have an associated spin[13]. In fact, elementary particles can be entirely characterized in terms of three numbers: their mass, charge and, indeed, their spin. For some elementary particles, their spin is intrinsically 0, for others 1/2 and yet others 1. For example, a photon always has spin 1, and an electron always has spin 1/2. In the context of the Stern-Gerlach experiment these numbers will tell you something about the magnitude of deflection of the particle. Spin can only be measured against a freely chosen axis[14]. When doing this for an electron, a spin 1/2 particle, the result can either be up (+1/2) or down (-1/2) with respect to that chosen axis, but nothing in between. This why Pauli had referred to spin as *Zweideutigkeit*. In English: 'two-valuedness'. Nevertheless, you will always measure a spin magnitude of 1/2 irrespective of the orientation of your measuring axis, as was also found by Stern and Gerlach. It seems to imply that either the particle carries some instructions of what spin-direction to produce for each possible axis it can be measured against, or that the particle only 'chooses' its spin direction upon measurement. This important interpretive dilemma for quantum mechanics can be demonstrated through spin quite nicely, and it will feature prominently in chapter 3.

So far, we have described that spin is a numerical property of all particles, as well as what spin *does*. Namely, its magnitude (electron: always 1/2) and direction (always +1/2 or -1/2 irrespective of measuring axis) determine how a particle is deflected from a straight path in the presence of an external magnetic field. But coming from a classical worldview, one might wonder what more there is to spin. What *is* it, and what does this number that makes things happen describe in physical reality?

Formally, spin is a purely quantum mechanical concept. It cannot be described in classically visualizable terms. However, there is a well-known classical *analogy* for spin. This analogy works in some ways, but it fails critically in others[15], and one may

[13] For composite particles, these spins can add to different numbers.
[14] Or so one would say, but as a 'sneak peek' of what is to come: it is precisely such a 'free' choice of an axis that superdeterminism forbids.
[15] There are numerous possible examples for this. For one, if a particle like an electron is viewed as a point particle, what is doing the spinning? If an electron is instead viewed as a small sphere spinning about an axis, many problems arise, such as the fact that the spinning motion would need to be faster than light to account for the measured angular momentum. Then there is the issue of how a massless particle such as a photon can be spinning. A final example is that classically, it makes no sense that spinning motion is discrete. In the electron example, we *always* find the spin to be up or down, with the same numerical value.

legitimately ask whether it ultimately does more harm than good. It will now be introduced anyway due to its commonality and primarily due to it working in a way that enhances intuitive understanding for the key concept of chapter 3. Nevertheless, it remains important to realize that it is, in the end, only an analogy of a concept that is in essence quantum mechanical, and this relationship between classical and quantum concepts will in turn be elaborated upon when discussion Bohr's doctrine of classical concepts in chapter 2. In any case, one can imagine a particle as a little ball spinning about an axis, as seen in figure 4. One must remember that the spin of any rotating object is, explicitly or not, always described with respect to an axis it is measured against. The spinning motion can always be in one or two directions: counterclockwise ('spin up') or clockwise ('spin down'). In the analogy, it is the magnitude and direction of this spinning that can be interpreted as quantum spin.

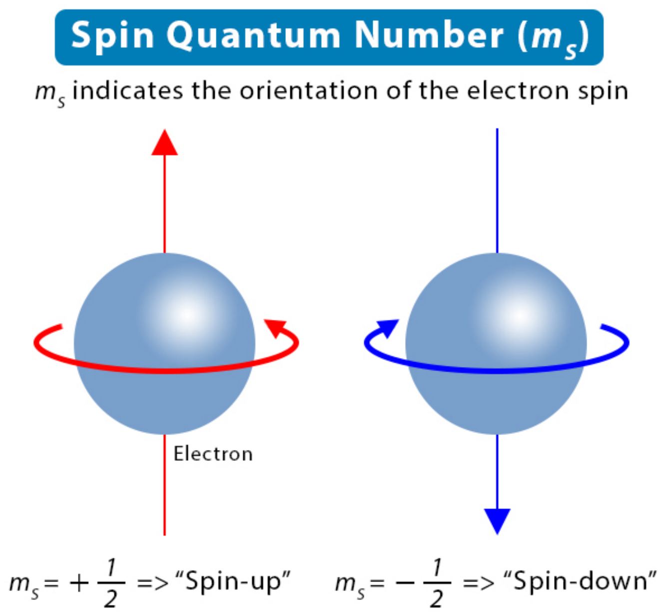


Figure 4: A pictorial representation of the classical analogue of spin. In this analogy, the electron is a small sphere spinning about an axis (here the z-axis), with the counterclockwise spinning being called 'spin-up' (+1/2) and the clockwise spinning being called 'spin down' (-1/2) (Spin Quantum Number, 2022).

Armed with knowledge of the Born rule, the uncertainty principle and quantum spin from quantum mechanics for later, we can proceed to the third and final phase of the development of quantum theory.

Quantum field theory

The third phase of development of quantum theory was that of quantum field theory. It was kickstarted by confronting shortcomings of quantum mechanics. One such example is the inability of the theory to describe the creation and annihilation of particles, such as the spontaneous decay of one particle into others. Quantum mechanics had introduced the concept of quantization to classical physics, whereby many observables could only take on certain discrete values rather than any value from a continuous range. Solving its shortcomings amounted to making two further modifications to classical physics beyond this quantization step, the combination thereof accumulating into modern quantum field theory.

One of these modifications was to make quantum theory relativistic. Einstein's theory of special relativity is built upon the postulates that the laws of physics are the same in all inertial reference frames[16], and that the speed of light is the same for every observer. Without getting into the details, the theory makes different predictions from classical physics, and these predictions are all experimentally tested to extreme precision, with validating results. In order for physics as a whole to be consistent, physicists set out to combine the non-relativistic quantum mechanics with special relativity into one unified framework. In a way, this can be compared to how quantum mechanics itself unified the disparate models of the old quantum theory, yielding one description. The first person to so successfully was Dirac, already in 1928. From the merging of both quantum mechanics and special relativity, he was able to predict the existence of antimatter and open the gates to the development of contemporary quantum field theory (Kuhlmann, 2020).

The second modification was to stop looking at quantum systems as single particles with associated wave functions, interacting with each other, but rather, to look at fields. A field is a physical quantity that ascribes a number (or vector, tensor, etc.) to each point in space and time. More concretely, one could consider the gravitation field in classical physics. Irrespective of where you are and when, you are always under the influence of the gravitational (vector) field, which pulls you towards a certain direction with a certain strength. The field description is able to deal with the aforementioned problem of particle creation and annihilation, as it describes, for instance, all electrons as excitations of a single underlying electron field. Another

[16] A reference frame can be viewed as a coordinate system used to describe events in space and time. For example, I can define a coordinate system whose origin is the center of a park, and express the location of trees with coordinates in this system. An *inertial* reference frame is one that that does not undergo any acceleration.

one of the arguments for the field-approach is that all instantiations of an elementary particles of a type are identical everywhere all the time, perhaps hinting at a single description that spans all of space and time. The central point is that the switch was made from quantum particles to quantum fields (Kuhlmann, 2020).

Taken together, incorporating special relativity and field theory into quantum mechanics makes quantum field theory. Quantum field theory was initially not without problems and has had a long development, but it still stands as the state-of-the-art quantum theory. In fact, the Standard Model of particle physics, the most fundamental description of nature known to physics as of today, *is* a quantum field theory. Given this is more or less the contemporary state of affairs, this is where our historical tour comes to an end. Not because there are no problems left to solve, as there are clear reasons to look for theories beyond standard model. From its inability to incorporate the *general* theory of relativity to the lack of an explanation for there being more matter than antimatter in the universe, our current understanding of the quantum world may be said to have left us with as many questions as it has answered, if not more. As we will see, this is another reason some have set their hopes on superdeterminism.

Crucially, it is not just physics problems that quantum field theory leaves us with, but also plenty of philosophical ones[17]. In particular, since quantum field theory is based on quantum mechanics, the key interpretive issues we concern ourselves with simply carry over from one to the other. So, when referring to the interpretation of quantum *mechanics* in the remainder of this thesis, it may be assumed that these same issues are at play in quantum field theory, unless mentioned otherwise. This is commonplace for the philosophical debate surrounding quantum theory because quantum mechanics is more accessible than quantum field theory while already containing most key interpretive issues.

This concludes our overview of the history of quantum theory. It is by no means complete, but rather a minimum of necessary context, physics and concepts that are a prerequisite to delve into the philosophy and interpretation thereof. There is no better way to introduce this than through the famous Bohr-Einstein debates, which will be the subject of the next chapter.

[17] That is, in so far as a clear distinction between the two can even be made given the concern with a description of the natural world at the most fundamental level we know of.

Chapter 2: The Bohr-Einstein debates

From the advent of the quantum era, physicists have debated about the interpretations of its theoretical and experimental content. This was especially the case pre-WWII on the European continent (Kaiser D. , 2004) (Baggott, 2021), most famously in the form of the Bohr-Einstein debates. These consisted of a series of influential (but cordial) public disputes about quantum theory and its interpretations between Bohr and Einstein. This chapter will first introduce the camps these men represented in face of the new theory. After this, the single-slit thought experiment, the EPR paradox and the Schrödinger's cat thought experiment will be covered, as these introduce centrale themes in the philosophy of quantum theory that later went on to motivate superdeterminism.

The Copenhagen school and its sceptics

After the formulation of quantum mechanics by both Heisenberg and Schrödinger, one could roughly say that two camps had formed. On the one hand, there was the Copenhagen camp. Men like Bohr, Heisenberg, Pauli and Born are often placed here. All of them had for some time worked at Bohr's Institute for Theoretical Physics of the University of Copenhagen, hence the name. The Copenhagen camp believed in what would, decades later, come to be known as 'the Copenhagen interpretation of quantum mechanics'. This was a loose set of ideas about the new quantum mechanics, first coherently combined by Bohr in a lecture he delivered in 1927 at the International Physics Congress in Como, Italy (Kumar, 2008, p. 368). Today, this interpretation of quantum mechanics is still the most popular under physicists. A 2013 survey[18] at a conference on the foundations of physics yielded the result that can be seen in figure 5 (Schlosshauer, Kofler, & Zeilinger, 2013).

[18] Some critical reflection on the source is warranted. First of all, the number of respondents is only N=33, the poll is 10 years old and due to location Austrian physicists are overrepresented in the poll. Perhaps more seriously, the answers might not be all that representative for the physics community as a whole. Interestingly, however, I would argue that if anything this strengthens the central point about the popularity of the Copenhagen interpretation since the particular unrepresentative elements may *undersell* precisely that. The conference where the poll was taken was already on the foundations of quantum mechanics. Those who answered the question, therefore, are likely to be acutely aware of possible problems (that will be expanded upon later) with the Copenhagen interpretation and the existence of several alternatives. Such awareness may be argued to increase the probability of this subgroup of physicists not preferring the Copenhagen interpretation. On the other hand, physicists who are not into the foundations of physics are less likely to know about this and have often been taught just standard Copenhagen quantum mechanics as the way it is (Baggott, 2021). It then seems reasonable to assume that a smaller portion of this group would explicitly prefer alternative interpretations. In conclusion, while there may be plenty of valid criticism of the survey, I believe it

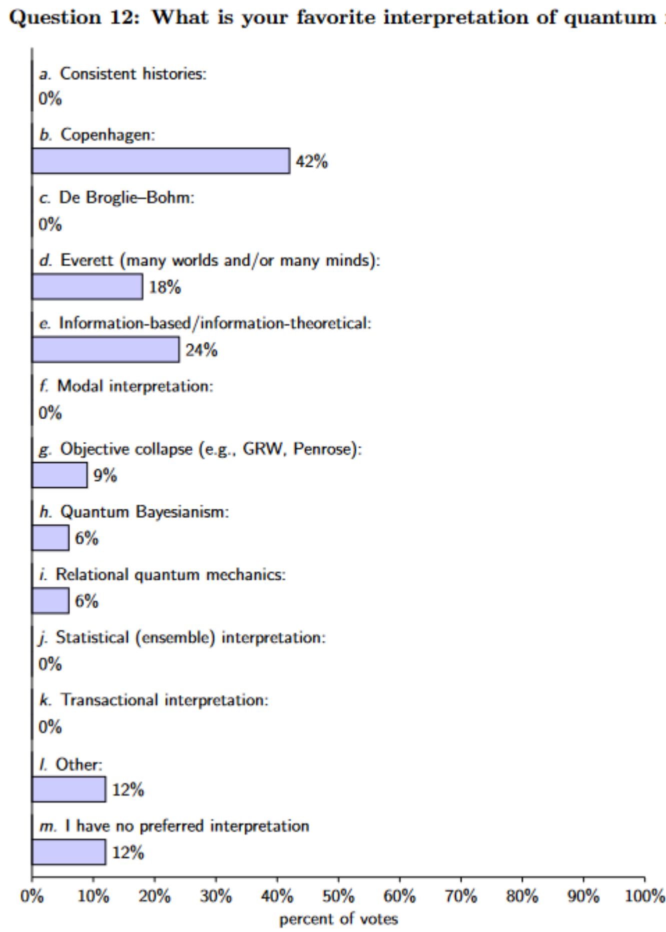


Figure 5: Poll results on the popularity of several interpretations of quantum mechanics (Schlosshauer, Kofler, & Zeilinger, 2013, p. 8).

Before discussing the content of the Copenhagen interpretation, it should be noted that the people associated with it did certainly not agree on everything. Trying to dub any one set of ideas as '*the* Copenhagen interpretation' has by experts in the field been referred to as "hopefully ambiguous" (Bacciagaluppi, Lecture: Introducing Foundations of Quantum Mechanics, 2021), but it has undeniably been highly influential in spite of this. I shall try to provide a simplified and brief overview, being sensitive to relevant differences.

---

can be argued that is still supports the claim that the Copenhagen interpretation is the most popular popular, and even majority, position under physicists.

Firstly, however, the opposition will similarly be introduced. Those skeptical of the Copenhagen interpretation were not necessarily a unified front as was sometimes the case with Bohr, Heisenberg and Pauli. Nevertheless, they all levied criticism against it, sharing some common themes such as critiquing its anti-realism (see below). While some of the skeptics developed alternative theories of their own, others came up with poignant critiques challenging the Copenhagen interpretation. They rarely denied the empirical success of quantum mechanics, i.e., the correctness of its predictions. Their qualms were usually strictly with the interpretations of experiments as proposed by the Copenhagen school. Influential voices in the skeptic camp were those of Einstein, Schrödinger, Planck and the Broglie[19]. The focus here will lie primarily on the ideas of Einstein, as his thought will later be seen to kickstart an avenue that later led some to superdeterminism. Einstein can be put in the camp supporting so-called 'hidden-variable theories'. Such theories assume that quantum mechanics is an incomplete theory, that is merely a statistical average of a more fundamental underlying theory (Kumar, 2008, pp. 354-355). Aspects such as intrinsic probability could then disappear, since not yet known hidden variables could out nature as deterministic after all. The quantum probabilities were, in fact, merely a result of our ignorance of this underlying theory, as was the case with classical probabilities. A historical example of hidden variables are the velocities of microscopic atoms and molecules making up a gas. A gas has macroscopic properties such as temperature, pressure and volume. The theory of thermodynamics describes the relation between such quantities. However, it was later discovered that there exists an underlying hidden-variable theory that was able to reduce the macroscopic theory of thermodynamics to the microscopic theory of statistical mechanics. The motion of microscopic particles in a gas turned out to allow one to explain the source and behavior of the familiar macroscopic quantities such as temperature, pressure and volume. Thus, the velocities of these particles functioned as the hidden variables constituting the theory of statistical mechanics underlying thermodynamics. Superdeterminism, as will be shown eventually, is also a specific type of hidden-variable framework.

In the table below are three columns. For each row, the first column contains a 'central tenet' of the Copenhagen interpretation, the second will explain it and the third will lay out the opposing view (Kumar, 2008).

[19] While de Broglie later accepted the Copenhagen interpretation, before this he was the first to propose what is now known as 'pilot wave theory'. Being, in some sense, superdeterminism's sibling that walked another path, it will get some more attention later.

Central tenet	Context	Criticism
Anti-realism	Quantum particles fundamentally do not have definite properties prior to measurement.	To Einstein, the existence of an observer-independent reality in time and space is the very thing physics is supposed to represent.
Probabilism from the Born rule	The probabilities from the Born rule are not a result of ignorance of an underlying theory, but nature is fundamentally probabilistic. Unlike in classical physics, there is no deterministic causal law allowing us to predict, for example, the position of a particle at all times.	Einstein famously proclaimed that "God does not play dice". As was the case in his theory of relativity, Einstein believed that quantum mechanics would ultimately be replaced by a deterministic theory obeying causality.
The uncertainty principle	The uncertainty principle is correct and shows that there is a fundamental limit to the degree of certainty we can have about non-commuting observables (e.g., position and momentum).	While the uncertainty principle is nowadays uncontroversial, sceptics, including Einstein, have tried to undermine it through thought experiments[20]. Even Copenhagen adherents themselves did not always agree on the nature of the principle. While for Heisenberg the principle resulted from the discontinuous nature of quantum mechanics and the influence of observation on the quantum

[20] The most famous of which were his double-slit with movable screen thought experiment and his light box thought experiment. While the former tried to argue it was possible to both know the momentum and position of a quantum particle in an interference experiment through the use of conservation of momentum, the latter tried to show that one could know the energy and time interval of an event involving a photon to arbitrary precision by applying $E = mc^2$ to a photon escaping a box of light. While it is generally accepted Bohr was able to refute both thought experiments, some experts claim that he contradicted himself in the process (Kumar, 2008, pp. 271-273).

		system, Bohr insisted that it was a more fundamental consequence of the complementarity of the wave and particle nature of quantum objects.
Complementarity	This is one of the most central convictions Bohr expressed about quantum mechanics. According to it, quantum systems can have complementary properties, such as wave and particle behavior or position and momentum. That is, both are a valid description of the object, but in each experimental setting only one of these properties can be displayed. In his own words: *''Evidence obtained under different conditions cannot be comprehended within a single picture, but must be regarded as complementary in the sense that only the totality of the phenomena exhaust the possible information about the objects.''* (Bohr, Discussion with Einstein on Epistemological Problems in Atomic Physics, 1949)	Many of the sceptics, including Einstein, rejected the principle of complementarity. Einstein repeatedly tried to show that complementary properties could, in fact, both be displayed in a single experimental setting. De Broglie's pilot wave theory also provides an ontology counter to that of the complementarity thesis. Other Copenhagen adherents like Heisenberg did not reject complementarity, but emphasized it far less than Bohr did.
The doctrine of classical concepts	Even though quantum objects do not behave classically, the essence of the doctrine of classical concepts	Rather than the use of classical concepts being inescapable, Einstein thought they would have the be replaced by

	is that experimental results must still be described in classical terms. This is an epistemological requirement, as, according to Bohr (and Heisenberg), it is a unique new feature of quantum theory that it is not possible to distinguish clearly between the behavior of an observed object and the way in which it is observed. There are different views on what the exact reason was for Bohr's strong belief in the doctrine of classical concepts, but establishing it as such suffices for our purposes (Faye, 2002).	radically new ones. Concepts like position and momentum had to be given up due to their "shacky" meaning in quantum mechanics. He referred to the doctrine of classical concepts as a "tranquilizing philosophy" (Kumar, 2008, p. 321).
The correspondence principle	As stated before, the correspondence principle consists of the idea that when applying the quantum formalism to everyday situations (where there are many particles and Planck's constant is negligible), it should reproduce the results of classical physics.	While this principle is often advanced as a core idea of the Copenhagen interpretation, Einstein would not have had a problem with this as his own theory of relativity made use of a similar principle. There exist, however, some modern critiques of the principle not holding up for quantum mechanics when applied to large chaotic systems (Zurek, 2003).
Epistemic wave function	The wave function in Schrödinger's equation, according to the epistemic view (or Ψ-epistemic) of the wave function, does not refer	Einstein would agree with a Ψ-epistemic view, albeit for very different reasons. While the Copenhagen interpretation does so because they think the

	to any real-world object. This symbolic view is held by the Copenhagen school. In contrast, the ontological view (or Ψ-ontic) view of the wave function is of the opinion that the wave function corresponds to something really out there in nature.	wave function only says something about the *knowledge* one has of a system, Einstein would say that the wave function is only a statistical average emerging from and reducible to properties of an as of yet unknown deterministic hidden-variable theory. This contrasts with the Ψ-ontic view of Schrödinger who thought of the wave function squared as a cloud-like distribution of a particle's charge and mass, or with de Broglie for who the wave function is a physical wave guiding the trajectory of particles.
Collapse of the wave function	In the Copenhagen interpretation a quantum particle does not have a position prior to measurement, and we can only say something about the probability distribution $\|\Psi\|^2$ of it turning up somewhere. Yet, after measurement, the particle is *really* at a specific location. It is then said that the wave function has 'collapsed' or 'reduced', and that $\|\Psi\|^2$ must be updated if it is to accurately describe all that can be known about the system.	Einstein and Schrödinger had serious problems with the idea of wave function collapse. Their critique has been most famously stated through the thought experiment of 'Schrödinger's cat' that we will get to later in this chapter. In addition, the collapse of the wave function is another example of disagreement between Bohr and Heisenberg. While Heisenberg adopted this idea, Bohr did not. Since the wave function was merely symbolic for Bohr, the idea of

	Visually, the curve in figure 1 becomes an infinitesimally small peak at that location.	a collapse did not make sense to him.
Quantum mechanics as a complete theory	For the supporters of the Copenhagen interpretation, quantum mechanics was a complete theory. This means that the fundamental features of it would not be altered in the future and that the theory fundamentally contained the ingredients to uncovering all that can be said about nature.	Many sceptics did not agree. As we shall see in a moment, Einstein's EPR-paper addressed exactly this claim.

One more note on the above table is that it is important to take into account that not all alternative interpretations disagree on all of these tenets with the Copenhagen interpretation. Just like they usually accept the *empirical content* of the Born rule and Heisenberg's uncertainty principle, they may accept some of these tenets but have significant qualms with other ones.

Now that the contents of the Copenhagen interpretation are clearer, as well as the positions of its sceptics, it will be easier to contextualize the debates that follow and to recognize interpretive issues later on. The following three subsections will discuss Einstein's single-slit thought experiment, the Einstein-Podolsky-Rosen (EPR) paradox and Schrödinger's cat.

The single-slit experiment and the 'peculiar mechanism of action at a distance'

Solvay conferences are large gatherings of experts in physics and chemistry, devoted to tackling major open problems in certain subdisciplines. One important Solvay conference took place in Brussels, 1927. With the topic being 'Electrons and Photons', the idea was to discuss the new quantum mechanics and its meaning. As can be seen on the iconic imagine below, all of our 'main characters' so far were present.

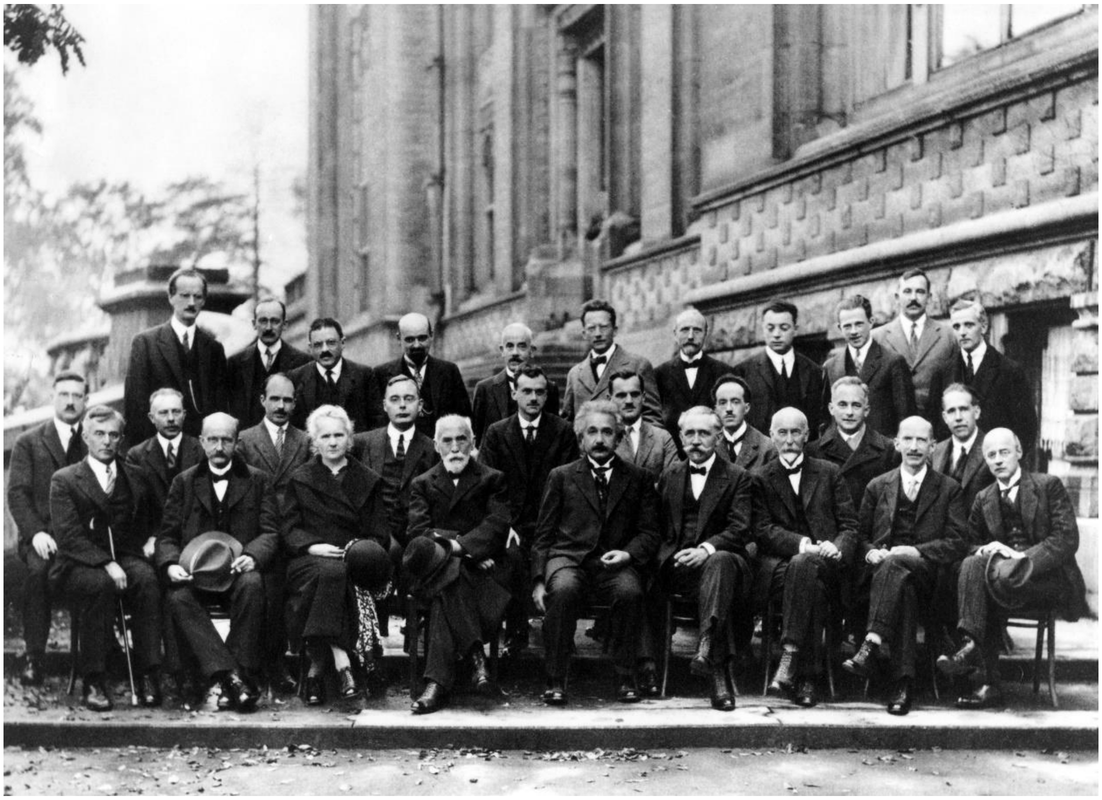


Figure 6: Attendants of the 1927 Solvay conference on Electrons and Photons. Bohr can be found at the right end of the middle row, while Einstein is seated in the middle of the third row (Solvay conference 1927 (group photograph), 2020).

Although many interesting discussions on the interpretation of quantum mechanics took place during this and the later 1930 Solvay conference, we will focus on a small part of the debate between Einstein and Bohr (the latter, by extension, joined by Heisenberg and Pauli) most relevant to our purposes.

Through the use of thought experiments, Einstein attempted to lay bare perceived problems of the Copenhagen interpretation. His first attempt at this was through the use of his single-slit thought experiment. A graphical display can be seen in figure 7 below.

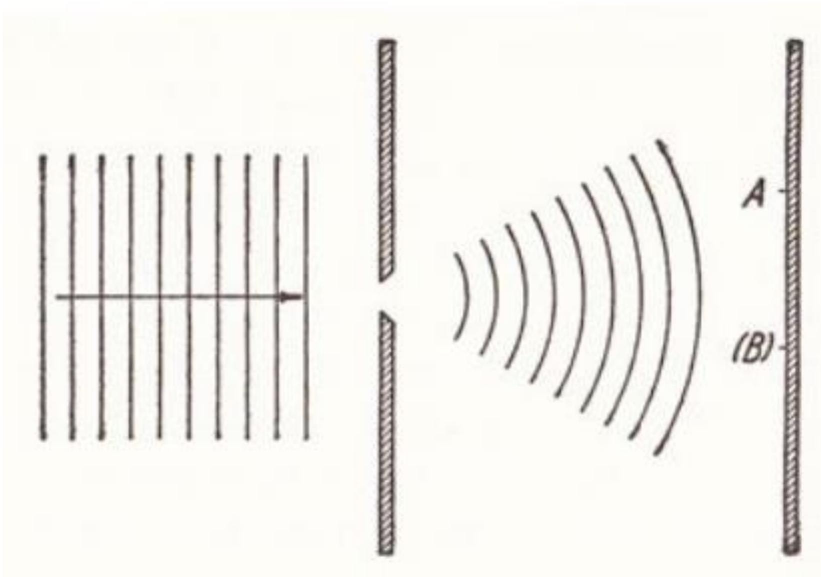


Figure 7: A visualization of Einstein's single-slit thought experiment. A photon passes through the slit in a screen and diffracts. The second screen is a photographic plate that detects the photons at a specific point (Kumar, 2008, p. 264).

In his thought experiment, one can imagine individual photons being fired one at a time at a screen with a very small slit. We view the setting from above. We then see the wave nature of light at play, as the wave diffracts through the slit and spherically propagates towards a photographic plate. This plate registers the photon somewhere as the wave hits the screen. At that moment of measuring, the particle nature of the photon manifests itself instead. Suppose the photon is detected at point A on the screen. The photon, of course, has an associated wave function giving the probability of finding the photon at any place on the screen. But once the photon is measured at A, the probability of the particle being in that spot must *instantaneously* be updated to 100% for A while simultaneously to 0% for any other spot, such as B in the figure. If this were not the case, there would be a non-zero probability of a second detection occurring at B or elsewhere. The wave function has 'collapsed' to the point A. With this background information, we can try to understand Einstein's own words on the matter at the general discussion of the conference:

"If $|\Psi|^2$ were simply regarded as the probability that at a certain point a given particle is found at a given time, it could happen that the same elementary process produces an action in two or several places on the screen. But the interpretation, according to which $|\Psi|^2$ expresses the probability that this particle is found at a

given point, assumes an entirely peculiar mechanism of action at a distance, which prevents the wave continuously distributed in space from producing an action in two places on the screen." (Bacciagaluppi & Valentini, Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference, 2009, p. 487)

Thus, quantum mechanics must either be incomplete or nonlocal. Here, 'nonlocal' refers to the faster-than-light wave function collapse. The 'peculiar mechanism' responsible for this would, after all, have to be instantaneous to prevent multiple actions from conceivably being detected on the screen. But such instantaneous action at a distance is often understood to be at odds with relativity. If we choose to reject the nonlocal character of this process, the quantum mechanical description of the event is incomplete in the sense that is must be modified to account for the observation of only ever observing one signal.

Einstein's point, however, was not entirely clear. Bohr's direct response was recorded to be as follows: *"I feel myself in a very difficult position because I don't understand what precisely the point Einstein is wants to make."* (Bacciagaluppi & Valentini, 2009, pp. 487-488). This may also have been the case because Bohr understood the wave function as an abstract symbolic wave of probability, to which, by virtue of its non-physicalness, instantaneous exchanges of information do not apply. Later that evening, he gave a reply that rather than addressing the point about locality, aimed to advance the consistency of the uncertainty principle when applied to the slit experiment. Einstein went on to 'work with what he got', and both at Solvay 1927 and 1930 many more discussions were had on issues such as the validity of the uncertainty principle. Years later, however, Einstein came back with a clearer and more well-known thought experiment that was more of a continuation of the above argument using the principle of locality. We turn our attention to this now.

The EPR-paradox and the (in)completeness of quantum mechanics

After the Solvay debates, Einstein gave up on trying to undermine the uncertainty principle. He accepted the *correctness* of quantum mechanics, in the sense of acknowledging the accurateness of its many predictions. Rather, together colleagues Boris Podolsky (1896-1966) and Nathan Rosen (1909-1995), he attempted to show that the theory was *incomplete*. This was the goal of the 1935 paper the three ('EPR') published (Einstein, Podolsky, & Rosen, 1935).

At the start of the paper, they state what they take to be a necessary condition for the completeness of a physical theory, namely that *"Every element of the physical reality must have a counterpart in the physical theory."* This makes at least intuitive sense, as a physical theory could hardly be called complete if nature was full of observable physical quantities that are not described by it. As this condition makes reference to ‘physical reality’, their next step is to provide a criterion for precisely that: *“If, without in any way disturbing a system, we can predict with certainty the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.”* Here, ‘disturbing the system’ is an effect of performing a direct measurement, but the idea is precisely that if the value of the quantity being measured can be predicted without doing so, it is an element of reality. While the authors admit that this is not the only way of characterizing physical reality, a precise definition of that is not the purpose of the paper. It simply supposes that most people will indeed intuitively agree with the idea that if I can, for example, predict the position of an electron precisely without measuring it, ‘position’ seems to be an element of physical reality. It is enough to challenge the Copenhagen contention that elements of physical reality possess no definite values of physical quantities like spin and momentum prior to measurement. From these two criteria, the strategy of the EPR-paper becomes clear. After all, if it were to be the case that one was able to predict such quantities with 100% certainty without disturbing the system, clearly these are elements of reality that are not described by quantum mechanics. It follows then that the theory would be incomplete.

For the next step of their argument, it would be instructive to introduce the concept of ‘quantum entanglement’. Two quantum states can be said to be entangled when it is not possible to describe the state of one independently of the state of the other. A brief example invoking spin might make this clearer. Suppose we have the following quantum state[21]:

$$\Psi = \psi_A(1)\psi_B(2) + \psi_A(2)\psi_B(1)$$

Here, we have one composite state Ψ which is a superposition of two possible states. More details will follow later, but for now it suffices to say that with 50/50 probability, either the composite state in the first term will be found upon measurement[22], or the one in the second term. Think of A and B as two quantum particles, with ψ_A being the wave function of particle A and ψ_B that of particle B.

[21] For those familiar with entanglement: I have left out a normalization factor or complex coefficients for didactive purposes.
[22] This particular measurement involves measuring the spin of one of the two particles. This can always be done through a Stern-Gerlach experiment.

Both particles can be in state 1 or state 2. For our example, state 1 can be interpreted as ‘spin up’, and state 2 ‘spin down’.

Now suppose that I take a measurement of particle A and find it to be in state 1. Then I immediately know that the system as a whole is in the composite state seen in term 1, i.e., $\psi_A(1)\psi_B(2)$. The system is, after all, either in this state or in the other, but in the other particle A is in state 2 so that cannot be the case.

The measurement of the state of one particle thus immediately tells me the state of the assumedly separate particle B. If, in the example, I measure A to be ‘spin up’, I immediately know that B is ‘spin down’. Thus, it is not possible to independently describe these states. They are entangled.

Entanglement is not necessarily a counterintuitive, strange concept. If I put a red ball and a blue ball in two separate boxes and give you one, then upon opening your box you immediately know not only the ‘color state’ of your own ball, but also of mine. The correlating aspect, even over large distances, it not the ‘strange’ thing about quantum entanglement. We shall see later when covering Bell’s theorem that this lies in the fact that these correlations are stronger than classically explainable.

Returning to the EPR-paper, the following argument is now made. It is in reality more generalized and abstract than how it is portrayed here, but the idea is the same. Suppose I have two quantum particles resulting from a decay of a composite particle at rest. The law of conservation of momentum tells us that if we know the momentum of one particle, we can infer the momentum of the other. They are entangled due to the conservation of momentum. Now if I make a measurement on the momentum of one particle, I instantly know what the momentum of the other particle is with 100% certainty. The particles can be separated so far that it would take light years for a signal from one to reach the other, such that influence between them at the moment of measurement would be impossible. The latter, of course, constitutes the principle of locality. Again, one must keep in mind that according to the Copenhagen interpretation, neither of these particles possesses a definite momentum before they are measured. The fact that by measuring one of them, certainty could be gained about the other, however, had to imply that either there must have been a reality to the momentum of that particle to begin with, or quantum mechanics in nonlocal, the measurement of the momentum of one particle instantaneously influencing the momentum of the other. Einstein opted for the former option.

The authors now present that one could reply with the objection that *"two or more physical quantities can be regarded as simultaneous elements of reality only when they can be simultaneously measured or predicted"*. In our example these physical quantities are the momenta of the particles. The EPR-authors respond that this conception of the reality of physical quantities is unreasonable. It would mean, after all, that the reality of the momentum of one particle is dependent upon whether I measure the momentum of another particle, which is implied to intuitively be an *ad-absurdum*.

In conclusion, it is argued to be possible to predict a physical quantity of a quantum particle with certainty without disturbing (measuring) it. The Copenhagen interpretation states that there is no reality to physical quantities of quantum particles prior to measurement. Given the criteria of reality and completeness, this means quantum mechanics is an incomplete theory.

Later that year, Niels Bohr published a response to the EPR-paper by the same title (Bohr, Can Quantum-Mechanical Description of Physical Reality be Considered Complete?, 1935). In it, he argued that the EPR-paper contained *"an essential ambiguity [concerning the] criterion of physical reality formulated in the paper by Einstein, Podolksy and Rosen when applied to quantum phenomena"*. Bohr's strategy was to question the presumption of 'separability' in the EPR-paper[23], which is the idea that the two systems when spatially removed from one another can be treated individually. Due to the entanglement, Bohr argued separability does not hold, and rather than two distinct systems it is one two-particle system one is treating. It is then not possible to do a measurement on one particle 'without disturbing the other', as not being separate entities, measuring one particle in an entangled pair necessarily also disturbs the other, albeit not in a mechanical way. To this end, Bohr also invokes complementarity and the doctrine of classical concepts.

Other than with separability, Bohr's argument may still seem to be at odds with locality, for now roughly viewed as the idea that faster-than-light influence is forbidden. It may be argued that regardless of whether the two-particle quantum system is considered as one whole or as two individual ones, an instantaneous influence from A on B over great distance is still required, upon consideration of the

[23] For Einstein, the separability principle was an important means of grounding his realism about physics, and even for physics to be able to describe the world at all. The idea of the state of one system being dependent upon whether another is measured, was unacceptable to him given their spatial separation. He found the separability principle to take far too much of a back seat in the eventual EPR-paper, which was written primarily by Podolsky (Howard, 1985).

EPR thought-experiment. Bohr seems to bite the bullet on this in his response, writing: *"But even at this stage [of measuring] there is essentially the question of an influence on the very conditions which define the possible types of predictions regarding the future behavior of the system."* While this passage may be found to be somewhat confusing, one can rest assured that they are not alone in this feeling, and it is in, in fact, rather commonplace. Nevertheless, some conclude from the text that Bohr acknowledges a kind of 'quantum nonlocality' by recognizing the 'influence' in question (Kumar, 2008, pp. 310-312).

This strongly clashed with Einstein's view, who stated in a letter to Born in 1947 that:

"I see of course that the statistical interpretation, the necessity of which in the frame of the existing formalism has been first clearly recognized by you, contains a clear degree of truth. Yet I cannot seriously believe in it, because the theory cannot be reconciled with the principle that physics has to represent a reality in space and time, without spooky action at a distance." (Einstein & Born, Born-Einstein Letters, 1916-1955: Friendship, Politics and Physics in Uncertain Times, 2004)

To him, Bohr's view admitted a certain 'spooky action at a distance', in direct contrast with locality. This spooky action does not, as is sometimes believed, refer to entanglement as such. Rather, it refers to the 'peculiar mechanism of action at a distance' inherent to wave function collapse in the Copenhagen interpretation that he had already addressed in the single-slit thought experiment years earlier. He believed that there was a (hidden-variable) theory underlying quantum mechanics that respected realism, locality and determinism, and that it would one day come to replace it.

For decades, this debate was unresolved. It was only after the death of both Bohr and Einstein that John S. Bell, inspired by the EPR-setup, discovered that quantum mechanics and hidden-variable theories actually produced testable differences. The result of the experiments following this directly led to thinking of superdeterminism, as will be discussed soon.

Schrödinger's cat and the measurement problem

The measurement problem may well be the most discussed problem in quantum foundations. Many argue that it is not just a matter of philosophy, but a serious inconsistency with the physics itself. Consequently, many interpretations of quantum theory are in large part motivated by the ambition to solve it. In this subsection, I will explain what the measurement problem is and where it comes from, using the famous analogy of Schrödinger's cat.

In chapter 1 it was explained that if one finds a solution Ψ_A to the Schrödinger equation, and then you find another distinct solution Ψ_B, the linear combination $(c_1\Psi_A + c_2\Psi_B)$ is also a solution to the equation, c_1 and c_2 generally being complex numbers. This addition of wave functions was called the 'superposition principle'.

This is, so far, rather abstract. A concrete example could, once again, be a quantum particle that can be either spin-up or spin-down:

$$\chi = \frac{1}{\sqrt{2}}\chi_{up} + \frac{1}{\sqrt{2}}\chi_{down} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$$

Here, χ is the spin-wave function. It is a superposition of the spin up solution and the spin down solution. The spin up solution can be represented by a column vector $\binom{1}{0}$, with a 1 in the 'up-direction' and a 0 in the 'down-direction', and vice versa for the spin down solution $\binom{0}{1}$. These can also be added together, as is done in the last step. The factors $\frac{1}{\sqrt{2}}$ are a concrete implementation of the numbers c_1 and c_2.

A quantum particle can be said to exist in a superposition of states. This follows from the Schrödinger equation. Yet, if we were to perform a Stern-Gerlach experiment to find out in what spin state the particle is, in fact, in, we would *always* find it to be either spin up or spin down. As we have seen, it is said that the wave function collapses to one specific state. Measurements do not yield superpositions of states, they yield particular states. Knowing this, the numbers in front of the states can be clarified. The squared value of these numbers corresponds to the probability of finding the quantum particle in the associated state. In this case, since $\left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}$, both spin states have a probability of 50% of occurring. This is an easy example, but in general quantum superpositions can include many individual states, each with a different probability.

Now if the mathematical formalism of quantum mechanics is applied to the measurement process, we also find a superposition of macroscopic detector states (for example: detector reading spin up or spin down). The question then rises how we get from these superpositions to the clear measurement outcomes we observe in reality. It appears that the act of measurement 'does something' to the quantum state causing it to collapse. But if quantum mechanics is a fundamental theory, the behavior of these detectors should follow from quantum mechanics.

A further issue that has not yet been discussed at all is that of what precisely constitutes a ‘measurement’ in quantum mechanics. Yet, a clear conception of this term is very much relevant if we state that wave functions collapse ‘upon measurement’. Bohr’s view on measurement was that it essentially comes about through an irreversible interaction of a quantum system under consideration and a classical measuring apparatus[24]. It should be noted that for Bohr, and most of his companions, measurement is independent of any (conscious) observer. This view highlights Bohr’s doctrine of classical concepts, where the formulation of quantum mechanics still requires classical mechanics. He recognized that this put quantum mechanics at a “very unusual place among physical theories”, since it “contains classical mechanics as a limiting case [the correspondence principle], yet at the same time requires this limiting case for its own formulation.” (Bell J. , Against 'measurement', 1990). From this, however, more questions can follow. One may ask what exactly is the nature of this ‘irreversible interaction’ Bohr speaks of. Another question that follows is where exactly this line is between the classically described measurement apparatus and the quantum system in question. John Bell, the ‘main character’ of the next chapter, referred to this as the ‘shifty split’ between classical and quantum mechanics.

In brief: how and when do wave functions collapse? If quantum mechanics only gives us superpositions, why do we observe singular measurement outcomes instead? Why is it that a classically described detector is invoked to describe collapse, even though the functioning of the detector should itself be subject to the fundamental description of quantum mechanics? And what even constitutes this distinction between a macroscopic detector and a microscopic quantum system in the first place?

Rather than being one clearly defined question, it is this group of related questions that is usually referred to as ‘the measurement problem’ (Bacciagaluppi & Valentini, 2009, p. 155).

---

[24] Note that terms like ‘apparatus’, ‘detector’ and ‘observer’ are sometimes used interchangeably as the system does the ‘observation’ or ‘measurement’. These should not be viewed as ‘conscious human’ or ‘big metal machine’ necessarily. They can be, but it is more instructive to think of anything that can interact with the quantum system under consideration in a way that allows for the derivation of information about that system.

Everybody knows about Schrödinger's cat[25]. What not everybody knows, was that Schrödinger invoked this thought experiment as an intuitive way to display the supposed absurdity of the Copenhagen answer when confronted with the measurement problem. The setup can be seen in figure 8 below.

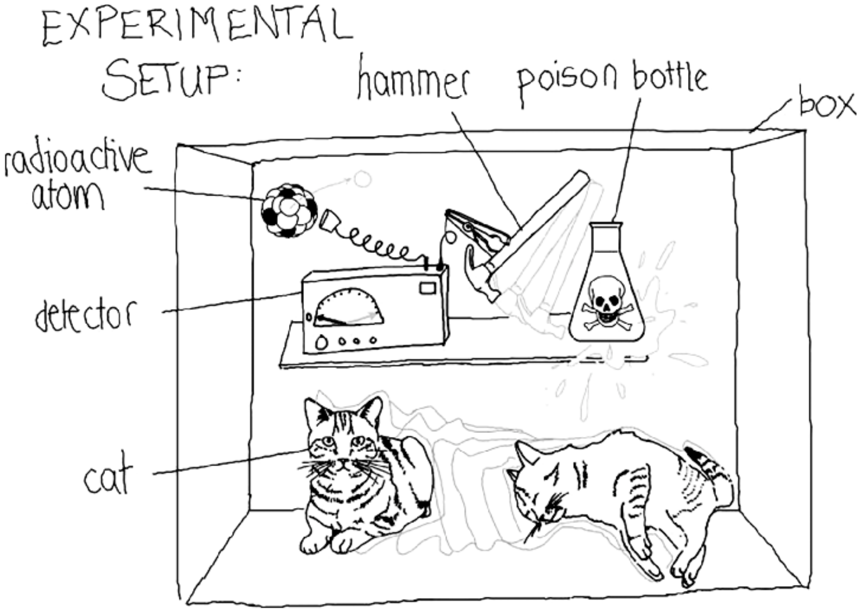


Figure 8: The experimental setup of the Schrödinger's cat thought experiment (Hartjes, 2016). *The fate of the cat is decided by the probabilistic process of a radioactive atom decaying or not.*

A cat is in a small room. The room contains a radioactive atom that, if it decays, will activate a Geiger counter. This is a device that can measure radiation (the detector in the figure). The Geiger counter is connected to a system with a hammer that will strike and break a small flask of hydrocyanic acid if activated. This would kill the cat.

It should be noted that the exact decay time of a radioactive atom is another quantity whose value quantum mechanics cannot predict with certainty. As in the case of position, momentum, spin direction, etc., we can only say something about the *probability* of decay within a certain time interval. Here too, a realist like Einstein would claim that there may be hidden variables to the atom that, if they were known,

[25] Nevertheless, even this thought experiment can be thought of as a part of the Bohr-Einstein debates, since Einstein had more or less already described the same thought experiment to Schrödinger before Schrödinger had done so, just without the cat (Kumar, 2008, pp. 315-317). In addition, Bohr gave some attention to this issue as well.

would allow us to predict when the atom was going to decay. The Copenhagen interpretation, however, views the indeterminacy as fundamental.

Schrodinger invites us to imagine that the particular radioactive elephant in the room has a 50% probability of decaying within an hour, and a 50% probability of decaying after an hour. According to the mathematical formalism of quantum mechanics we can write down the wave function of the system under consideration after precisely an hour as something like this:

$$\Psi = \frac{1}{\sqrt{2}}\psi_{cat}(alive)\psi_{atom}(not\ decayed) + \frac{1}{\sqrt{2}}\psi_{cat}(dead)\psi_{atom}(decayed)$$

Like before, the state of the system of the room is an entangled superposition of two individual quantum systems. According to quantum mechanics, prior to observation the cat is in a *superposition of being dead and alive.* But let us contrast this with common, everyday sense. If we leave the cat alone for an hour in this room and then open it to check, we will not find this superposition of a dead and alive cat. We will always find the cat to be either dead or alive. The Copenhagen interpretation in its anti-realism seems to imply, however, that *before this opening process the cat is in fact in this bizarre superposition*.

In addition, the thought experiment also invites us to consider how and where the collapse here actually happens, and how to delineate between classical and quantum systems.

Unlike for Heisenberg, for Bohr there was no line to be found between the observed and the observer. To him, the situation was akin to a blind man and the cane he uses to gain sensory information about his environment. They are “inextricably bound together” through the act of measurement (Kumar, 2008, pp. 317-318). In addition, as stated before, he did not believe in actual collapse since he viewed the wave function as symbolic. In this way, Bohr ‘solved’ the problem, but not without leaving many with questions such as how measurements are then made. This included Einstein and Schrödinger who were convinced it all went to show the incompleteness of quantum mechanics.

Nowadays, it is sometimes believed that the measurement problem has been solved through the introduction of *quantum decoherence* (Sivasundaram & Nielsen, 2016, p. 6). This is the process whereby the entanglement between a quantum system and the environment, quantum phenomena such as interference are suppressed. A quantum particle does often not exist in a vacuum, and the interactions it has with other particles are an example of the ‘environment’ that is referred to here. It can then mathematically be shown that these interactions lead to entanglements that have an effect on certain quantum behaviors. This, however, does not mean that the measurement problem is thereby solved. Quantum decoherence, when accounting for it, shows how the complex probabilities that appear in quantum mechanics average out to real, classical probabilities for what we measure with our macroscopic devices. But it does not show how it is that, in the end, one of these probabilistic outcomes is realized in nature over another. On the contrary, some experts argue that decoherence only serves to widen the measurement problem (Bacciagaluppi, The Role of Decoherence in Quantum Mechanics, 2020) (Landsman, Foundations of Quantum Theory, 2017, pp. 440-444).

Much more could be said about the Bohr-Einstein debates and the many interesting experiments, be they in thought or in reality, that have been created about quantum mechanics over the years. These include, among others, Wigner’s friend, the quantum Zeno effect, the delayed-choice quantum eraser, the Elitzur-Vaidman bomb tester and most famously the double-slit experiment. But for our purposes, the single-slit, EPR and Schrödinger’s cat thought experiments provide a sufficient basis to understand the road to superdeterminism. The next chapter will introduce developments in the foundations of quantum theory in the latter half of the 20th century that directly led to the introduction of superdeterminism.

Chapter 3: Bell's theorem

Up until now, the debate between Bohr and Einstein had been one of theory and philosophy. Both men used logic and thought experiments to try and advance their own view. But as it turned out, their views came with empirically testable differences. This allowed for experimental physics to have a say on the matter, which it did to great effect. The person who made this possible was physicist John Stewart Bell (1928-1990). Frustrated with the lack of foundational analysis in Copenhagen dominated physics and inspired by alternative interpretations of quantum mechanics by figures like Einstein and David Bohm (1917-1992), he set out to derive the inequality that would allow different views on quantum theory to be put to the test. The final result came to be known as 'Bell's theorem'.

In this chapter, the road to Bell's theorem will briefly be described. After this, his inequality will be derived, allowing us to look at the experimental results that have come in its wake. Finally, we take a deeper look at what these results mean for the debate surrounding the interpretations of quantum theory.

Bohm's reformulated EPR-setup

In 1932, mathematician John von Neumann (1903-1957) published his book "Mathematical Foundations of Quantum Mechanics". It was highly influential as it provided a very mathematically rigorous account of quantum mechanics, using the framework provided by the Copenhagen school[26]. The book also contained a mathematical proof of the statement that it was impossible to reproduce the statistics found by measurements of quantum systems through a deterministic hidden-variable framework. This was an important factor in the popularity of the Copenhagen interpretation and the little interest in investigating hidden-variable alternatives (Kumar, 2008, pp. 336-338).

Von Neumann's proof, however, rested on some assumptions that made it so that it, consequently, only ruled out a particular subclass of such theories. This was explicitly brought to attention by Grete Hermann (1901-1984) in 1935, but her work never received mainstream recognition (Crull & Bacciagaluppi, 2016). It was Bell who independently made the same observation in 1966. Bell was inspired to investigate the interpretations of quantum mechanics in important part through the work of Bohm. Bohm had shown that it was possible to come up with a hidden-

[26] Which is not to say Von Neumann agreed with Bohr on all issues, as unlike Bohr, Von Neumann took the measurement problem very seriously and believed subjective observers were involved in wave function collapse.

variable theory that could reproduce the results of quantum mechanics. This 'pilot wave theory' will be discussed later. He had also, together with Yakir Aharonov, recasted the EPR-setup in a simplified manner (Bohm & Aharonov, 1957). In their paper on the topic, they imagine a molecule made up of two atoms whose spins sum up to zero. They then separate these two atoms over large distances. Both atoms have a nonzero spin, for example 1/2. Due to conservation of angular momentum, the total spin of the two atoms is zero. Therefore, we know that if the spin of one of the particles is measured as 'up' in one direction, the spin of the other particle must be 'down' in that direction. Thus, we have an entangled state. This concept has been introduced in chapter 2 with, for obvious purposes, an example that is just like this. It can be expressed mathematically as follows:

$$\Psi = \frac{1}{\sqrt{2}}(\psi_A(up)\psi_B(down) + \psi_A(down)\psi_B(up))$$

In this form, the EPR statement becomes something of the form that quantum theory cannot be a complete description of reality, as without assuming spooky action at a distance, we come to the conclusion that the atom not disturbed by measurement must already have had a definite spin value all along, rather than being in a superposition. After all, when measuring +1/2 for one atom, we know that the other one must be -1/2, were we to measure it.

To derive his famous inequality, Bell used the above setup provided by Bohm. It will now be shown how this can be done.

The derivation of Bell's theorem and the experimental verdict

In the following subchapter, Bell's theorem will be derived for a simple experimental setup. The theorem results from a mathematical inequality that follows from a few assumptions. For the sake of completeness, it must be noted that there are many setups and derivations of Bell's inequality possible. Bell's own proof is more abstract and mathematical than the one that will be demonstrated below, but the core idea and implications are equivalent (Bell J. , 1964). First, the setup and the physics behind it will be described. Secondly, this will be used to derive the theorem.

Mermin's experimental setup

For the explanation of Bell's theorem, I will use the experimental setting portrayed in figure 9 and 10, devised by N. David Mermin. This setup is almost identical to ones that have been used for real experiments and has the benefit of clarity (Mermin, 1981). It utilizes the EPR-paradox in the way it was presented by Bohm and further specifies the details, resulting in what is, in my view, the most accessible method for arriving at Bell's results.

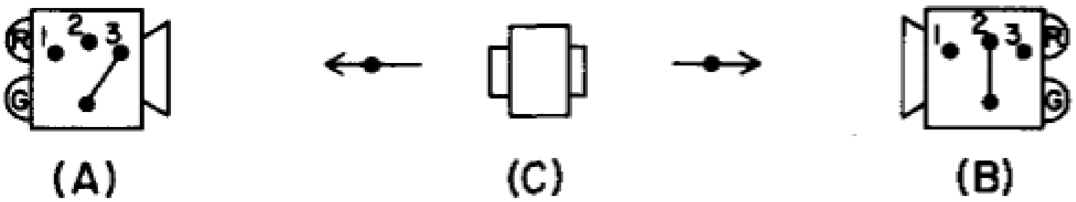


Figure 9: Mermin's realization of Bohm's EPR-setup. Two detectors A and B measure the spin of two entangled particles emitted from a source C. All parts are unconnected (Mermin, 1981, p. 400).

In the above figure, a particle source C contains molecules of spin 0 that can at the press of a button be split into two atoms with nonzero spins, one emitted to the left and the other to the right. The particles then go on their journey to device A and B. These need not be at the same distance and A might well be 10 lightyears away while B is just 1 meter removed from C. Device A and B are the measurement apparatuses, seen more closely in figure 10.

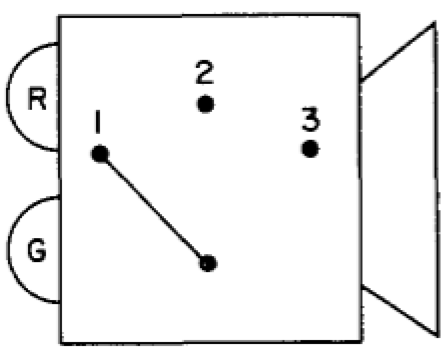


Figure 10: The spin measurement apparatus in the Mermin setup, particularly the one on the left in figure 9. Atoms come in from the right and their spin is chosen to be measured against one of three axes. When 'up' is measured, the detector flashes a red light ('R'). When 'down' is measured, the detector flashes a green light ('G') (Mermin, 1981, p. 399).

When an atom enters the detector, it can be measured against one of three axes. Consider the (z,x)-plane. Option 1 is just the z-axis. For option 2, the axis has been rotated 120º parallel to the plane, and for option 3 this is 240º. Rotating by another 120º just returns one to option 1, so the axes can be conjoined into an equilateral triangle. In reality, this ‘machine’ could be realized by having a Stern-Gerlach setup in the detector for which the orientation of the magnets can be changed. The one of these three settings the device is in can be installed through a button, but it can also be left up to a *randomizer*. For example, one could connect each device to a random number generator producing a real number between 0 and 1. If the number is between 0 and 1/3 it would be set to setting 1 where the spin is measured against the z-axis, if it is between 1/3 and 2/3 setting 2 and 2/3 and 3 will yield setting 3. One final detail is that one may very well decide the measurement setting of one of these devices only after an atom has left the emitter on its journey towards it. In the case that one of the detectors is much further away from the source than the other, it is even possible to set the measurement setting of the farther-away detector after the closer-by detector already measured one of the atoms in the entangled pair.

When an atom’s spin is measured, the red light R on the left side of the detector flashes if the spin is up with respect to the chosen measurement axis, and the green light G flashes if it is down. We can already infer that if the measurement settings on the two detectors are the same, the opposite light will always flash. After all, due to the conservation of angular momentum, if one atom is measured as having spin up with respect to a given axis, the other one must be spin down when measured against that same axis, and vice versa.

Note that in total there are 9 detector configurations. For example, detector A might be on setting 1 while B is on 3 (notation: 13). Or perhaps both are in setting 2 (notation: 22). Of the 9 settings, 3 will always yield opposite lights (namely 11, 22 and 33) as explained in the previous paragraph.

However, one may wonder what happens in the other 6 cases (12, 13, 21, 23, 31, 32), where we measure one atom against one axis and the other against a different one. In this case, given the spin of one particle, it is a matter of probability[27] what the spin of the other one will be. The predictions of quantum mechanics here are as follows. Suppose we have measured one of the atoms to be up with respect to the z-axis, and then have yet to measure the spin of the other atom. This one, however, will be

[27] Whether this is an intrinsic probability or a result from ignorance about the real spin of a particle is a matter of interpretation that we shall get to in a moment.

measured against the x-axis, 90º to the z-axis[28]. Classically, one may think the spin will be 0 in that direction, since it must be fully down in the z-direction given the measurement of the other atom. But one should remember the inadequacy of the classical analogy for the quantum mechanical phenomenon called ‘spin’. Measuring it will always show the spin to be completely up or down in the direction of measurement. Because the x-axis is perpendicular to the z-axis, the spin of the other atom being up in the z-axis says nothing about what the spin of the now measured atom in the x-direction will be. The result is that it is a 50/50 probability of being up or down, again *with respect to the x-axis*. Now suppose that we rotate our measuring axis 30º counterclockwise with respect to the x-axis, which means it makes an angle of 60º with respect to the z-axis[29]. In this situation there is some correlation between the measuring axis and the positive z-direction the other atom was found to be in spin up against, i.e., they are not perpendicular. Quantum mechanics then predicts that this produces statistics where the atom will be spin down 75% of the time[30] but spin up 25% of the time. As a last example: if we rotate the measuring axis for the second atom 180º with respect to the positive z-axis, it is possible to know with certainty again what the spin of the atom would be, as it would always be up. This may seem strange since the other atom was up as well. But consider a clock with a long arm pointing to 12PM and a short arm pointing to 6PM. You might say “the long arm is up, and the short one is down”, and then I say “indeed, the long arm is up, but look, now I rotate the clock by 180º degrees, and the short one is ‘also’ up!”

With this information, the derivation of the theorem can be understood.

From Bell’s inequality to Bell’s theorem

At the beginning of this chapter, it was stated that Bell showed Einstein’s and Bohr’s view to produce different empirical predictions. This can be demonstrated by using that given two assumptions, any hidden-variable theory will yield different statistics than those of quantum theory when doing experiments with the setup previously described.

[28] This is not one of the three settings on our detector, but this is meant to serve as a general description of what happens when entangled spins are measured not with respect to the same axis, but to others.

[29] For everyday visualization: turn the hand of a clock from 3PM to 2PM.

[30] Using the quantum mechanical result that the probability ‘P’ of the spin measurements of the two atoms giving opposite results is $P(\theta) = \cos^2(\theta/2)$, where θ is the angle between the measurement axes (Mermin, 1981, p. 943).

First, the two[31] assumptions will be introduced:

1. Principle of locality: No two physical objects can send information to each other faster than light.
2. Statistical independence: The probability distribution of the hidden variables is independent from the detector settings. This assumption is sometimes also called ‘measurement independence’ or ‘free choice’.

The meaning of the assumptions will become more apparent through the explanation of the hidden variable perspective on the described experiment. It will be shown later that the argument hinges upon these assumptions and will be invalid when these are discarded.

Let us once again consider what can be the reason for the fact that in the Bohm-Mermin setup, it can be the case that if both detectors share the same measurement setting (be they 11, 22 or 33), they will always flash opposite colors when measuring the atoms from the entangled pair. The perhaps most intuitive answer to this question is that at the moment of their creation, the atoms ended up with pre-existing ‘instructions’ of what value to reveal given measurement. Remember that in hidden-variable theories, there is no intrinsic probabilism, but rather there are, as the name implies, elements of reality to these atoms that make it so that they always had a certain spin, momentum, position, etc. independently of being observed. We just do not know what these are yet, but their discovery would reveal a deterministic theory underlying quantum theory, unveiling the ‘randomness’ to just be the result of our ignorance of certain properties. The key here is that the argument about hidden variables we are about to make works regardless of what exactly this hidden variable is. The property that they are the elements of reality missing in our understanding that would allow us to determine an atom’s real state if all that is needed.

Using this insight, let us now introduce the following notation. An atom described by the state ‘GRG’ will flash green when measured in setting 1, red if measured in setting 2 and green if measured in setting 3. It logically follows that there are 8 possible states for an atom upon creation: RRR, RRG, RGR, RGG, GRR, GRG,

[31] Some authors (such as Hossenfelder and Palmer (Hossenfelder & Palmer, 2020)) argue that the locality assumption can be split into two distinct ones (output and parameter independence), while others (like Landsman (Landsman, Indeterminism and Undecidability, 2021, p. 20)) argue that the use of probability theory itself introduces more implicit but nontrivial assumptions for the theorem. Nevertheless, this ‘mainstream’ portrayal of two assumptions will for now be suitable and comprehensive enough for our purposes.

GGR, GGG. Given that we always have opposite flashes when atoms are measured in the same direction, it logically follows that if the atoms have real states, they will always carry the opposite state from one another. If the atom measured by detector A is in state RRG, the atom measured by detector B will be in state GGR. The need for the locality assumption in this deduction already becomes clear. Suppose atom A (defined as the atom measured by detector A) is measured first. If it were to be the case that atom A could upon measurement instantaneously influence atom B, this influence could assure that in the case of equivalent setting, atom B could be 'changed' such that it would flash opposite when itself measured. It would not be necessary to have definite opposite states upon creation already.

Now we turn our attention not to the three situations with equivalent settings (11, 22, 33) for both detectors, but the six situations (12, 13, 21, 23, 31, 32) where the detector settings are different. One could now ask the following question: *what percentage of time do detectors A and B flash opposite colors (i.e., the atoms have opposite spin) when their settings differ when repeating the experiment many times, randomly sampling a setting-configuration?* This question is experimentally decidable.

Before we do so, we can analyze what the hidden-variable framework would predict in the case of hidden variables and quantum mechanics. For each of the six detector settings under consideration, we can determine whether A and B will flash a light of the same or of different color. For example, suppose my measurement settings are '13', and the atom measured by detector A is in the state RGR, with the atom at B consequently having the state GRG. Then detector A flashes red and detector B flashes green. After all, the real instructions carried by atom A are such that it outputs 'R(ed)' when measured in setting 1, and atom B revealing G(reen) when measured in setting 3.

With 6 possible options for which the detector settings are different and 8 possible real hidden-variable states, the above can be done 6x8=48 times. The results of this are shown in the table below:

		12	13	21	23	31	32	
RRR	**GGG**	N	N	N	N	N	N	1
RRG	**GGR**	N	Y	N	Y	Y	Y	1/3
RGR	**GRG**	Y	N	Y	Y	N	Y	1/3
RGG	**GRR**	Y	Y	Y	N	Y	N	1/3
GRR	**RGG**	Y	Y	Y	N	Y	N	1/3
GRG	**RGR**	Y	N	Y	Y	N	Y	1/3
GGR	**RRG**	N	Y	N	Y	Y	Y	1/3
GGG	**RRR**	N	N	N	N	N	N	1

Here, the first column is the state of atom A and the second column is the associated state of atom B. The top row displays the detector settings. The 'Y' stands for 'yes', meaning that both detectors flash the same light in that case. The 'N' thus stands for 'no', which naturally implies both detectors flash different lights. Finally, the last column answers the question we asked and aimed to answer from the perspective of a hidden-variable theory, for all 8 possible states. The combination of these results and our two assumptions is enough to derive the inequality. This is the case because even if we do not know what states atoms are in, and we have no idea what the hidden variables are, the fact that, contrary to the Copenhagen interpretation, these states and hidden variables exist, is enough to derive it. We find that:

$$P_{N,HV} \geq \frac{1}{3}$$

Here, P_N refers to the probability that for any random measurement of an atom in Mermin's setup, both detectors flash different lights ('N'). The subscript 'HV' refers to the fact that this is an equality derived for 'Hidden-Variables'. Put alternatively, $P_{N,HV}$ is the fraction of times 'No' happens when repeating the experiment an infinite number of times[32]. Moreover, the inequality also applies in the case not present in the table where both detectors are in the same measurement setting, as it is then just P(N) = 1, i.e., the lights always flash a different color.

This is an example of a *Bell inequality*. Depending on the setup it can look different, perhaps more abstract, but the essence is the same: from the perspective of a hidden-variable theory and given our two assumptions, the given experimental setup always produces differently colored light flashes at least one third of all measurements. This

[32] This statement does imply an interpretation (frequentism) and the assumptions of probability theory, but this will do for our purposes.

can be seen in the last column of the table, where this is the case 1/3 of the time for most configurations of state and measurement settings and even more so (all the time) for all remaining configurations. That is why we have an inequality rather than an equation.

Bell's inequality, however, is incompatible with the predictions of quantum theory, which was of course viewed as a complete description of physical reality by the Bohr and the Copenhagen school. The math of quantum mechanics shows that $P_{N,QM}(\theta) = \cos^2(\theta/2)$, where θ is the angle between the measurement axes and the subscript 'QM' refers to the fact that this is the derived probability from quantum mechanics (Griffiths, 2014) (Mermin, 1981, p. 943). Since in the cases where the settings are not the same for both detectors this angle is always $\theta = 120°$, we find that $P_{N,QM}(120°) = \cos^2(60°) = \frac{1}{4}$.

Thus, we are left with the striking observation that:

$$P_{N,QM} = \frac{1}{4} < \frac{1}{3} \leq P_{N,HV}$$

Or put more simply:

$$P_{N,QM} \neq P_{N,HV}$$

With P_N testable by the experiment that has been described.

These tests have been done using many different experimental setups, closing more experimental loopholes over the years until none were left (Kaiser D. I., 2022). In October 2022, Alain Aspect, John Clauser and Anton Zeilinger were rewarded the Nobel prize in physics for their key contributions to these experiments (Amos, 2022).

All experiments have found results in perfect *agreement with the predictions of quantum theory*. In the language of the setup we have used above: $P_N = \frac{1}{4}$.

Thus, this allows one to conclusively formulate Bell's theorem as follows:

No local hidden-variable theory obeying statistical independence can reproduce the (correct) results of quantum theory.

The legacy of the theorem

For many in the field, Bell's theorem is taken as proof that quantum theory rules out 'local realism', and that Einstein was wrong and Bohr was right. This is one reason for the popularity of the Copenhagen interpretation as seen in figure 5.

While it is understandable how this sentiment has arisen, now having gone through the derivation of the theorem, we can see that there is a bit more nuance to it, especially to superdeterminists. Firstly, as we will see in chapter 5, there are more 'anti-realist' interpretations of quantum theory than the Copenhagen interpretation. So while one may indeed say that Bohr's view, as opposed to Einstein's, produces the correct prediction for the statistics observed in tests of Bell's inequality, it does not vindicate all of Bohr's interpretive views and the Copenhagen interpretation as such. But secondly, and more importantly for the contents of this thesis, not all realist interpretations of quantum theory are ruled out by Bell's theorem. It has been explicitly stated that the theorem assumes locality and statistical independence. As we will see in chapter 5, De Broglie-Bohm pilot wave theory is an example of a realist framework that rejects the locality assumption (PL1), while superdeterminism does away with statistical independence. Given Einstein's use of the locality principle in many of his arguments we discussed, in addition to his famous rejection of "spooky action at a distance", Einstein would likely not have opted for giving up PL1. This is not something that needs to be guessed, as it is documented that Einstein was not too enthusiastic about this option (Kumar, 2008). It is, however, an open question as to whether he would have been willing to give up statistical independence when confronted with either that or rejecting the realist hidden-variable approach. Since in the remainder of this thesis the focus will be on developments from the second half of the 20th-century onwards, Einstein will no longer be one of its protagonists. However, an interesting side-question to return to in the discussion at the end will be what he would have thought about superdeterminism, a judgment we can make only after a complete evaluation.

At this point one may wonder about an apparent inconsistency between the vindication of Bohr's camp through Bell's theorem, and the important role the assumption of locality has played in the derivation of this result. One may note that back in chapter 1, it was explicitly stated that quantum field theory incorporates special relativity, which among other things claims that no information can travel faster than the speed of light. Yet, as even Bohr had admitted, the EPR-paradox does seem to imply "*…an influence on the very conditions which define the possible types* *of predictions regarding the future behavior of the system"*. If Bell tests[33] show that quantum theoretical predictions are correct, but quantum theory involves instantaneous wave function collapse of an atom hypothetically lightyears away the moment I do a spin measurement on its entangled partner, does quantum (field) theory not violate one of its own core postulates?

This paradox is commonly, although not necessarily uncontroversially[34], resolved by distinguishing between two variants of locality. In fact, while the principle of locality has a strong presence in physics, it is not always clear how it is to be characterized and understood correctly in different physical contexts (Berkovitz, 2007). While much deeper analysis is thus possible, a complete study of the role of locality in different quantum mechanical contexts is beyond the scope of this thesis. Crucially for the difference between PL1 and PL2 here, is how broadly the class of events that constitute an 'influence' is defined.

PL1: No two physical objects can influence each other faster than the speed of light.

PL2: No two physical objects can send information between each another faster than the speed of light.

The key is that special relativity, and by extension quantum field theory, forbids the sending of information faster than light, thus PL2. That means no particles, no gravitational waves, no sound, or anything else that could, in principle, be used to communicate at superluminal speeds. The ability to send information using a Bohm-Mermin like setup would mean that I could use entangled particles to write a message, and signal something instantaneously. If I could predict the spin of my atoms at site A, and previously made a number of agreements (e.g., a binary language) with my colleague at site B, I could instantaneously influence the measuring results they see by doing my own measurements. Under these conditions, faster-than-light communication would in principle be possible, which is incompatible with relativity and quantum field theory. However, due to the random nature of spin wave function collapse, this method cannot be used to signal messages. This is proved by the so-called 'no-communication' or 'no-signaling'

[33] A Bell test is a general experimental test of Bell's inequality. Mermin's setup is one that can easily be translated into a concrete experimental test of the inequality, but there are many other possible setups. Some have more setting options, some have more particles involved, some use polarization rather than spin, etc. Different setups will result in inequalities that look slightly different, an influential example being the CHSH inequality (after Clauser, Horne, Shimony and Holt).
[34] Although this is not fully uncontroversial (Myrvold, Genovese, & Shimony, 2019).

theorem (Berkovitz, 2007). Thus, quantum (field) theory crucially does not violate PL2. It only violates PL1 in the sense that a measurement of an atom's spin at detector A influences something about the other entangled atom, as Bohr and Einstein already noted. Although the influence is instantaneous, it cannot be used to communicate, at which point most physicists shrug off the issue of this 'quantum nonlocality'. A crisis of inconsistency whereby quantum field theory is founded on the type of locality it itself does not adhere to is thereby generally accepted to be averted. But the question as to what the nature of the 'influence' in PL1 is, and if really makes its presence fully unproblematic, is thereby, however, unanswered. It is therefore not the last time the issue will be mentioned (Berkovitz, 2007).

Another set of very important questions not explicitly answered yet is then: how do the rejection of locality in the physically crucial sense of PL2 as it was assumed for Bell's theorem one the one hand, and statistical independence on the other, allow one to formulate a realist hidden-variable theory compatible with the predictions of quantum mechanics after all? Answering this is possible through the now gained understanding of why Bell's theorem is true, and because this constitutes its origin it will be an important part of chapter 6 when we talk about the motivation for superdeterminism.

Before continuing, it should briefly be noted that Bell's theorem has inspired many 'no-go theorems' after it. As the name implies, these are theorems characterized by the fact that they show that something in physics is *not* possible. They must therefore be taken into account by all interpretations of quantum mechanics, including superdeterminism when relevant. In the following, I will very briefly state the conclusions of some no-go theorems important for quantum foundations.

- Leggett's inequality: constrains hidden-variable theories further than Bell, by also ruling out a subset of nonlocal hidden-variable theories (Kumar, 2008, p. 354).
- The Bell-Kochen-Specker Theorem: states that deterministic hidden-variable theories that attempt to exactly reproduce the results of quantum theory must be contextual. Quantum contextuality refers to the fact that values of physical quantities in quantum mechanics depend on which other (non-commuting) quantities are measured at the same time, which is not the case for classical mechanics. In other words, the way in which one tries to measure a particular quantity of a quantum system, is not always independent of the value that is revealed when that measurement is performed (Budroni & al., 2022).

- The Pusey-Barrett-Rudolph (PBR) Theorem: states that, for similar assumptions as those of Bell’s theorem, hidden-variable theories that attempt to exactly reproduce the results of quantum theory must be psi-ontic (Rizzi, 2018).
- The Free Will Theorem: states that the conjunction of determinism, PL2-locality, statistical independence and key quantum mechanical results such as the Born rule are incompatible (Conway & Kochen, 2006) (Landsman, Foundations of Quantum Theory, 2017, pp. 202-204). Many interpretations of quantum mechanics deal with this by giving up on determinism, but this is not the only option.
- Landsman’s Theorem: no truly deterministic hidden-variable theory is compatible with Born’s rule (Landsman, Randomness? What randomness?, 2019). This theorem will be important in chapter 9.

With the knowledge of Bell’s theorem (and others), it is now possible to see why superdeterminism is being discussed. It is the only local hidden-variable framework that cannot be ruled out. It does so, however, by giving up statistical independence. Why, and what this leads to, will be extensively discussed from chapter 6 onwards.

Chapter 4: Interpreting quantum theory

The previous chapters have introduced the history, content and philosophical issues of quantum theory insofar required to understand interpretations of quantum mechanics. But as was hinted at in the introduction, it is not always straightforward whether to call a given framework an ‘interpretation’, and this certainly also holds for superdeterminism. Some conceptual analysis of this issue might help to provide a clear picture of what there is to interpret, what an interpretation must provide and whether superdeterminism can actually be classified as such. The goal of this chapter is to answer these three questions.

The meaning of an interpretation

Discussing interpretations of quantum theory, and the very fact that superdeterminism is often mentioned in that list, thus begs the question: what even *is* such an interpretation?

There is no consensus on this question. Physicist and philosopher Tim Maudlin claims that we should not be talking about ‘interpretations’ of quantum mechanics in the first place. To him, the essence of any physical theory starts with describing what *is*, that is, providing an ontology. But it is precisely matters of ontology that are an important part of where we find the differences in so-called ‘interpretations’ of quantum theory. Thus, according to Maudlin, we should refer to these as different *physical theories* rather than *interpretations* (Maudlin, 2022, p. 02:20:18).

One contrasting rigorous attempt to answer this question has been formulated by philosopher of natural science F.A. Muller. In his paper on the topic, he claims that:

“To provide an interpretation of quantum mechanics is to add postulates to those of ‘minimal quantum mechanics’ (QM_0) so as to provide answers to questions about physical reality, that we deem meaningful and that pertain to physical systems falling within the purview of QM_0, extending QM_0 may very well involve changing and usually extending its sparse vocabulary.” (Muller F. , 2014, p. 14)

Here, QM_0 is understood as the set of postulates enabling one to calculate no more or less than measurement outcomes and their probabilities accurately.

In this conception, interpreting quantum theory is more than just the intuition of ‘assigning meaning to something’. Muller lates gives a step-by-step algorithm for how any attempt at an interpretation is to do this. At its core, any set of coherent ideas that can be called an ‘interpretation’ of quantum mechanics should accept a minimum body of empirical truths described by the formalism and then build an ontology upon that. As we will see in the next chapter, on this view most popular ways of looking at quantum mechanics can indeed be called an 'interpretation' thereof.

In the remainder of this thesis, I will use 'interpretation of quantum mechanics' in the Mullerian sense. The first reason for this is Muller's rigid analysis of the concept, which seemingly frees it from any inconsistencies or grey areas. The second reason is that it is both meaningful and intuitively sensible language to speak of a different physical theory from quantum mechanics if it actually makes diverging empirical predictions in given physical situations, while referring to an 'interpretation' if the empirically adequate core remains unchanged but views on how to understand the mathematical formalism producing those differ. An example would be that we do not refer to special relativity as an interpretation of classical mechanics, even though the former reproduces the latter in a particular limiting case. Special relativity can be used to make accurate predictions about more phenomena, its empirical core is clearly different. At the same time, the many-worlds interpretation of quantum theory (see next chapter) is always referred to as an interpretation. Its empirical content is equivalent to QM_0 and competitors such as the Copenhagen interpretation, but it makes different claims on what it is the theory says about reality. One final note on the use of 'theory' here is that this term is usually reserved to a consistent set of ideas offering a scientific explanation, that is already well-tested. This is, of course, not currently the case for superdeterminism. It is now more so a collection of models, many of which do not yet incorporate all relevant details. Nevertheless, as will be seen below, the distinction between a model and a theory as used here is still useful. Moreover, some other alternative theories to or theoretical extensions of quantum mechanics seen in the next chapter use 'theory' in this way all the time, even though they cannot claim to be well-tested either. Examples are objective-collapse theories and pilot wave theory.

With these views in mind, how can superdeterminism be classified? One thing already made clear in the introduction is that superdeterminism is a class of hidden-variable theories or interpretations that share the central tenet of rejecting the 'statistical independence' assumption made in the derivation of Bell's theorem. Thus, it may be the case some specific superdeterministic theories may be called 'interpretations' while others may be called 'physical theories'. This is exactly what we find when we ask prominent superdeterminists. For example, 't Hooft explicitly refers to this Cellular Automaton model as an "interpretation" of quantum mechanics, and claims that this superdeterministic framework can indeed reproduce all empirical predictions of quantum mechanics such as the Born rule ('t Hooft, The Cellular Automaton Interpretation of Quantum Mechanics, 2016). In contrast, Hossenfelder refers to her Future-Bounded Path Integral model as an effort to produce an actual physical theory, rather than an interpretation. This theory, in the language of Muller, modifies QM_0 and explicitly claims to contain situations where empirical predictions between the two diverge. Quantum mechanics is only reproduced as a limiting case of this underlying hidden-variable theory, but outside of the limiting scope they predict different experimental outcomes (Hossenfelder S. , Superdeterminism: A Guide for the Perplexed, 2020, p. 7). For Maudlin, however, both would classify as physical theories.

Lastly, one sometimes also comes across terms like ‘model’ and ‘framework’, and these have already been used in this thesis from time to time. A model can broadly be viewed as any representation of a system, in physics often some demarcated set of (mathematical) rules that allow for calculations. The model need not be a fully accurate representation of reality, but it can be concretely worked with to produce results with utility to its user. For example, there are multiple models of superdeterminism with different mathematical rules with which they describe physical systems, sometimes even leading to different predictions as in the example of ‘t Hooft and Hossenfelder in the previous paragraph. In turn, ‘framework’ can be seen as an umbrella term referring to any conceptual structure, here that of superdeterminism.

In summary, there exist superdeterministic *theories* that make different predictions from QM_0, and superdeterministic *interpretations* that aim to reproduce the predictions of QM_0, each with several possible *models* to do so, but all of this falls under the *framework* of superdeterminism. Graphically, the distinction between these terms is shown in figure 11. With each crossing of a grey line, a new subset of the set of superdeterministic models is entered. Returning to our initial question on whether superdeterminism is an interpretation of quantum mechanics, we now see that the answer depends on the superdeterministic model in question. At the same time, we have seen that what an interpretation of quantum mechanics is, is a matter of controversy in the first place, so one’s conception here also influences one’s answer to the question. In the remainder of this thesis, I will however use Muller’s approach and consider whether a model modifies QM_0 as a standard of whether to refer to it as an interpretation or physical theory.

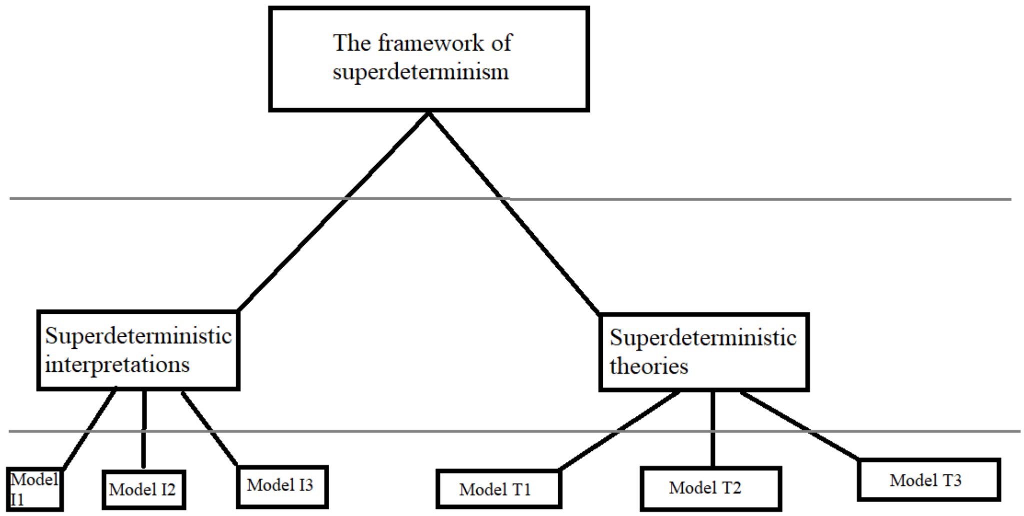


Figure 11: A visualized conceptual scheme for superdeterminism. The upper grey line divides the set of all superdeterministic models in the subset of theories and the subset of interpretations. The second grey line further divides these subsets in one-element sets describing individual models.

The interpretive issues in a nutshell

In the previous chapters the main interpretive issues of quantum theory have been introduced. Interpretations of (and alternatives to) quantum theory usually attempt to address these from their own perspective. This brief subchapter exists to summarize the main issues we explored in a structured, seven bullet point list of questions for all frameworks to answer. While not an exhaustive representation of the entire subject area of the interpretations of quantum mechanics, this list was formulated with the intent of, to a significant degree, characterizing all the contenders in this great debate. This will be done extensively for superdeterminism in chapter 6. The list is as follows:

- What is the framework's reply to the measurement problem? That is, what does it say about how, if and when collapse happens? Are detectors reducible to quantum physics or do they play a separate role?
- If applicable, is the framework affected or constrained by no-go theorems such as Bell's?
- Is it deterministic or probabilistic?

- Is it realist in the sense that it always ascribes definite physical quantities to all quantum particles, irrespective of the system being measured or not? Or is it anti-realist in the sense that they, prior to measurement, are in superpositions of distinct possible outcomes?
- What is its view on the nature of the wave function?
- Is quantum theory viewed as being a complete description of nature? Are the foundations of the theory here to stay?
- Is the framework local, and in what sense?

In the next chapter, some popular non-superdeterministic interpretations of quantum theory will be briefly reviewed. What are they about, what is their view on some of the issues above and what are possible criticisms leading some to superdeterminism instead?

Chapter 5: Popular frameworks and their alleged shortcomings

In previous chapters, the core ideas of the Copenhagen interpretation and the general hidden-variable framework have already been discussed. Furthermore, we have learned that superdeterminism is a class of hidden-variable models that reject the statistical independence assumption in Bell's theorem. In the next chapter, superdeterminism and the meaning of this rejection will be covered extensively. For now, some other interpretations and theories, as well as their challenges will be briefly introduced. The point here is not to bring forth the full picture of these frameworks, but to have some basic understanding of superdeterminism's 'competitors' and what reasons one may have to be critical of them. Moreover, some arguments for or against superdeterminism refer to other frameworks in comparison, and understanding these will aid in evaluating the viability of superdeterminism in the broader context. While there are certainly more interpretations of quantum mechanics out there, three well-known ones will be covered, namely objective-collapse theories, the Everett (or 'many-worlds') interpretation and De Broglie-Bohm pilot wave theory.

According to philosopher and physicist Tim Maudlin, all three of these take a different one of three possible answers to the measurement problem (Maudlin, 2022, p. 02:08:36). This is, according to him, a result of the following logically necessary statement:

$$(P_1 \wedge P_2 \wedge P_3)$$

$$\Rightarrow \perp$$

Here,
P_1 = "The evolution of the wave function is fully determined by the Schrödinger equation."
P_2 = "Wave function collapse is a real, physical process that is triggered upon measurement."
P_3 = "The wave function provides a complete description of any quantum mechanical system."

The idea is that since we are confronted with *definite measurement outcomes* and a wave function that supposedly tells one *all there is to say* about the physical system, that wave function being *determined by the Schrödinger equation that allows for superposition of states*, we are confronted with a contradiction. That contradiction lies at the heart of the measurement problem that these three premises, often associated with the Copenhagen interpretation, lead to. For structure's sake, each subchapter will be titled using what premise it denies.

¬P_1: Objective-collapse theories

Objective-collapse theories are first and foremost attempts to solve the measurement problem by adding something to the quantum mechanical formalism that can explain the collapse process. They propose that the wave function is a complete description of any physical system, but that the value of the wave function of a system at a point in space and time is not always determined by just the Schrödinger equation. They also hold that the wave function really corresponds to something in nature, and that collapse happening upon measurement is thus a physical process. Different objective-collapse models then each propose their own physical mechanism that causes collapse. This mechanism is represented as a mathematical extension to the Schrödinger equation[35] that 'takes over' when the described collapse conditions are realized. Such an objective mechanism would then answer all the question contained in the measurement problem, since it would then be clear when a wave function collapses and why superpositions are not observed in macroscopic conditions (Ghirardi & Bassi, 2020).

There are several objective-collapse models. For example, the Ghihardi-Rimini-Weber model was the first objective-collapse theory, introducing a collapse rate and localization distance to the math of quantum mechanics that describe random collapses. Another well-known model is the Dióse-Penrose model. This model proposes that the fluctuations of the one force that is not described by a quantum field theory, namely gravity, is responsible for wave function collapse. Penrose argues that this is a sensible approach since quantum field theory has not yet been combined with general relativity, and that in this eventual union the answer to the measurement problem may likely be entailed. Since these models modify the Schrödinger equation, they also claim to make testable predictions diverging from standard quantum theory (Ghirardi & Bassi, 2020).

Not everyone is sold on the objective-collapse framework. Some (but not all) common criticisms of it are that it is still indeterministic, that it would be incompatible with the locality condition required by special relativity and that it would violate the law of conservation of energy. There is also the more unique 'tails problem'. The wave function can extend out very far in space, even though its amplitude might greatly fall off at such distances. Think of a Gaussian curve sharply

[35] For the mathematically-minded: this is done by adding nonlinear terms to the Schrödinger equation.

peaking at the origin, which can be thought of as having two 'tails' on its left and right end. In many collapse models, this tail is still there after the localization process of the wave function upon collapse, which is problematic for reasons already discussed in the single-slit experiment in chapter 2. A system cannot be said to localize if it still has nonzero amplitude away from its center, however small it may be (Ghirardi & Bassi, 2020).

As can be seen, criticisms can be levied on the basis of physics and that of philosophy, an example of the latter being indeterminacy. As will be explored later, advocates of superdeterminism like Hossenfelder and Palmer, however, would argue that there are also physical reasons not to give up on determinism (Hossenfelder & Palmer, 2020, p. 3). Whether one may think such criticisms legitimate or not, it may seem that plenty of people believe that there are reasons to look for interpretations of quantum mechanics outside of objective-collapse theories.

$\neg P_2$: The Everett interpretation

The Everett interpretation, often referred to as the 'many-worlds' interpretation, has recently gained quite some attention and advocates. It was introduced by Hugh Everett (1930-1982) in 1957. Like objective-collapse theories, it is first and foremost an attempt to solve the measurement problem but in doing so answers other interpretive problems of quantum theory as well.

In the Everett interpretation, rather than attempting to discover how collapse works, collapse is thought not to happen in the first place. It is therefore often promoted as a simplification of quantum theory, being in line with Ockham's razor as it assumes the existence of fewer entities (here: collapse) to explain the same phenomena. Without collapse, there is no measurement problem. There is only the wave function evolving through the Schrödinger equation. The obvious next question is how the interpretation then explains measurement outcomes. Everett's answer is that every possibility in a superposition actually happens. Every time a measurement is made, the world splits into multiple worlds. We just happen to be on one of these world-splitting 'branches', among many existing ones. In the example of Schrödinger's cat, we can say that upon measurement the world branches into two. In one world, the cat is dead, while being alive in the other. One may even consider the wave function of the entire universe, being an objectively real entity that describes everything that can happen, and all these things do happen in different worlds. Everett concluded from this that there exists an uncountably infinite number of worlds (Werner, 1964). Moreover, since in Everett's view there is no probabilistic collapse but only a deterministic Schrödinger equation that dictates the evolution of this universal wave function, his interpretation restores determinism (Vaidman, 2021).

The Everett interpretation is certainly not without its criticism. On a metaphysical level, the existence of an uncountably infinite number of worlds may be some heavy ontological baggage to accept. On an epistemological note, this baggage might also turn the tables on the invocation of Ockham's razor by Everett supporters. They may not need a collapse postulate, but when it comes to ontology, assuming an infinite number of universes may not necessarily constitute the scientific explanation introducing the lowest number of entities. Another argument in the epistemological category could be that the Everett interpretation is unfalsifiable, since in contrast to the other two lines of thinking in this chapter, it does not predict empirical divergence from standard quantum theory. At the same time, it is an interpretation of quantum theory just like Copenhagen. This criticism is thus not unique to it, and since in addition 'QM_0' needs an interpretation in any case, one may wonder if rejecting the Everett interpretation on grounds of unfalsifiability is a fair criticism (Vaidman, 2021).

On a physical level, one important critique is that it is not clear what the probabilities provided by quantum mechanics mean in the context of the many-worlds interpretation. If for every possibility one world is created, what do these empirically supported probabilities refer to? If more worlds are created for more probable outcomes, what happens if one outcome is, say, 4.2395 times as likely as another? Do we get 4 and 0.2395 worlds? And what if we move from finite discrete outcomes as in the case of spin, to a continuous variable, as in the case of position measurements? In addition, one may still ask when exactly the branching process happens, such that there is arguably still a measurement problem around. This argument is made by Hossenfelder (Hossenfelder S. , The Trouble with Many Worlds , 2019). Moreover, the criticisms on the basis of the law of conservation of energy applies here as well: if I open the room and the cat is alive, where does the energy come from to 'create' the world where the cat is dead as well (Vaidman, 2021)?

Thus, there are also exist plenty of reasons people give for not adhering to the Everett interpretation.

$\neg P_3$: De Broglie-Bohm pilot wave theory

Lastly, we look at the De Broglie-Bohm pilot wave theory, sometimes just called 'pilot wave theory' or 'Bohmian mechanics'. These are the same De Broglie and

Bohm that have been mentioned before in different contexts. Pilot wave theory was introduced by de Broglie at the 1927 Solvay conference, but with little support he abandoned the idea later. It was independently reinvented and improved by Bohm in 1952, but again it received little support and reception[36]. Popularity only increased late 20th century, in part due to Bell's sympathetic attitude towards it (Kumar, 2008, pp. 335-336).

Like superdeterminism, pilot wave theory rejects the idea that the wave function is a complete description of physical systems. It is a deterministic hidden-variable theory in which no measurement problem exists, because quantum particles already have definite properties at all times.

The ontology of pilot wave theory is twofold: there are particles and there is a guiding wave that 'pilots' the particle evolution. There is a mutual interdependence here, since while the guiding wave determines the particle motion through time, the configuration of those particles in turn determines what the guiding wave looks like. From the math of pilot wave theory, Born's probability rule is a derived result rather than an axiom. The guiding wave is a kind of wave that makes it so that you find particles in positions in agreement with what Born's rule would tell you. When this is the case, the system is said to be in 'quantum equilibrium', and the math and empirical predictions of pilot wave theory are indistinguishable from standard quantum mechanics. However, there may be cases where there is not an equilibrium because the system has not yet settled for one, and the theory would yield different predictions from standard quantum theory. This would solidify pilot wave theory as a real alternative theory rather than just an interpretation (Goldstein, 2021).

The dual ontology and mathematical model of pilot wave theory recreates, sometimes surprisingly, some key quantum effects. The best way to get a visual picture of this is not a single image, but a video showing a classical analogue of the motion of the particles. An excellent example of this can be found on the Youtube Channel 'Veritasium' and it is recommended for readers looking for a more visual understanding of pilot wave theory (Muller D. , 2016).

The De Broglie-Bohm pilot wave theory brings a very clear ontology to the table, but also has opponents. Metaphysically, it has a more extensive ontology than many other interpretations, because it adds a guiding wave to it in addition to just particles.

[36] It is often stated that this had to do with Bohm's communist affiliations at the time of red scare America (Peat, 1997, p. 133).

Also, in the math of the theory, position as a physical variable has a preferred role to momentum, leading to the justification that in reality we can only perform position measurements in the first place. Philosophers of science may wonder whether it is indeed true that all measurements in physics come down to position measurements, although there is an interesting debate on the philosophy of measurement to be had there (Goldstein, 2021).

Physically, the biggest problem for pilot wave theory is that it is seemingly incompatible with relativity. In the terminology of chapter 3, relativity assumes PL2 rather than PL1. The guiding wave depends on the present configuration of all particles in the universe, which implies faster than light signaling, strongly violating locality even in this PL2 sense. This means that the theory cannot reproduce the empirical successes of quantum field theory. Another reason for the latter is that it also has trouble accounting for the creation and annihilation of particles, a strength quantum field theory was explained to have over quantum mechanics (Goldstein, 2021).

Therefore, again, the situation is such that there may exist reasons for people to look for alternatives to pilot wave theory. On a more philosophical level, no framework seems to grant justified belief in the classical world of realism, determinism, locality[37] and reductionism[38]. The one exception to this is superdeterminism, but at the cost of statistical independence. How this is done, what this means, and if this cost is worth it will be the subject of the remainder of this thesis.

[37] In truth, classical physics was also nonlocal because Newtonian gravity includes action at a distance. The general theory of relativity does successfully preclude this, as even gravitational waves cannot travel faster than light. To be fully accurate, it should therefore be said instead that classical physics + relativistic extensions accommodate a belief in realism, determinism, locality and reductionism.

[38] Reductionism is the idea that a system can, in principle, be explained entirely through its constituent parts. In the case of the Copenhagen interpretation, we showed that as part of the measurement problem, reductionism seems to be in trouble as it has difficulty accounting for detectors. A framework that can do this would restore this reductionistic principle otherwise common in physics, leaving debates on mental phenomena aside here.

Chapter 6: Superdeterminism and statistical (in)dependence

This chapter will be about developing a detailed understanding of what superdeterminism is. First, building on chapter 3, it will be explained how exactly Bell's theorem can be circumvented by rejecting statistical independence. Next, I will show clearly what the meaning of this abstractly formulated assumption is, which will make it far easier to introduce the critiques of superdeterminism in chapter 8. Lastly, with all the pieces in check, the framework itself will be clearly defined and subjected to the interpretive questions posed in chapter 4.

How statistical dependence avoids Bell's theorem

In chapter 3 it was derived that $P_{N,HV} \geq 1/3$, using the table resulting from assuming hidden variables. Repeating the experiment described by Mermin many times, one will always find that at least one-third of the time, the two detectors in the setup will flash a different light. The empirically verified prediction from quantum mechanics, however, was that $P_{N,QM} = 1/4$, hence we arrived at Bell's theorem. One way to get around this conclusion from the hidden variable point of view is the idea that if one atom is measured, it instantaneously influences the state of the other one such that a measurement where the lights of both detectors flashed differently would instead flash the same color. Since it helps to confirm the reasoning that follows in this paragraph as well as the reasoning a few paragraphs later, the table from chapter 3 has been inserted as a reminder below. As an example, suppose the detector settings are 12. The state of atom A is assumed to be RRG and thus the one of atom B is GGR. Also suppose that detector A is closer to the atom source than detector B. Given these configurations, atom A will cause detector A to flash 'red' here, and if no communication between the atoms is possible, atom B will soon cause detector B to flash 'green'. However, since the frequency of differently colored flashings must be brought down to 1/4 in order for the hidden variable approach to be compatible with the data, atom A may instantaneously influence the state of atom B, changing it to, for example, GRR. Now, the detectors would flash the same light. This is why the locality assumption was necessary for the derivation of Bell's theorem. Without it, the theorem can be circumvented by postulating that the measurement of one atom causes it to instantaneously change the state of the other, which can be exploited to create a model which yields the 'correct' statistics. While this may sound quite ad hoc on first reading, this is the approach of de Broglie-Bohm pilot wave theory, whose adherents attempt to provide a solid theory to explain why this happens.

		12	13	21	23	31	32	
RRR	**GGG**	N	N	N	N	N	N	1
RRG	**GGR**	N	Y	N	Y	Y	Y	1/3
RGR	**GRG**	Y	N	Y	Y	N	Y	1/3
RGG	**GRR**	Y	Y	Y	N	Y	N	1/3
GRR	**RGG**	Y	Y	Y	N	Y	N	1/3
GRG	**RGR**	Y	N	Y	Y	N	Y	1/3
GGR	**RRG**	N	Y	N	Y	Y	Y	1/3
GGG	**RRR**	N	N	N	N	N	N	1

Other than locality, the derivation of the inequality assumed statistical independence, which, as mentioned before, is sometimes also called 'measurement independence' or 'free choice'. The formulation of this assumption did not appear explicitly in Bell's initial paper[39], but was first noted by Abner Shimony (1928-2015) and others (Shimony, Horne, & Clauser, 1976), after which mathematical physicist Carl H. Brans (1935) first developed a local hidden-variable model rejecting the assumption and re-evaluating Bell's theorem (Brans, 1987). Statistical independence is a condition defined with respect to two variables. Statistically independent variables are uncorrelated, while statistically dependent variables in some way are. Therefore, when superdeterminism is stated to reject statistical independence, it is specifically referring to that of *the probability distribution of the hidden variables* and the *settings of the detectors* in Bell-type tests. It should be noted that on its own, there is nothing special about variables being statistically dependent. Nature is full of variables between which that is the case. My body's acceleration towards the Earth when I jump, is statistically dependent on the Earth's gravitational field, because the Newtonian law of universal gravitation states that the Earth's gravitational field *directly causes* a downward force on my body. The monthly ice cream consumption in a country is statistically dependent on the number of hours spent in swimming water in the country that week, because these two phenomena are *correlated due to the common cause* of seasonal temperature affecting human decisions. But in the case of superdeterminism, statistical dependence relations are assumed that, for reasons explained shortly, do not have such simple and uncontroversial explanations.

Mathematically, statistical independence in this context can be written as follows:

[39] Although he commented on it later in a separate paper, as we will come to.

$$\rho(\lambda|D) = \rho(\lambda)$$

Here, $\rho(\lambda)$ refers to the probability distribution of the hidden variables denoted by λ. What exactly the hidden variable is depends on the superdeterministic model, and even there it is not always equally clear. Crucially, though, they allow for the specification of a real state underlying any quantum mechanical system. If this seems abstract, it may be helpful the keep the chapter 2 example of thermodynamics and statistical mechanics in mind. There, the velocities of microscopic particles in a gas where the source of and explain the behavior of macroscopic quantities of the gas like pressure, temperature and volume.

Being a probability distribution, $\rho(\lambda)$ represents how much a specific value of the hidden variable quantity occurs when considering many quantum particles. Keeping with the previous analogy, this can be compared to the distribution of velocities of the many particles making up a gas. Note that the mathematical concept of a 'probability distribution' has nothing to do with the notion of intrinsic probabilism, which superdeterminism rejects.

On the other side of the inequality, $\rho(\lambda|D)$ refers to the probability distribution of the hidden variables *given detector settings* D. The line stands for a conditional probability. In Mermin's experimental setup used in chapter 3, D would consist of $\vec{a}$ and $\vec{b}$ if those are taken as the orientations of the axes the spin is chosen to be measured against. When, as superdeterminists proclaim, statistical independence is violated, the equality sign becomes an inequality sign. Thus, the probability distribution of the hidden variables and the detector settings are correlated.

Consequently, when statistical independence holds, it is *not* the case that some states (determined by λ) may appear more frequently given specific detector settings, and vice versa. If the latter were to be the case, there could, for example, exist a situation where if atom A is in state RRG, the measurement setting 12 and 21 occur less frequently, as for these settings opposite-colored flashes would ensue. This can be checked with the table. The other way around, it may be that if the detectors are in configuration 31, atom A will be less likely to be in one of the following states: RRR, RGR, GRG, GGG. These states, after all, cause different flashes when the detector configuration is 31. Assuming that these correlations between the hidden variable distribution of the atoms and the detector settings exist, i.e., rejecting statistical independence, Bell's inequality could conceivably be brought down to be in agreement with quantum mechanics, as just like in the earlier case where locality was rejected. Thus, a local hidden-variable framework rejecting statistical independence can be made to reproduce the empirically verified measurement statistics of quantum mechanics and thereby get around Bell's theorem. That is how superdeterminism 'gets the job done'.

Understanding statistical (in)dependence

The correlations that result from rejecting statistical independence are far from obvious to assume. As we shall see in chapter 8, most, if not all, critiques against superdeterminism stem from its dismissal of statistical independence. Before proceeding, it is therefore important to get a firm grasp on the concept, so these arguments can later be understood more easily.

Statistical independence assumptions of the kind rejected by superdeterminism are generally thought to be very important, and they are ubiquitous in science (Dattani, 2022). An (albeit idealized) example outside of the quantum realm might aid in understanding why. If I want to investigate the effect of eating 10 hamburgers a day on the probability of developing heart disease, I could realistically[40] go about it by doing an experiment with two large groups, one eating the hamburgers daily and the other being the control group. Subsequently, the fraction of both groups developing heart disease later in their lives can be documented. Any such experiment, however, works because we assume the groups to be statistically independent. Here, the 'state of the quantum atom' and the 'angle of the Stern-Gerlach apparatus I use to measure its spin against' from Mermin's experiment could translate to, respectively, 'the genetic heart disease disposition of a person in a group' and 'whether I assign a person to the hamburger or control group'. If I assume statistical independence between these variables, there is no correlation between the 'state' (i.e., genetic predisposition to heart disease) of the group members and the setting I use to measure them (i.e., subdividing them into the control group or experimental group). Given they are large enough, the groups are then statistically equivalent in every way except what group they have been assigned to. We can infer that if we see a higher fraction of group members in the hamburger-eating group suffering from heart disease later, this must be due to the only variable changed for this group in contrast to the other: the eating of the many hamburgers each day. If, however, we cannot assume statistical independence, I cannot draw this conclusion. This would mean that there already exists a correlation between the 'state' of the group members and which of the two groups I sign them up for. It would be impossible for me to subdivide the group truly randomly. Upon understanding the role statistical

[40] Not morally.

independence relations of this kind play in science in general, one might see potential problems with not having these. While this foreshadows chapter 8, the goal for now is to understand what statistical (in)dependence means and does.

Postulating that correlations between the states of quantum particles and the detector settings used to measure them exist, as in the case as the Mermin experiment used before, is one thing. So far, it is akin to a mathematical trick. A subsequent physical question is where these correlations could come from. Superdeterminism saves an observer-independent reality from the restrictions of Bell's theorem by giving up statistical independence rather than locality. Thus, it makes sense to keep locality intact when making a superdeterministic model. It logically follows that the correlations in question *either* must have been created when the quantum particle and whatever 'chooses' the detector setting were still in causal contact[41], *or* they must be a direct consequence of the initial conditions of the universe at the big bang. Indrajit Sen and Antony Valentini, two physicists working on quantum foundations, differentiate between type I and type II superdeterministic models based on this distinction, the former being type I and the latter type II (Sen & Valentini, Superdeterministic hidden-variables models I: nonequilibrium and signalling, 2020, pp. 2-3). Below, we will first consider the meaning of a type I superdeterministic model for what our universe is like, to then move on to a type II model.

The phrase 'causal contact' for type I models can roughly be taken to mean that somewhere in the past, interactions respecting locality must have taken place between these systems. The implications of this can be nicely demonstrated through the cosmic Bell test. This was an experiment performed quite recently by a group of scientists (Rauch et al., 2018). In brief, the team did a series of Bell tests using the random polarization of light of distant galactic events (quasars) to determine the measurement settings of detectors in a Mermin-like setup. This can be done by writing an algorithm that connects, e.g., vertically polarized photons to a specific detector setting. Due to the large distance of these quasars the light used to set detector settings must have left the quasars a long time ago, billions of years in this experiment. However, they found that using this light, Bell's inequality was still violated. From the perspective of superdeterminism, that must mean the correlations between the detector-setting photon polarizations of the quasars and the spin (determined by λ) of the quantum particles being measured on Earth have been

[41] This concept is most clearly explained through the use of 'light cones', but in order to minimize the introduction of technical concepts I will go for a less-rigorous heuristic explanation.

created a long time ago. The team found that this must at least have been around eight billion years ago, which is over half the age of the universe.

It is then not unlikely that using increasingly older light, as has been done, we may eventually be led to the conclusion that all these correlations must already have been created during the very early universe, shortly after the big bang: the beginning of our universe where all matter and energy content started out infinitesimally close together. If one is a type I superdeterminist, this then likely entails the view that already in the very early universe, causal interactions between systems took place resulting in, effectively, a set of 'instructions'. These instructions consist in correlations between system that make it so that Bell tests give us the 'right' results, which arguably has significant implications for how we understand our universe. A visual representation of this can be found in figure 12.

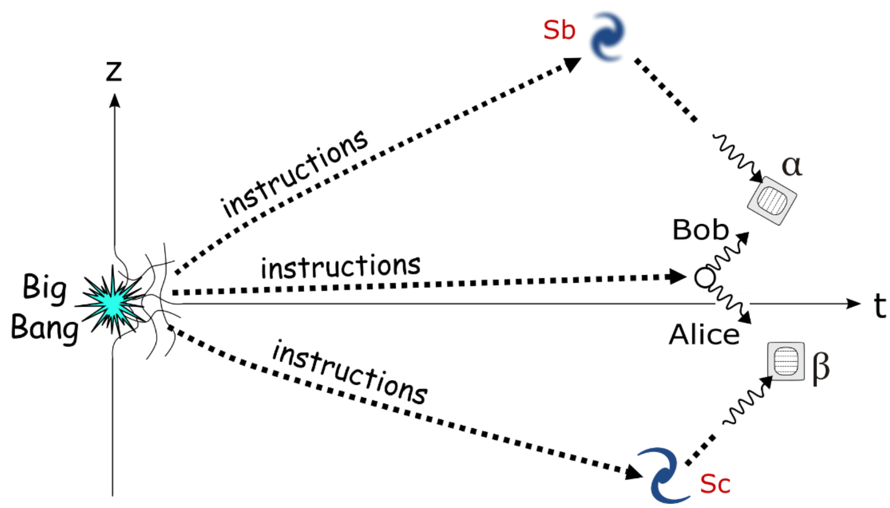


Figure 12: A visualization of correlations present in the universe from the big bang onwards (Dalton, 2021).

In the figure, we see a galaxy 'Sb' and 'Sc' sending out light determining measurement settings α and β in a Mermin-like setup. After this, two entangled atoms 'Bob' and 'Alice' arrive at the detectors. The horizontal axis represents time, while the vertical axis represents the distance in space. At the beginning of time, systems influenced each other such that this resulted in instructions that were sent out both to the atoms[42] with their real state and to the photons in the galaxies determining the measurement settings.

The situation is slightly different for type II models. In general, the evolution of a deterministic physical system can be predicted from the combination of causal laws and initial conditions. To make it more concrete through a classical example: if I throw a ball up in the air and it falls down, I both need Newton's laws of motion and gravitation as well as data on the initial speed the ball was thrown up with, the height from the ground at which I threw it, etc., to calculate at what time it will hit the ground again. If I want the ball to hit the ground 5 seconds after I throw it up in the air, either I need to make sure it leaves my hands at the right position and velocity for it do so, or I must wish that the laws of mechanics had been different, or there had been more laws, just in the way that my goal with the ball is achieved. While in the type I model just considered the correlations are a result of a yet unknown law making it so that quantum systems influence each other in a way through which 'the right correlations' are established, in a type II model, there are no such causal interactions. Rather, the universe, like the ball, also has initial conditions. Just like I need to change those of the ball to get the desired result, the initial conditions of the universe are required to be such that the superdeterministic correlations will come out. Note that the example breaks down in that we cannot change the laws of mechanics, but we can change the initial conditions with which the ball is thrown. In the superdeterministic case, we have no control over either the laws of nature or the initial conditions of the universe, such that it is not a given that one of the two must be more likely. In any case, the conclusion is as follows: in a type II superdeterministic model, one may imagine that all particles in the universe start out with an initial position and velocity[43] just so that they will develop in a way that we see the correlations necessary to violate Bell inequalities in our experiments ('just so that the ball hits the ground after 5 seconds of throwing it up'). Note then that figure

[42] One may wonder how the instructions could, in the past, have been given to the atoms when the Mermin experiment describes how they are created from a molecule, to be measured in the present. The idea here is that the molecule also had these instructions and simply carried them over to the atoms. Given that molecules were not around shortly after the big bang, the molecule itself likewise must have had its instructions carried over from whatever systems combined to constitute the molecule. This goes on until we reach the early systems that influenced each other and were subject to influences that locally created these instructions.

[43] A picture of the early universe as a collection of classical particles with a definite position and velocity is of course rather outdated, and in reality, we may not yet even know what kind of quantities are involved in the initial conditions of our universe. It is merely meant as a visual with the goal of promoting understanding.

12 can be interpreted to be both a visualization of a type I as well as a type II model depending on whether you think the correlations were created in, likely, the early phases of the universe through causal interactions, or whether they are a direct consequence of the initial conditions during the big bang respectively.

The type I and type II models give us some understanding of what statistical *dependence* in this context can be taken to mean for what our world is like. But as the reader may already have considered, it leads to an interesting state of affairs. It would seem that the universe is 'just so' to give us the predictions of quantum mechanics every experiment we perform. As if the world is involved in one gigantic conspiracy to make sure we violate Bell's inequality in every possible test of it we can imagine. Either having it or understanding where this intuition comes from, will make it much easier to grasp one of the most important critiques of superdeterminism in chapter 8. Before we get into the heat of the debate, however, first a little more about where exactly superdeterminism stands on the many interpretive issues within quantum mechanics.

Defining and 'interviewing' superdeterminism

Superdeterminism is a class of local, Ψ-epistemic, deterministic hidden-variable models that, critically, assume that the probability distribution of the hidden variables determining the values of physical quantities of a quantum system is correlated with the measurement settings involved in trying to measure these physical quantities. This definition is in line with how advocates of the framework themselves, explicitly or implicitly, introduce it (Hossenfelder & Palmer, 2020, p. 4) (Andreoletti & Vervoort, 2022, pp. 1-2) ('t Hooft, 2016, p. 10). As stated in chapter 4, some of these models are alternative theories to quantum theory as they make different empirical predictions, while others are interpretations thereof. The origin of the 'super' prefix is never mentioned. It is assumedly there because due to some consequences of rejecting statistical independence, it may appear, or 'feel', 'more deterministic' than classical mechanics. Of course, from a logical point of view, determinism is determinism, and there is no such thing as 'extra large determinism'. This is why I feel that it is rather that superdeterminism taps into a subjective association with the concept[44] where the origin of the prefix is to be found. One such example of this is that our own 'free choice' of measurement settings is also crucially restricted in the theory. Counter to our intuition, these choices must

[44] These associations may be familiar to those who often engage themselves with the free will debate. Determinism sometimes tends to be conflicted with distinct notions such as fatalism and predestination.

necessarily also be correlated with the probability distribution of the hidden variables of atoms whose quantities we aim to measure. Moreover, the aforementioned intuition of a 'universal conspiracy' also comes with a highly unintuitive metaphysical outlook. These examples, and the degree to which these intuitions may be judged to be reasonable, will be thoroughly discussed in chapter 8.

To get a complete picture of the superdeterministic framework, we will now 'interview' it and see how it answers the interpretive questions posed in chapter 4 in order.

- Superdeterminism solves the measurement problem because in it, there is no wave function collapse happening. Quantum objects always have a definite state, and there are no superpositions whereby only after measuring a value is found. Applying statistical dependence to the table resulting from Mermin's experimental setup is one example of how observations can be explained without needing to resort to collapse. The macroscopic world is thus fully reducible to the quantum world and detectors play no special role in the formalism.
- The restrictions imposed on hidden-variable theories by Bell's theorem do not apply to superdeterminism, as it rejects its assumption of statistical independence. The same goes for the Kochen-Specker theorem and the PBR theorem. Leggett's inequality does not apply as superdeterminism is local, and the contradiction in the free will theorem is also avoided by rejecting free choice, i.e., statistical independence. Landsman's theorem does affect superdeterministic *interpretations*, as will be discussed in chapter 9.
- Superdeterminism is deterministic, meaning that knowledge of the state of any system at one time and of the evolution law of that system through time enables one to, in principle, infer the system's state at all future times.
- Superdeterminism is a realist framework. It assumes that there exist observer-independent real states to systems that evolve deterministically over time.
- Superdeterminism has a Ψ-epistemic view of the wave function. It views the wave function as a statistically emergent feature of an underlying theory, not as corresponding to a fundamental entity in nature, as the many-worlds interpretation would.
- Being a hidden-variable framework, superdeterminists do not believe that quantum theory is a complete description of reality. Rather, there are as of yet unknown hidden variables that would yield new insights that current quantum theory does not have.

- Superdeterminism is compatible with locality in the sense of PL2, but also in that of PL1. There is no influence that instantaneously collapses the wave function of an object, but objects already contain instructions on what values to reveal upon measurement. In fact, they truly possess these values.

While all superdeterministic models generally agree on these questions, they may disagree on issues such as the legitimacy of the Born rule in all contexts or on what the hidden variables are. Several models have been built, pointing out different hidden variables and concrete mechanism accounting for the correlations. An example would be Gerard ‘t Hooft’s cellular automaton model, which works in a cellular grid changing the cells around it discretely and deterministically through time ('t Hooft, 2016). A famous example of a cellular automaton is Conway’s game of life. In this thesis, however, the viability of the framework as a whole will be evaluated, with individual models sometimes being invoked only in service of this goal. The ontology of the models may very well also have a philosophical element, but the questions we have asked ourselves and the philosophical debate surrounding superdeterminism focus mostly on the framework in its entirety. In addition, getting into the models would inevitably entail getting deeply into the mathematics that they employ, which is not the goal.

Now that we understand what superdeterminism is, we can evaluate the debate surrounding it. Starting with chapter 7, arguments in favor of superdeterminism will be considered, to then move on to several critiques in chapter 8.

Chapter 7: Evaluating arguments in favor of superdeterminism

Superdeterminism is clearly not the most popular position in quantum foundations, for reasons that will be extensively considered and evaluated in the next chapter. At the same time, there is an increasing number of voices speaking out in favor of the framework. Arguments in favor of superdeterminism are advanced by physicists and philosophers such as Gerard ‘t Hooft, Sabine Hossenfelder, Tim Palmer, Louis Vervoort, and plenty more (Andreoletti & Vervoort, 2022, p. 2). In this chapter, the main arguments advanced for superdeterminism will be laid out and evaluated on their strengths and weaknesses, such that we may get a nuanced picture of the supportive case in the end. That also means that from this point onward, my own attempts at analysis, arguments and evaluation of subject matter will feature more frequently.

Solution to the measurement problem

Expanding on what was already stated in the previous chapter, one argument for superdeterminism is that it (dis)solves the infamous measurement problem. For Hossenfelder and Palmer, this is the central motivation for adopting superdeterminism (Hossenfelder & Palmer, 2020, pp. 2-4). If quantum systems have real states and evolve deterministically through time, there are no superpositions and, subsequently, no collapsing wave functions. After all, the Schrödinger equation, whose linearity allows for the superpositions, is not viewed as a fundamental description of physical reality. Rather, it emerges from an underlying superdeterministic model as the limit in which the hidden variables are distributed randomly, like how thermodynamics resulted from statistical mechanics when taking the thermodynamic limit. Thus, to the superdeterminist, the cat was always dead or alive. The question as to why and how collapse happens then dissolves. The apparently privileged role for macroscopic detectors that somehow ‘do something’ to the microscopic quantum system does as well, as there is no longer any need to ‘do that something’. Reductionistic physics is seemingly restored by doing away with a macroscopic classical system that is taken as necessary for the formulation of a quantum theory that claims to be fundamental, thereby also having to be able to describe that macroscopic classical system from first principles. Relating this to earlier theory, in this sense superdeterminism concurs with Einstein’s view, who referred to Bohr’s doctrine of the classical concepts as a “tranquilizing philosophy” and claimed that instead, radically new quantum concepts were needed to solve existing problems (Kumar, 2008, p. 321).

It seems to me undeniable that under these assumptions there is no measurement problem. In that sense this is an unambiguous win for superdeterminism.

Determinism as an advantage

In the philosophy of quantum theory, superdeterminism represents a minority position of being deterministic. For many superdeterminists, however, determinism is not just an accident of the framework, but a strongly motivating property to pursue it in the first place. To ‘t Hooft, there are some demands that the universe obeys that are “*nearly inevitable and non-negotiable*”, the first of which is determinism ('t Hooft, Free Will in the Theory of Everything, 2017). He starts from the thought experiment that you are a God with the task of running a universe. He states that a vast space of possibilities takes much more calculating power and ‘administration’ than a universe that is in one definite and deterministic state. This then results in a far less efficient universe in the absence of determinism. In addition, when we consider experimental settings such as particle accelerators, we ultimately always observe particles to choose just one path and moment at which collision happens. There must therefore be some rule for that. Probabilism and ambiguities on this front are not forced on us by nature, but they reside in our theories. This shows that there must be some rule determining when a collision takes place. Lastly, in response to the misconception that Bell’s theorem would rule out *all* hidden-variable theories, ‘t Hooft states that it is very well possible to have a deterministic formulation quantum mechanics. We can, therefore, build a framework that gets rid of the ambiguities in measurement outcomes. This, to him, speaks in favor of superdeterminism, as it manages to account for quantum mechanics by doing this.

I am not quite convinced by ‘t Hooft’s arguments for determinism as laid out here. First of all, ‘t Hooft’s arguments hinges on some propositions that are neither proven, nor intuitively evident. Are ‘calculations’ of any kind actually needed to ‘run’ the universe? Why would a universe that, metaphysically speaking, is fundamentally probabilistic, necessarily require more ‘administration’? Is that not a too human way of thinking about fundamental processes beyond understanding in these terms, brought about by the overextension of the ‘being a God and running a universe’ thought experiment? Moreover, why should the universe be compelled to run as efficiently as possible? Concerning the latter, take the example of an empty universe. This would be more ‘efficient’ than ours in the sense that far fewer information is needed to encode its state and calculate what happens (nothing). After all, the positions and velocities of all particles in the universe would not need to be specified.

Some nuance, however, is in place. ‘t Hooft’s arguments here would be stronger given specific metaphysical presuppositions. For example, if the universe is a simulation created by someone, or created by God with finite computing power, it could be that from the perspective of such entities it would make sense to run fewer calculations such that more things can happen more quickly. ‘t Hooft’s thought experiment is, of course, precisely this situation. Nevertheless, this does not justify one in establishing a connection between this thought experiment and the real world. Moreover, the claim that a deterministic universe would require less calculations is *still* by no means trivial. Perhaps this God, transcending the universe we find ourselves in, operates with very different rules and powers that could create a fundamentally random number generator for quantum events in a way that no more calculations would have to be made.

Secondly, I do not think ‘t Hooft is justified in his claim that there exists a rule determining what happens when particle collision takes place in a collider from the mere fact that experimentally all particles end up choosing one path, with probabilism remaining as a mere mathematical peculiarity of an incomplete quantum theory. There is indeed a rule, whereby quantum field theory allows for the calculation of probabilities of particular interactions and outcomes resulting from collisions. Given the previous paragraph, it does not follow that the rule must instead necessarily be deterministic.

Lastly, one may of course provide a deterministic model of quantum mechanics and this may show determinism is an option. ‘t Hooft is correct in pointing out that this is perfectly compatible with Bell’s theorem, whether it is a Bohmian model or a superdeterministic one. But given we have both probabilistic and deterministic models, mere existence thereof is not sufficient to decide on the matter.

In conclusion, I do not think ‘t Hooft’s arguments conclusively show that the superdeterministic property of being deterministic is itself an argument for the framework.

Fertile ground for a Theory of Everything

Multiple superdeterministic authors propose that the radical revision made by superdeterminism could provide a new pathway towards a ‘Theory of Everything’.

In physics, there are two state of the art theories with great explanatory power in their respective domains. One is quantum field theory, which accurately allows one to predict (statistical) outcomes of (mostly) the microscopic world. The other one is general relativity, which allows one to predict gravitational effects taking place at larger scales. However, these two theories rest on different foundations and run into problems when one tries to combine them. There are four fundamental forces in physics, three of which are described by the Standard Model of particle physics (a quantum field theory) and the remaining one, gravity, by general relativity. But while the Standard Model describes forces as resulting from the exchange of force carrying particles, general relativity describes the gravitational force as resulting from the geometry of spacetime. This makes the attempt to unify these forces in one description very difficult, and straightforward attempts lead to infinities in calculations. On a more philosophical note, general relativity is a realist, deterministic and fully local (PL1) theory with no reference to wave functions and superpositions.

Their apparent incompatibility tends not be a problem, as in most situations in nature either quantum theory or general relativity is applicable. However, some situations exist where both are. These are black holes and the very early universe. It is then not a surprise that many fundamental questions remaining unanswered with regards to these subject areas. This is one reason physicists desire to find a more fundamental theory underlying both quantum theory and general relativity. Two other more explicitly philosophical reasons for this pursuit are as follows. Firstly, unification has been the norm in the history of physics. Newton showed that both the phenomena of an apple falling to the ground and the Earth orbiting the sun can be described by a single underlying theory. Now, we know that electricity and magnetism are, in fact, part of the same force, and this force can be further unified with another fundamental force (the 'weak nuclear force'). Extrapolating this historical process may lead one to believe that we can eventually find a theory unifying all phenomena in the universe under a single framework. A second reason may be the idea that the universe must ultimately operate in a logical and consistent manner, captured by a single all-encompassing and self-consistent theory. This is not a mere intuition. It would, after all, defy logical principles such as non-contradiction if two contradictory theories would be applicable to the same phenomenon, as is currently the case.

This idea of a fundamental, self-consistent and complete theory encompassing all natural phenomena is referred to as a 'Theory of Everything', the holy grail of fundamental physics that would revolutionize the field. Physicists have been looking for such a theory in many ways, but it has not yet been found.

That is why some physicists argue that the ways in which such a theory has so far been looked for are insufficient, and that some fundamental revision to our understanding of the natural world is required. This is, according to some proponents, where superdeterminism could come in. As a radically different vision on quantum theory that, moreover, relinquishes an unquestioned principle such as statistical independence, may according to them just be the revolutionary step physics needs. This could be compared to Planck introducing quantization, or Einstein rejecting the idea that velocities can simply be added to yield an object's total velocity[45]. In addition, making quantum theory deterministic like general relativity does some work to align the metaphysical premises these theories are based on, which might be what is needed for unification to succeed.

These suggestions can be found in texts by both ‘t Hooft, Hossenfelder and Palmer ('t Hooft, Free Will in the Theory of Everything, 2017) (Hossenfelder & Palmer, 2020). While none of these authors claim to know with certainty that this will turn out to be true, the very suggestion of possibly providing such a holy grail is an ambition that should not be immune to scrutiny.

The conceptual coherence and objective existence of such a theory are not uncontroversial, let alone the implicit conviction in this search that humankind is capable of discovering it. More interesting for our purposes, however, is see whether given these things, superdeterminism may indeed be up to the task.

Some potential problems with the view that superdeterminism is the road towards this theory will become clear when analyzing the consequences of statistical dependence. They are corollaries of arguments meant to show flaws of superdeterminism due to its lack of statistical independence. These corollaries will become visible in chapter 9.

A more general problem is of course that there is no way of knowing that out of many possible radical steps one could take in trying to construct a Theory of Everything, the assumption of statistical independence was, in fact, the one thing standing in its way. Loop quantum gravity, another research program aiming to unite general relativity and quantum theory, takes the radical step of quantizing spacetime. String theory, on the other hand, brings a radically new ontology to the table in which all particles are actually one-dimensional vibrating strings. This includes a

[45] To be precise, this is not the core of special relativity but rather a consequence thereof. It is, however, a simple to understand way in which relativity radically changed our thinking about mechanics.

mode called a ’graviton’ responsible for the gravitational interactions, just as how in quantum field theory the photon is responsible for electromagnetic interactions. Perhaps Penrose is right on the money when he formulates a mechanism relating gravity and quantum theory, where gravity causes the collapse of the wave function, the consequences of which may open new research pathways to a deeper theory. And maybe a breakthrough in our understanding of dark matter and dark energy is what will eventually show the way. The point is of course that without knowing what a Theory of Everything looks like, the probability that superdeterminism fits the mold increasingly falls of as we imagine what other radical moves might be open to the physicists. And that all operates under the, while perhaps likely, unproven assumption that some radical divergence from our current metaphysical and ontological understanding is a necessary condition to the establishment of such a theory.

Another noteworthy comment concerns the degree of complexity a superdeterministic Theory of Everything would likely contain. While simplicity tends to be put forward as an important epistemic value in how a physicist judges a theory, philosopher and physicist Eddy Keming Chen argues that superdeterminism is very unlikely to produce this (Chen, 2020, pp. 13-16). After all, a superdeterministic model will need to make sure that regardless of what setting mechanism one uses in a Mermin-like experiment, the correct correlations with the state of atoms being measured, resulting in violations of Bell’s inequality, must be present. These setting mechanisms can be anything, from the polarization of light from stars at incredible distances to the even- or oddness from the millionth digit of π onwards. The constraints this leaves a superdeterministic model with are likely to make it so complex that Chen, given the criterium of simplicity, concludes it to be much more sensible to consider pilot wave theory or objective collapse models like that of Penrose (Chen, 2020, p. 16). As will be seen when discussing the ‘science-invalidation argument’ in chapter 8, some superdeterminists deny that this complexity would be required. In response, Chen argues that their proposed solutions to get rid of it just serve to move the complexity to other places, such as the ontology required to formulate such a model (Chen, 2020, pp. 16-18). The details will thus be worked out in chapter 8, but this sufficiently serves to makes Chen’s point regarding complexity in relation to our current topic.

A final small argument I want to make against the superdeterminists’ theoretical ambitions is that the superdeterministic framework itself, which is what is being talked about in the articles after all, gives one no clue as to how a connection between general relativity and some new theory underlying quantum mechanics is to be seen. The bridge between these two regions is clearer in approaches like string theory and loop quantum gravity, which explicitly start out from a radical move that connects the two somehow[46]. Perhaps a very concrete superdeterministic model with its own ontology and mechanisms can somehow bake in this connection more explicitly, but the reference to opportunities for a Theory of Everything is made with regard to the framework itself. And with that regard, it is not very clear when compared to the competition.

In conclusion, I do not think that the possibility of finding the Theory of Everything is a strong argument in favor of superdeterminism. That does not logically exclude its ability to do that. But for the above reasons, too few appealing reasons for this are given to justify supporting superdeterminism in particular.

Similarity to the Liouville equation

Classical physics is deterministic. However, it can also display chaos[47]. This is the case when in a physical system, extremely small differences in initial conditions can lead to enormous differences in the evolution of the system through time. A concrete example is a double pendulum[48]. Even though the difference in angle at which you let the pendulum fall might be arbitrarily small, the motion will still develop in a very different way. This, however, does not mean that the double pendulum is indeterministic. If you know the initial conditions such as the starting angle precisely, Newton's laws of motion will still tell you exactly how it will move over time. However, in practice it is extremely difficult to precisely predict the motion due to the high sensitivity to such initial conditions.

Luckily to physicists, there is the Liouville equation. This equation allows one to formulate the probability of finding, for example, the pendulum at a certain position with a certain velocity at a certain time after it has started to move. Therefore, while despite being deterministic it is extremely hard to predict these physical quantities

[46] Since loop quantum gravity applies quantization (quantum realm) to spacetime (general relativity realm) and string theory hypothesizes an elementary particle force carrier called a graviton (quantum realm) that mediates gravitational interactions (general relativity realm), this ambition is clear. The connection between the two realms is very visible as it is baked in the projects' DNA from the onset.
[47] The notion of (non)linearity is very important to this discussion, but for the sake of comprehensibility I will try to qualitatively explain the essence of the argument.
[48] A visualization makes this much easier to grasp. A good visualization can be found in the following simulation on Youtube (Twice, 2017).

precisely, we can use the Liouville equations to find the probability distribution of where to find the pendulum and at what velocity at a specific time.

Now it would be a mistake to claim that the position and velocity of the double pendulum is fundamentally probabilistic, and that in addition, the Liouville theorem is a complete description of physical reality as it pertains to that system, allowing no underlying theory. Because, in reality, there is an underlying deterministic theory at play, and the Liouville equation tells us something about statistics emerging from the dynamics at play at that level.

One may already see where this is going. The argument for superdeterminism as provided by Hossenfelder and Palmer is that the Liouville equation looks "remarkly similar" to the Schrödinger equation[49] and that this "strongly suggests" that quantum theory may just be the probabilistic description of a chaotic deterministic theory that underlies it (Hossenfelder & Palmer, 2020, p. 3). Without going into the details but just to give the reader a feeling and show what is being talked about, the equations are as follows:

Liouville equation: $$\frac{\partial \rho}{\partial t} = \{H, \rho\}$$

Schrödinger-Liouville equation: $$i\hbar \frac{\partial \rho}{\partial t} = [H, \rho]$$

In essence, the argument proposes that the formal similarity between these two situations hints at that the relationships between these two levels of description are equivalent.

The obvious thing to be said is that formal similarity does not guarantee equivalence. However, I think the argument certainly has merit when interpreted not as a proof, but as a motivation to investigate something. Therefore, it is not an argument for the correctness of superdeterminism, but it seems to me reasonable to accept that such a striking similarity is more than interesting enough to function as a hint, warranting further investigation.

Beyond the obvious, a point must be made on how looks can be deceiving. Hossenfelder has written an entire book on how arguments from beauty "lead physics astray" (Hossenfelder S. , Lost in Math: How Beauty Leads Physics Astray,

[49] To be accurate: remarkably similar to the Von Neumann-Dirac / Schrödinger-Liouville / Quantum-Liouville equation, which for our purposes can (slightly reductively) be understood to be 'the Schrödinger equation written in a different way'.

2020). One of its core messages is that physicists should be careful with constructing arguments on the basis of a subjective notion of 'mathematical beauty'. One may however ask whether the two equations are as similar as they appear to the eye in the way they are expressed above. For one, the imaginary unit and Planck's constant appear prominently in front of the partial derivative for the Schrödinger-Liouville equation. Planck's constant is the hallmark of quantum physics and displays a very real difference with classical theory. The imaginary unit is hard to interpret straightforwardly but it is unique to quantum theory that it features prominently in the equation of motion of the theory, i.e., the Schrödinger equation. Perhaps more importantly, the small difference in brackets can be deceiving. The square brackets are commutator brackets, while the curly brackets are Poisson brackets. These are defined in a different way:

$$[A, B] \coloneqq AB - BA$$

$$\{A, B\} \coloneqq \sum_{i=1}^{N} \frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i}$$

Again, it is for our purposes not important to know exactly what all these symbols mean, but it is important to see that these seemingly alike small brackets are, in reality, shorthand for mathematical expressions that are not quite that similar. Moreover, the contents of the letter 'H' in the equations (the Hamiltonian) are also quite different for the classical and quantum case, and as was noted in footnote 49 the Schrödinger-Liouville equation is in a sense just a recast Schrödinger equation. One may well write it in a different way such that the beautiful elegance and similarity greatly diminishes.

At the same time, Hossenfelder and Palmer could, I think legitimately, counter that the very existence a way to write these equations in the above form provides enough motivation to investigate, and that the equations nevertheless share certain additional mathematical properties that inspire this[50]. Letting the validity of superdeterminism aside, the increasing amount of correspondence on this topic seems to have achieved this motivation to investigate at the very least.

[50] Here the linearity mentioned in an earlier footnote comes in. A central idea of the authors is that a *nonlinear* deterministic description might underlie the associated *linear* probabilistic equation, not just in the classical Liouville case but also in the quantum case.

Total compatibility with the principle of locality

In chapter 3, the distinction between PL1 and PL2 was introduced. PL2, stating that no *information* can travel faster than light, was explained to be a core tenet of relativity and quantum field theory. The fact that both our most fundamental and well-tested theories of physical reality have this principle built in so strongly suggests that physics is unlikely to part with it anytime soon. It was also mentioned in chapter 5 that the difficulty for pilot wave theory to incorporate this principle is considered that theory's greatest challenge.

PL1, however, is generally not thought to be obeyed by mainstream ways of understanding quantum theory. In the Copenhagen interpretation, even though quantum particles have no real states prior to measurements[51], the fact that if I measure one physical quantity of a particle in an entangled pair, the state of the other one can be predicted with certainty as well, means that some influence has instantaneously travelled between them. This was the argument made by the EPR-paper under their criterion of reality, and Bohr had recognized some influence as well. We now know that this influence does not carry *information* faster than light as the probabilistic nature of the collapse prevents two experimenters from communicating through this process.

If we accept, as most do, that some instantaneous influence is still present here, one may still wonder if the answer to this should really be as simple as stating that it is not in direct contradiction with relativity and then doing away with the criticism that quantum mechanics allows for some sort of nonlocality.

At this point, a pragmatically-minded physicist could say that the PL1-PL2 distinction works and that there is no physics problem left here. However, while a popular distinction, it is not uncontroversial. There exists plenty of disagreement on how far-reaching relativistic locality really is (Berkovitz, 2007, p. §10). For physics, it is an open question whether this 'quantum nonlocality' is as unproblematic as often assumed.

On a philosophical note, one can at least ask the question how coherent it is that some forms of nonlocality are strongly forbidden, while others are fine. Instantaneous influences from any kind have been a contentious issue in the history of physics, from the Clarke-Leibniz correspondence and before to contemporary

[51] Remember that if they do, Bell's inequality holds, and the only way to keep insisting that they do is to do away with locality or statistical independence.

quantum theory. Our understanding and lenience towards forms of action at a distance has changed over time (De Regt, Understanding Gravitation, November 2021). A related topic is whether ‘information’ as a concept employed in PL2 is a fundamental entity of a universe, or a mere instrument with, so far, utility for physicists. While not diving into the ontology of information, it is clear that different positions exist, and that the latter would make it harder to justify why the universe would join us in our (in that view) constructed PL1-PL2 distinction.

In any case, for people who do have a problem with instantaneous influences allowed by quantum theory, superdeterminism may be an attractive option. Unlike other interpretations[52] that were discussed in chapter 5, superdeterminism can accommodate PL1 and PL2. This issue could therefore, depending on one’s view of the content and importance of the principle of locality, be counted as an argument in support of superdeterminism.

A final note on the locality issue, however, is an apparent contradiction between two statements made concerning superdeterminism. On the one hand, it was stated that the Copenhagen interpretation does not allow for superluminal communication due its probabilism preventing its ‘spooky action’ from encoding information. But on the other hand, superdeterministic theories may reject the Born rule in some cases and do not accept probabilism. Therefore, one may be mistaken to believe that superdeterminism could be exploited to communicate by sending information faster-than-light. Interestingly, while it may allow for superluminal communication, no information is ever sent throughout this conversation since the apparent sending of information is, in fact, explained away by the existence of pre-existing correlations implied by the rejection of statistical independence. This seemingly bizarre state of affairs can certainly be argued to expose a flaw within superdeterminism, but that flaw is *not nonlocality*. This issue will be further explored slightly more in depth when discussing Valentini’s and Sen’s conspiracy argument in the next chapter. Knowing that at least the consistency of superdeterminism may be saved in this regard, we can move on to the final subchapter dealing with this.

[52] For the many-worlds interpretation, the answer to whether it contains any nonlocal influences is “not straightforward, as it depends on one's particular reading of the interpretation” (Berkovitz, 2007, p. §5.3.3). The next subchapter will however show that not having such ambiguities might also be an attractive point about superdeterminism.

Completeness and consistency regarding interpretive questions

Whether one agrees with its contents or not, in the last chapter it was seen that superdeterminism is able to provide consistent answers to all the formulated interpretive questions of quantum theory with relative ease. From the premises of the framework it clearly follows what kind of philosophical view emerges from it, irrespective of whether one agrees with it. I provide this as a philosophical argument in favor of superdeterminism because its own consistency, as well as its ability to provide answers to the main interpretive questions in quantum theory, was advanced as one of the criteria for its very viability. Neither of these two are guaranteed for all interpretations of quantum mechanics. After all, where does the line between macroscopic observer and quantum system under consideration lie in the Copenhagen interpretation? Do all supporters of the many-worlds interpretation answer the question of realism in the same way? Do pilot wave theories necessarily allow for PL2 violations, meaning we live in a universe with relativity violating instantaneous communication-allowing action? And does determinism return in the objective collapse mechanisms that tend to be proposed, or do these collapse processes still come down to fundamental probability?

The point of these questions is that while the answers to them are not immediately clear even among supporters of these frameworks, this charge is more difficult to levy against superdeterminism. Its consistency and straightforward applicability to philosophical questions may be counted as a plus.

At the same time, specific models within these other frameworks may very well answer these questions in a clear manner as well. Secondly, superdeterminism being clear and consistent with respect to these particular philosophical questions is not equivalent to it being correct. Thirdly, the list of questions considered in chapter 4 was admittedly not exhaustive with respect to all of the foundations of quantum theory. And perhaps most importantly, could the rejection of statistical independence not throw up new philosophical questions faced with the same (if not worse) ‘problems’ than mentioned here for other frameworks?

Finally, the following two chapters will provide some reasons to be a little more doubtful of the positive claims made here. An example is that as will be discussed in chapter 9, it is not so clear that all superdeterministic models can really be deterministic after all.

These are all relevant nuances against this argument. A complete evaluation of superdeterminism’s self-consistency and ability to provide answers to the philosophical questions arising from quantum theory can only be performed by an extensive study of the consequences of statistical dependence. Therefore, now having a good grasp on the arguments in favor of superdeterminism and how these can be evaluated, we will proceed to turn our attention to this study in the next chapter.

Chapter 8: Evaluating arguments in opposition to superdeterminism

Despite the arguments in favor of superdeterminism discussed and evaluated in the previous chapter, superdeterminism enjoys relatively little support from the physics community. Naturally then, there exist several arguments against it that are judged to be convincing by many people. The understanding and explicit evaluation of these arguments in this chapter is therefore crucial in coming to a judgment on its viability. Due to most critiques against superdeterminism being model-independent, these arguments are usually very philosophical in nature, targeting mainly metaphysical and epistemological concerns with the framework. Three of them will be extensively reviewed in this chapter. The two critiques I consider to have most merit will be the first two and have, in fact, already been partially set up in chapter 6. These are the more metaphysically inclined conspiracy argument and the more epistemologically inclined science-invalidation argument. The latter one, the argument from free will, is argued to be less convincing. For the sake of completion, two further but less common arguments will be briefly introduced at the end. Issues surrounding future expectations of empirical testability will be covered in the next chapter.

The conspiracy argument

The conspiracy argument is one of the most frequently used arguments against superdeterminism. The argument is primarily metaphysical is nature and starts from the intuition that superdeterminism requires an extremely high number of 'coincidences' to work. This is then shown to imply that the framework comes with an extreme fine-tuning problem. In this subchapter, the conspiracy argument will first be clearly formulated and evaluated, using the input of Bell himself and subsequently that of Antony Valentini and Indrajit Sen. Due to its commonality, most superdeterminists have some replies to the argument, making for quite a lengthy list. The second half of this subchapter will then be dedicated to exploring these counters to the conspiracy argument, to then be analyzed and see whether they hold up.

The central intuition to the conspiracy argument has already been discussed in chapter 6. Whether this is realized through a type I or type II model, the universe contains correlations that are always 'just so' that superdeterminism can reproduce the measurement statistics of quantum theory. This has to do both with their systematic nature and the apparent unlikeliness given the systems that the correlations apply to. It will now be shown where this idea comes from explicitly.

Bell's conspiracy argument

There are many ways in Bell tests have been performed, specifically if we look at the methods used to determine measurement settings. As has been mentioned before, one method can be the use of the polarization of light from distant quasars that cannot have communicated over half the age of the universe ago. Yet, from the perspective of superdeterminism, there must still be a correlation between the light used to determine the settings and hidden variables of a quantum system considered here on Earth. Measurement settings can also be set by letting people freely select them themselves, and yet again we observe the correlations[53] (Collaboration, 2018). Lastly, the most peculiar mechanisms can be appointed to choose the measurement settings, such as random number generators, the even- or oddness of digits of π beyond the millionth or the coloring of a single pixel of your television over time in your favorite movie. A concrete example of how to implement the latter suggestion could be that, using the Mermin setup, I use setting 1 if the movie's 100th pixel is red, setting 2 if that pixel is green and 3 if blue. Every second I evaluate the color of the pixel to determine the settings in this way, using all kinds of movies. But alas, even in bizarre setups such as this, the correlations will be there. Across all the three categories just considered, all kinds of wildly different variables turn out to have to be correlated for superdeterminism's to consistently have statistical dependence produce violations of Bell's inequality. Examples of these sorts were advanced by Bell in a paper he wrote addressing the possibility of statistical dependence observed by Shimony et al. (Bell J. , Free variables and local causality, 1987). He describes a random number generator that uses chaos, being technically deterministic but having widely diverging outputs for minute changes in the inputs. Bell tells us that:

"..nothing is forgotten. And yet for many purposes, such a device is precisely a 'forgetting machine'." (Bell J. , 1987, p. 102)

Here he refers to the fact that due to the sensitivity of the random number generator to these minute changes in input, for all practical purposes these inputs are forgotten regardless of if such a system is deterministic. He goes on:

"With a physical shuffling machine, we are unable to perform the analysis to the point of saying just what peculiar feature of the input is remembered in the output. But we can quite reasonably assume that it is not relevant for other purposes. In this

[53] More discussion on this particular issue as it pertains to the apparent 'free' part of the choice will be present in the third subchapter.

*sense the output of such a device is indeed a **sufficiently free variable for the purpose at hand**."* (Bell J. , 1987, p. 103) (Emphasis added)

Bell's thesis here comes down to the fact that the determined-ness of a randomization process to generate measurement settings such as in the examples listed above, does not imply that this cannot be independent of the state of some measured system in the practical sense. These inputs are 'peculiar' precisely because they seem so very unlikely to be correlated with that state. It seems reasonable to assume that there is no correlation between the color of a pixel in a movie and the spin state of an atom prepared in a lab maybe years after the show's release. Bell, acknowledging that the output of any randomization process is technically still determined, refers to such variables to install the settings as *sufficiently* free. It goes strongly against intuition *not* to think such variables and hidden variables of the sort we discussed must be statistically independent. He concludes that:

"Of course it might be that these reasonable ideas about physical randomizers are just wrong – for the purpose at hand. A theory may appear in which such conspiracies inevitably occur, and these conspiracies may then seem more digestible than the non-localities of other theories. When that theory is announced I will not refuse to listen, either on methodological or other grounds. But I will not myself try to make such a theory." (Bell J. , 1987, p. 103)

In conclusion, Bell does not logically exclude superdeterminism, but thinks it requires a highly unlikely conspiracy for the reasons stated above. The reason that it is called a conspiracy then becomes clear: the rejection of statistical independence may be said to lead to an image of the universe where the world has consistently 'set itself up' in such a way that we always find quantum statistics in Bell tests, no matter the number and bizarre nature of the correlations this requires. This seems so unlikely that we are inclined to conclude that it is as if we are being 'conspired against'.

A final illustration I believe can aid the understanding of this point that is very familiar to the philosopher is the invocation of Leibniz's parallelism. Faced with the interaction problem between the material and the mental substance, Leibniz proposed that the assumption of an interaction being there in the first place may be the root of the problem. Rather than the physical process of me stepping on a Lego block *causing* the mental effect of an experience of pain, these events happen concurrently within their perspective realm, independent of one another. This naturally leads one to ask how we, nevertheless, systematically experience correlations between physical and mental events [54]. To explain this, Leibniz presented his doctrine of pre-established harmony, arguing that in the absence of any interaction between the substances, God had synchronized all events in both realms at creation. God knew that a physical body would step on a Lego at a particular point in time and space, and therefore already 'programmed in' the presence subjective pain. No interaction needed. From a secular perspective[55], however, it becomes rather difficult to account for psychophysical harmony. I would argue that an analogous situation plays out in superdeterminism. Likewise, we are confronted with a possibly infinite number of correlations bringing about very specific results that hardly seem to be coincidences. Yet we know that there is no immediate causation between that which is correlated, namely some random number generator yielding a measurement setting and the, e.g., spin state of a particle to be measured. If that were to be the case, it would be a pilot wave theory whereby the chosen setting sends a faster-than-light signal to another particle to be measured. We are left with seemingly inexplicable but systematic correlations of bizarre nature given the ways in which settings can be chosen.

One should also realize that the randomization processes discussed by Bell can be compounded to arbitrary degree. You could make an algorithm in which three dice I throw determine which digits π beyond the millionth I use to evaluate their even- or oddness and then use this as an input to select a pixel in a TV screen with a certain movie playing, where the color the pixel is associated with a number which is one of the inputs of a sum determining the measurement setting for a Bell test. The possibility of extending this algorithm, conceivably even indefinitely, implies that it might make sense to attempt to quantify the notion of 'conspiracy'. This is exactly what Valentini and Sen intend to do (Sen & Valentini, Superdeterministic hidden-variables models II: conspiracy, 2020).

Fine-tuning in physics

For this, the notion of fine-tuning needs to be introduced first, as the intuitive notion of a conspiracy leads to a, possibly very severe, fine-tuning problem. In physics, a parameter is 'finely tuned' if it needs to have a very precise value to be compatible with experiment or any particular state of affairs (De Vuyst, 2020). Small deviations

[54] This systematic correlation is often referred to as 'psychophysical harmony'.
[55] There will probably not be many superdeterminists inclined to invoke Godly powers to support the viability of their framework. This is an uncommon practice in contemporary science for many reasons, and if that ends up being the conclusion drawn by the framework one may as well link interpretive issues, such as wave function collapse in quantum mechanics, with God from the offset, therefore no longer requiring superdeterminism. For these reasons, I will carry by investigating the implications of the conspiracy argument from a secular perspective.

can turn out to produce enormously different consequences. An example could be the universal gravitational constant G that can be found in Newtonian gravity[56]. There, the magnitude of the gravitational force of a body with mass m_1 acting on another body with mass m_2, separated from one another by a distance r, can be expressed as follows:

$$F_g = G\frac{m_1 m_2}{r^2}$$

Thus, the Newtonian law of gravity is such that the force between two objects is determined by their masses and separation. That just leaves the universal gravitational constant. Experimentally, it is determined (in SI-units) to be:

$$G = 6.6743 \cdot 10^{-11}\frac{m^3}{kg \cdot s}$$

In the formula for the gravitational force, G determines its strength given a certain pair of masses and separation. The constant is not derived from any deeper necessary theoretical construct, and as far as we know it 'just is', its numerical value not changing throughout time and space. One may, in a sense, even consider these physical constants to be part of the *initial conditions* of our universe.

We can, however, imagine that G had been different. This would not contradict any laws of logic or physics. Such universal physical constants already vary by many orders of magnitude in our own world. The Coulomb force law that appears in classical electromagnetism, which describes the force between two electric charges, is identical in form to that of Newton's law of gravitation with the exception of a wildly different constant. One only needs to replace the masses of the two objects with their charges. The constant that appears here in place of the universal gravitational one ('Coulomb's constant') is *twenty* orders of magnitude larger[57].

The consequences of changing G for our world, however, can be enormous. Suppose G had been much larger. Then the universe may have quickly imploded soon after the big bang, like a spring made of metal instead of plastic pulling in much more rapidly after being extended. But had G been much smaller, large scale structures

[56] It also appears in Einstein's theory of general relativity, which after its introduction in 1915 replaced Newtonian theory as our must fundamental description of gravity known to date. Newtonian gravity makes for an easier example though.
[57] This relates to the concept of 'naturalness' in physics, which is related to the fine-tuning debate but not an aspect we will dive into here.

(including our solar system) may never have formed. When gravity is weak, matter like hydrogen may not accumulate, and even if it does, the gravitational pressure may not be great enough to fuse it into helium and form stars. Thus, what our universe looks like depends to a large extent on the value of such fundamental constants, of which there are many. Moreover, it is not just the numerical value of a given constant that matters, but also the ratios between different constants. Specific numbers aside, had G, rather than the other way around, been orders of magnitude greater than Coulomb's constant, then gravity would dominate atomic interactions. This would change (or probably: do away with) all of chemistry. Thus, to keep specific aspects of the world similar to the one we inhabit, both the values of and the ratios between all of the physical constants walk a tight rope. Given the above examples, the many consequences of the values of the constants can also be seen to include the aspect of the existence of life in our universe. A different arrangement of constants could have rendered it lifeless. While the magnitude of 'allowed' changes to this end is somewhat larger than is sometimes thought, particularly for G, this does not change the fact that one could imagine[58] an infinite set of sets[59] of constants for which that would be the conclusion (Adams, 2019, pp. 140-141). This all ultimately begs the question: why are the constants the way they are? More specifically: why do they inhabit the small range that allows us to be here? This constitutes the essence of the fine-tuning problem of physics.

Differentiating and quantifying conspiracy arguments

With our understanding of the traditional fine-tuning problem about parameter fine-tuning for life, we can now turn our attention to the rigorous conspiracy argument of Valentini and Sen. The core argument in their two-part publication on the topic is that superdeterminism as a general class of models is conspiratorial in at least two ways, and that it is possible to quantify the degree of fine-tuning this entails. On the basis of these arguments they conclude that superdeterminism is a "scientifically unattractive" explanation of violations of Bell's inequality. Valentini and Sen sometimes use a more formal mathematical approach in their treatment of superdeterminism[60], but the essence of the subject matter has already been discussed in qualitative terms before (Sen & Valentini, Superdeterministic hidden-variables

[58] Conceivability is one thing, possibility may be another. This distinction will come back later.
[59] Whereby one set lists all the values of the constants of a universe, and the set of all sets represents all possible configurations of values of the physical constants, i.e., all possible universes given changed constants. This set of all sets is infinite.
[60] An example is their discussion on 'marginal-independence' in section II. Their definition of a signal is discussed here as well, but without stating it in the form of a mathematical condition.

models I: nonequilibrium and signalling, 2020) (Sen & Valentini, Superdeterministic hidden-variables models II: conspiracy, 2020).

Important to their analysis is the concept of a nonequilibrium hidden-variable theory. As noted before, deterministic physical theories come with laws and initial conditions. The laws are usually understood as necessary: they are the rules of our universe whose content is given and uncontrollable. Initial conditions, on the other hand, are usually understood as contingent: they could have been different, and they can be controlled for as in the example of the ball and gravity before.

In superdeterminism, there are hidden variables λ associated with quantum systems. In concrete terms, these could be some property of quantum particles we do not yet understand, but in fact determine the spin of these particles with certainty in Mermin-like experiments. This was demonstrated using the color states such as 'RGR' before. When measuring the spin of a large amount of such particles, superdeterminists posit that the hidden-variable distribution over the particles are such that statistics equivalent to those predicted by quantum theory come out. This is the so-called 'equilibrium situation', which is demanded to hold by the superdeterminist, but is in reality only a small subset of possible statistical distributions. Given the contingency of this specific distribution of λ's, other distributions are possible in the superdeterministic framework. These will then *not* reproduce quantum statistics, and for this reason they are called 'nonequilibrium distributions'. For superdeterministic models with such a hidden-variable distribution, there will still be statistical dependence, just without correlations that give rise to the very specific statistical distribution compatible with what quantum theory predicts. As one may imagine, there are in principle many more of such distributions possible than the particular ones reproducing quantum theory. They therefore refer to this as a "nonequilibrium extension of a superdeterministic model" (or more briefly: the "quantum nonequilibrium"), something more general (Sen & Valentini, Superdeterministic hidden-variables models I: nonequilibrium and signalling, 2020, pp. 4-5).

The authors now note that for this general class of superdeterministic models, it is not at all given that the same hidden-variable distribution (i.e., the equilibrium one) will come out independently of the process with which a measurement setting is selected. In other words, given that the correlations are not already set up to reproduce quantum statistics, if I had repeated my experiment using the colors of TV pixels instead of a random number generators, the measurement statistics would in general come out differently. One may respond that in superdeterminism it is not possible to have done an experiment using different mechanisms than you actually ended up doing, so it does not matter that a different experiment from what really happened could have led to different statistics. This, however, is not the point made by Sen and Valentini, and the true point was already illuminated clearly by Bell himself:

"I would insist here on the distinction between analyzing various physical theories, on the one hand, and philosophizing about the unique real world on the other hand. In this matter of causality it is a great inconvenience that the real world is given to us once only. We cannot repeat an experiment changing just one variable; the hands of the clock will have moved, and the moons of Jupiter. Physical theories are more amenable in this respect. We can calculate the consequences of changing free elements in a theory, be they only initial conditions, and so can explore the causal structure of the theory." (Bell J. , Free variables and local causality, 1987, p. 2)

Thus, the point here is to change a parameter within the theoretical framework and see how it responds. This practice is perfectly compatible with determinism and fundamentally distinct from the concept of changing a parameter in the causal chain given in the real world. An intervention of the latter kind is, after all, impossible in a deterministic setting.

They conclude now that "the choice of setting mechanism can affect the measurement statistics for a nonequilibrium extension of a superdeterministic model", and this is the seed for the first fine-tuning problem the authors detect with superdeterminism (Sen & Valentini, Superdeterministic hidden-variables models I: nonequilibrium and signalling, 2020, p. 5). Superdeterminism, after all, requires that we can reproduce the measurement statistics of quantum theory in all experiments we do, no matter what mechanism we use to establish a particular detector setting. Given that this does not follow from the nonequilibrium extension, a fine-tuning of parameters, such as the initial conditions at the birth of our universe, is required to make sure that the measurement statistics do depend on the measurement settings, but not on the way in which these settings were chosen.

In their second paper, Valentini and Sen quantify this idea. Without getting too deeply into the mathematics, in essence they use combinatorics to derive an overhead fine-tuning parameter F. This is done by looking at the number of possible configurations of hidden variables given the necessary constraint that the measurement statistics must not depend on the way in which measurement settings were chosen. When F=0, the model is completely general as after applying relevant constraints all possible configurations are still possible. When F=1, the model is completely fine-tuned, as the imposed constraints leave no possible configurations at all. Lastly, there are N possible setting mechanisms for the measurement settings. By remembering the sort of examples of setting mechanisms, from digits of π to the pixels of a TV-series and their compounds, one notices that N is an arbitrarily large number. Crucially, when letting $N \to \infty$, the result is that F=1. Therefore, Valentini and Sen provide a quantitative argument for superdeterminism requiring complete fine-tuning.

As was mentioned, however, Valentini and Sen attempt to show that superdeterminism is conspiratorial in at least *two* ways. Other than the one just discussed, they argue that the second one can be found in the *apparent superluminal signaling* present in superdeterminism.

Valentini and Sen define an actual signal to be present in a Bell test if and only if (Sen & Valentini, Superdeterministic hidden-variables models I: nonequilibrium and signalling, 2020, p. 5):

1. The distribution of measurement results at one of the two detectors in an, e.g., Bohm-Mermin like Bell test is correlated with the measurement settings of the other detector[61].
2. The measurement setting of one of the two detectors is a cause of the measurement result at the other.

As a definition of signaling, this is in line with earlier discussions on locality. A causal influence will obviously imply a correlation as in the first point, but as the second point states that the origin of the correlation must lie in the direct causal influence of the setting. This is the case in, for example, pilot wave theory. But superdeterminism, according to the authors, is unique in that the first clause holds for it, but not the second (Sen & Valentini, Superdeterministic hidden-variables models I: nonequilibrium and signalling, 2020, pp. 9-10). The first clause if often referred to as the 'no-signaling constraint', which is why Valentini and Sen argue that it should be renamed the 'marginal-independence constraint'. Superdeterminism, after all,

[61] Mathematically: $p(A|M_A, M_B) \neq p(A|M_A, M'_B)$. Here, 'A' refers to a measurement result at detector A, while 'M' refers to the specific detector settings, the subscript denoting at which of the two detectors. On the right-hand side, the primed 'M' at detector B denotes a different measurement setting. Given the inequality, the measurement statistics at one side are therefore correlated with the settings at the other, as we would expect given the violation of Bell's inequality.

violates it, yet does not allow for genuine signaling of causal kind. It is only apparent, as it seems that changing a measurement setting on one detector *causes* changes in the distribution at the other, but this is actually the result of pre-existing correlations. This is visually represented in figure 13.

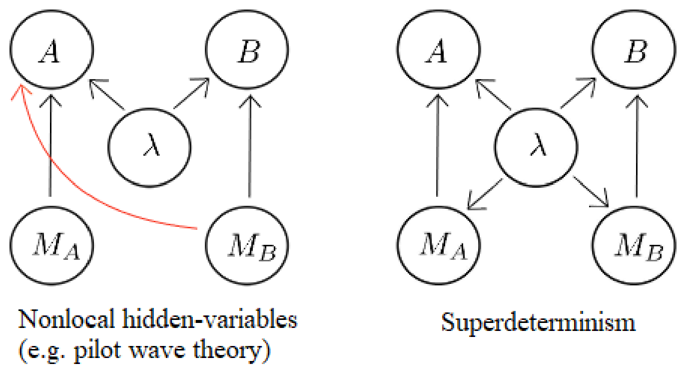


Figure 13: A nonlocal hidden-variable theory compared to a superdeterministic hidden-variable theory. In the nonlocal case, atoms have a real state characterized by λ, but changing the measurement setting at detector B causes a faster-than-light signal changing the spin of atom A later measured at detector A (again in a Bohm-Mermin type setup). In the superdeterministic case, no such signal is present, but rather, pre-existing correlations represented by λ account for observations that appear to imply signaling (Sen & Valentini, Superdeterministic hidden-variables models I: nonequilibrium and signalling, 2020, p. 10).

Given this observation, Valentini and Sen now ask themselves what the consequence is of a theoretical framework that rejects marginal-independence, yet insists that no direct causal relationship can be held responsible for the correlations this implies. This, they claim, leads to the second conspiratorial aspect of superdeterminism. It starts from the realization that someone controlling the measurement settings at one end can influence the measurement statistics at the other end. At this point one would do well to remember that we are currently referring to the nonequilibrium extension of superdeterminism, where it is not yet the case that we have demanded the correlations between measured states and settings to be such that it reproduces Born's rule. Now since these measurement statistics are different depending on this setting choice, it becomes possible, in principle, for someone choosing the measurement settings to send messages to an observer at the other end of the experiment instantaneously, even though they may be lightyears removed. They could, for example, agree to a certain code translating outcomes in sentences before they embarked on their journey lightyears apart. They could even use this to effectively construct a telephone. Nevertheless, their resulting communication, which is surprisingly at least conceivably possible in nonequilibrium superdeterministic models, will be nothing more than coincidences. The reason for this is that the setting at one end is not the cause of the outcome at the other. They are the result of pre-existing correlations. Having instantaneous but meaningful conversations over a phone while being arbitrarily far removed from each other, without the speaker's actions causing the sending of those messages, implies an arbitrarily long *"series of coincidences mimicking an actual conversation"*. This heavily implies fine-tuning which, like with the previous argument, the authors set out to quantify next (Sen & Valentini, Superdeterministic hidden-variables models I: nonequilibrium and signalling, 2020, p. 10).

This quantification is done from the quantum equilibrium point of view. The authors argue that even in this situation, an analogous fine-tuning problem turns up. The key here is that to satisfy Born's rule, each way of picking measurement settings is used only if it is correlated with λ. If this statistical dependence were, after all, not present, the correlations required by quantum theory cannot be reproduced. This, however, brings with it a fine-tuning where it can again be shown that the theory becomes arbitrarily conspiratorial[62] (Sen & Valentini, Superdeterministic hidden-variables models II: conspiracy, 2020, pp. 7-9).

Finally, they consider what superdeterminists could do faced with these problems. One option would be to do away with the possibility of quantum nonequilibrium, as the first of their conspiracies rests upon this notion. The authors, however, argue that this would entail doing away with the distinction between laws of nature and initial conditions as well. The contingency of any initial hidden-variable distribution was central to their argument for the possibility of nonequilibrium. Simultaneously, without the possibility of nonequilibrium, this leaves the initial conditions as somehow necessary, blurring the distinction under consideration. Yet, this distinction is, they argue, "a central principle in scientific theories" (Sen & Valentini, Superdeterministic hidden-variables models II: conspiracy, 2020, pp. 10-11). Moreover, I would add that rejecting the nonequilibrium is not an option for superdeterminists who aim to use their models to produce different empirical

[62] This is done by invoking the entropy of sub-ensembles consisting of the statistics obtained from a specific pair of methods to choose the measurement settings of the detectors. For the purposes of this thesis, diving into this too deeply will likely serve to obfuscate the larger point.

predictions from those of quantum mechanics. After all, the equilibrium distribution just reproduces the Born rule, making the two indistinguishable. This all becomes especially dire for superdeterminists when considering Landsman's no-go theorem in chapter 9, which state that superdeterministic models that do not sometimes violate the Born rule are impossible anyway. A second possible superdeterministic response is the suggest a mechanism that drives a nonequilibrium distribution towards an equilibrium one. However, this still allows for deviations from the nonequilibrium which could then be testable in the form of experiments that do not violate Bell's inequality. A final option for superdeterminists is to target the authors' second argument. The derivation thereof rests upon the assumption that the hidden variables are correlated only with the setting mechanisms that are actually used in Bell tests. Rejecting this assumption, they then argue, leaves one with just another conspiracy, namely that the hidden variables must be correlated with every setting mechanism that could conceivably be used (Sen & Valentini, 2020, pp. 10-11).

In conclusion, Valentini and Sen claim to have demonstrated that superdeterminism can be quantitatively shown to necessitate an arbitrarily large degree of fine-tuning in multiple ways, given certain assumptions. This is argued to be the case even without the specific demand that the systematic correlations in the theory must reproduce the Born rule, which just goes to show another fine-tuning problem. I believe they are successful in this pursuit and through their work have added to the discourse on the conspiracy argument by both differentiating different mechanisms through which conspiracies turn up in superdeterminism as well as by quantitatively backing up the central intuition already clearly conveyed by Bell.

This, then, allows us to consider the responses to the conspiracy argument by supporters of superdeterminism. In the following, I will consider two responses by ‘t Hooft, three by Hossenfelder, one by Andreoletti and Vervoort and finally a general one. These arguments will be evaluated to judge whether the conspiracy argument holds up.

Response 1: the quantum-dependence argument

One first argument made by ‘t Hooft appears to stress that apparent conspiracies are a feature of quantum mechanics, but that they do not appear once one adopts the presupposed underlying superdeterministic theory. An idea of this sorts seems to be expressed in at least the following two places:

“We must conclude that, if there seems to be conspiracy in our quantum description of reality, then that is to be considered as a feature of our quantum techniques, not of

the physical system we are looking at. There is no conspiracy in the classical description of the cellular automaton. Apparent conspiracy is a feature, not a bug.” ('t Hooft, The Cellular Automaton Interpretation of Quantum Mechanics, 2016, p. 74)

“’Conspiratorial’ could mean the modification in the past that is required to modify the present. In a deterministic theory, this ‘conspiration’ is difficult to object against, so presumably what is meant here is that Nature conspired to adapt the wave functions as well to the new situation. The point we wish to make is, that in a deterministic theory there are no wave functions, …” ('t Hooft, The Free-Will Postulate in Quantum Mechanics, 2007, p. 4)

One first and smaller comment here is that ‘conspiratorial’ in the sense used in the second quote is different from how it has been used up until now. ‘t Hooft seems to refer to ‘conspiratorial’ as an external intervention affecting a causal chain. However, the meaning of ‘conspiratorial’ in the conspiracy argument is more so to be found in the universe being full of correlations between quantum systems and systems determining measurement settings that appear to be ‘just so’ that we always see Born’s rule being confirmed in Bell tests.

More importantly, both quotes seem to imply that the conspiracy issue seems to arise as a consequence of thinking in quantum mechanical terms. The validity of this argument hinges, at minimum, upon the premise that one must necessarily invoke quantum mechanical concepts to arrive at the conspiracy argument. Yet, I would argue that this is not the case. Let us consider a simple Bell test. While it was often stated that in such experiments we measure the ‘spin’ of atoms, this is not what is directly observed. Instead, what we observe is an atom having moved up or down when exposed to a Stern-Gerlach setup. This behavior is directly observable, independent from what theory one may think describes it. Bell tests ultimately come down to evaluating the fraction of events an atom went up or down when moving through the magnetic field present in different configurations. To show that hidden-variable theories are incompatible with these observed fractions, one only needs to assume that the atoms have real states, that they obey locality and that they are uncorrelated with the direction chosen for external fields in the Stern-Gerlach experiments. These assumptions are therefore not uniquely quantum mechanical, but they do yield Bell’s inequality. It then turns out the observed fractions are incompatible with Bell’s inequality. Keeping locality and realism, as superdeterminism does, statistical independence must be rejected, but this then necessitates the type of correlations that solely lie at the basis of the conspiracy argument. In conclusion, it does not seem warranted to view conspiratorial outcomes as a remnant of quantum mechanics, as the charge at hand can be levied entirely without reference to it.

A stronger version of ‘t Hooft’s argument, however, may concern type I models of superdeterminism specifically. One may remember that while type II models attribute the correlations in superdeterminism to the initial conditions of the universe, type I models opt for interactions in the very early universe being responsible for them instead. As was explained in chapter 6, we can show experimentally that these correlations must have been created a long time ago, likely during the very early universe. ‘t Hooft appears to claim that conspiracies will not appear any longer if one adopts a superdeterministic over quantum mechanical framework. Another way of interpreting this is then that the correlations many refer to as conspiratorial are no longer conspiratorial if a clear causal law that created the correlations can be provided. The correlations are not bizarre coincidences, but perhaps result of a causal law that operates in a way that it systematically creates these correlations. This way of looking at ‘t Hooft’s comments can be supported by the following passage:

“Rather than saying there are ‘spooky signals’ going around, we could also say that the laws of nature cause correlations, these correlations may even be controlled by various types of conservation laws.” ('t Hooft, The Fate of the Quantum, 2013, p. 6)

On the one hand, this seems to be a reasonable strategy to avoid the conclusions of the conspiracy argument. As an intuition pump: most physicists will not claim that the universe must be conspiring against them because gravity exists, always making the universe behave ‘just so’ that massive objects attract one another. This hints back to the common distinction between laws of nature and initial conditions, with the laws often being characterized as necessary. This metaphysical notion of necessity here seemingly functions to immunize physical laws against the charge of being fine-tuned aspects of nature in the debate on fine-tuning. These attacks tend to be reserved for the initial conditions, categorized as contingent. After all, if prior to the establishment of the law of gravity, someone had claimed that objects on Earth always fall downwards due to the coincidence that the initial conditions of the universe happened to be the way they were, many would react to this person with the charge of bizarre fine-tuning. Therefore, ‘t Hooft, advocating for a type I superdeterministic model, might reasonably claim that the correlations resulting from the framework do not display a heavily fine-tuned universe on the basis of its initial conditions, but are merely the logical result of a physical law we have not yet grasped. From this point of view then, this physical law must in the early universe have created the superdeterministic correlations resulting in the quantum-statistical outcomes of Bell tests now, just like the physical law of gravity always causes correlations between the accelerations of massive bodies. Wherein, then, lies the conspiracy? Following (this interpretation of) his line of thinking, either ‘t Hooft is correct in there being no conspiracy in his superdeterministic model, or every law of nature is equally fine-tuned. In the latter case, why single out superdeterminism rather than either treating the conspiracy argument as a ubiquitous problem in physics, or re-evaluating the legitimacy of fine-tuning arguments in the first place?

On the other hand, one may also levy some arguments and comments against ‘t Hooft’s type I defense. I shall briefly name some of these I think could be considered, structuring them in a list:

- Quantitative arguments for fine-tuning also work against at least some type I models. This point is made explicitly by Valentini and Sen in their previously covered second paper on the superdeterministic conspiracy (Sen & Valentini, 2020, pp. 2-3). Thus, invoking a physical law as the origin of the correlations does not disperse all forms of fine-tuning associated with superdeterminism.
- Continuing with the work of Valentini and Sen, the authors argue that a law functioning to establish quantum equilibrium, as is the case with ‘t Hooft’s law in the very early universe, “may get delayed due to several factors” (Sen & Valentini, 2020, p. 11). This could lead to testable deviations from quantum mechanics, and since experiments such as the cosmic bell test use objects like quasars that are no longer in causal contact, this nonequilibrium situation cannot be fixed locally.
- Even when considering the more qualitative way of looking at the state of affairs as put forward by Bell, the ‘absurdity’ inherent to the superdeterministic worldview remains unchanged by ‘t Hooft’s approach here. If one deems the existence of compounded correlations between, for example, the TV-show pixels and the set of later created atoms whose spin I measure just metaphysically unacceptable, one will not be swayed by whether this outcome results from a type I or type II superdeterministic model. Their problem will be with the consequences of any superdeterministic view as opposed to those other theories and interpretations of quantum mechanics.

- Earlier, it was established that type I superdeterministic models are susceptible to empirical constraints. Specifically, it was explained that Bell tests using light from distant objects can push back the maximum time since the birth of the universe at which superdeterministic correlations must have been locally created by this causal law. Every test so far has moved this maximum closer to the big bang, and we already know that it would have had to happen at least half the age of the universe ago. Extrapolating, either we may sooner or later definitively falsify type I models, or the maximum can be pushed back arbitrarily close to the big bang. However, that would mean the causal law responsible for all the correlations was causally efficacious only at that moment, apparently never to be 'seen' again thereafter, at least from the perspective of the requirements this law would necessarily need to fulfill in a type I superdeterministic model. Perhaps one question one may then ask is whether a relevant distinction between type I and type II models remains in the first place. That is, how do we conceptualize the distinction between initial conditions and a law acting *for only an infinitesimal amount of time after the big bang* and functionally disappearing after that? While not necessarily devoid of a possible solution, this may be a problem type I supporters need to consider. Some thoughts related to questions such as this will be provided in the final bullet point.
- 't Hooft intends for his cellular automaton framework to be an interpretation of quantum mechanics, reproducing empirical predictions equivalent to those of any other interpretation. Given this fact, we run into the epistemological issue that beyond just an interpretation of data, we have a physical law that is untestable. This unfalsifiability with respect to a law of nature may be deemed exceptional and unacceptable by some, while at the same time some may invoke the principle of parsimony as an argument not to introduce physical laws one cannot test. Nevertheless, 't Hooft may rightfully object that this holds for many interpretations of quantum mechanics that aim to expand on 'QM_0'.
- It may be argued that 't Hooft's proposed law, independent of what it may look like, is of a strange character compared to other laws. The laws of nature, be it Faraday's law, the law of gravity or conservation laws, tell us how nature behaves given particular contexts. But rather than being born out of inductive observations of regularities in nature, 't Hooft's law is *required* to produce the familiar superdeterministic correlations. These correlations, in turn, are required for us to always observe quantum statistics in Bell tests,

instead of at least sometimes observing different statistics that would follow from Bell’s inequality. This seems to give the law a somewhat teleological character, with the telos being the display of the Born statistics. While some may consider the law ‘off’ in this way, ‘t Hooft could reply that such a view exists merely in the eyes of the beholder. We focus our attention on this in the next response.

- A comment that should be made is that we are primarily considering superdeterminism as a framework, rather than individual superdeterministic models. Much more may be said on the topic of this causal law as it appears in ‘t Hooft’s cellular automaton model and how it mathematically functions. Such an investigation is worthwhile, but beyond the scope of this work.
- Lastly, one may wonder whether it makes sense that if we replace a fundamental number (such as an initial condition of the universe) with a law, the problem of fine-tuning suddenly disappears. Why is it, that the question is always asked how it is the constants of nature are just so that live can thrive in our universe, and not how it is that the specific combination of lawful behavior of nature is just so that it can? This question may be even more prominent in the face of the unique character of ‘t Hooft’s hypothetical law. If we were to copy our world in every aspect expect the presence of the law of gravity, we know that life as we know it could not have arisen. A similar point can be made with respect to the superdeterministic discussion. One may ask why the existence of particular initial conditions leading to particular results are more conspiratorial than the existence of a law doing so.
  If this distinction does not hold up when considering the universe in its entirety, a type I superdeterministic model provides no more protection against the conspiracy argument than a type II model does. However, one need not agree with this way of looking at things. Perhaps the distinction is not just a pragmatic way of modelling physical systems, but a metaphysical truth at the most fundamental level. It may be that a universe can be just as consistent with different initial conditions (thereby being contingent), yet with the laws creating one complete and internally consistent description of nature (thereby being necessary). Taking these away, one may then introduce contradictions resulting in the logical impossibility of having it any other way, as if taking away an essential wooden beam from a Jenga tower. This is ultimately a metaphysical question I do not have the answer to, but nevertheless one a supporter of type I superdeterministic theories may consider.

In summary, there are several possible arguments that challenge the notion of a type I superdeterministic model getting around the conspiracy argument. Especially with respect to the argument as it was conceived by Bell and Valentini & Sen shown earlier this chapter, it is not clear how the conspiratorial difficulty disappears with a type I model. Therefore, although sensible, it is unlikely to rebut the conspiracy argument.

Response 2: the vagueness argument

We then move on to another comment of ‘t Hooft on the conspiracy argument. Here, he appears to argue that seeing a conspiracy in nature according to superdeterminism is a subjective matter. He states that:

“They howl at me that this is ‘super-determinism’, and would lead to ‘conspiracy’. Yet I see no objections against super-determinism, while ‘conspiracy’ is an ill-defined concept, which only exists in the eyes of the beholder.” ('t Hooft, Free Will in the Theory of Everything, 2017, p. 5)

A similar comment is made at a different place:

“This explanation is usually dismissed. It is called a ‘conspiracy theory’, and that is considered to be disgusting. But are ‘disgusting’, or ‘ridiculous’, valid arguments in a mathematical proof?” ('t Hooft, The Fate of the Quantum, 2013, p. 6)

To evaluate ‘t Hooft’s comment, two questions must be answered:

- Can a reasonable, by all agreed-upon definition of conspiracy be provided?
- Is the applicability of this conspiracy concept a subjective matter?

I do not think ‘t Hooft is justified in his claim that ‘conspiracy’ is ill-defined. The use this term is more of a framing device, aiding a particular intuition for a problem not itself ill-defined. One can imagine a practically infinite number of contingent possible distributions of correlations, between the practically infinite number of things can be used for Bell tests and require such correlations. Yet, the comparatively extremely small number of distributions perfectly reproducing quantum statistics every time just so happens to be the actual one. So far, while not necessarily ill-defined, the argument appeals to an intuition about the sizes of sets of possible distributions. However, with Valentini and Sen, quantitative means to express this problem have also been produced. I would argue this removes the remaining credibility of denoting the problem as ill-defined in the absence of further argument.

Similarly, the conspiracy exists beyond the eye of the beholder. If the problem is well-defined, and at the same time refers to an observation about the state of physical reality independent of what any observer thinks about this, I do not see what remains with which it can be called ‘subjective’. One way in which this could still be done is to state that for some reason all fine-tuning problems in physics are ultimately not real problems, but merely perceived at such. This is the next counterargument that will be looked at.

Response 3: the probability-conditions argument

Evaluations on the legitimacy of fine-tuning arguments in physics are much broader than just superdeterminism (Landsman, The Fine-Tuning Argument, 2015) (Hossenfelder S. , Screams for Explanation: Finetuning and Naturalness in the Foundations of Physics, 2018). However, when looking at Hossenfelder’s critiques of the conspiracy argument, we will start by considering how she attempts to rebut it by applying more general arguments against fine-tuning. In one of her papers on the topic she states:

“Finetuning arguments can be used when one has a statistical distribution over an allowed set of parameters and a dynamical law acting on this distribution.” (Hossenfelder S. , 2020, p. 15)

Hossenfelder does not expand much on this specific idea in this particular article replying to arguments against superdeterminism. Nevertheless, I will attempt to steelman the argument because I do think there is something there, and that she has made this argument in other places as well. The argument is, namely, to the one made in general against fine-tuning arguments in physics. On that, she says the following:

“These arguments are not mathematically well-defined because they refer to probabilities without a probability distribution.” (Hossenfelder S. , 2018, p. 15)

Let us unpack the meaning of this statement. At the core is the idea that fine-tuning arguments presuppose a number of conditions to hold that we are not epistemically entitled to. This is due to the appeal to probability theory made by fine-tuning arguments in their claim that a particular set of constants of nature, initial conditions (against superdeterminism) or even physical laws are somehow ‘improbable’ and thus require explanation. However, three conditions required for the notion of probability to carry meaning in the first place are as follows (Landsman, The Fine-Tuning Argument: "What-if" Physics, 2021):

1. *A space of possible outcomes*: to make meaningful statements about the probability of one outcome over another, there need to be multiple possible outcomes.
2. *A probability distribution on this space*: likewise, given that probabilities can be distributed over a space uniformly, Gaussian, etc., there needs to be a measure of how likely a particular outcome is over others.
3. *(Repeated) random sampling*: generally, one can take a number of individuals within a set in a way that every individual has equal chance of being taken, each of them carrying particular properties, and in this way estimating the distribution of properties in the set. To give a more intuitive example: we obtain the probability distribution of a coin flip by averaging over the hidden state variables, i.e., the initial conditions relevant to the coin toss. In other words: one is repeatedly randomly sampling.

While lacking even one of these conditions already causes trouble for the application of probability theory, we can observe that fine-tuning arguments, including the conspiracy argument, gets in trouble with each of them:

1. How can we know for sure that parameters like the initial conditions of the universe or the constants of nature could have been different? That there is some allowed range of values for all entries in these categories that all might have been chosen to 'create' a universe? We only have access to our own world, so this does not seem to be a given. In the case of initial conditions: we know empirically that we can test the law of gravity by throwing up a ball with different initial velocities. Creating universes with different initial conditions, however, seems to be beyond our abilities.
   Perhaps they could not have been different. This was also Einstein's view. On the fine-tuning problem in relation to the constants of nature being the way they are, he stated:
   "There are no arbitrary constants of this kind; that is to say, nature is so constituted that it is possible logically to lay down such strongly determined laws that within these laws only rationally completely determined constants occur (not constants, therefore, whose numerical value could be changed without destroying the theory)." (Einstein A. , Albert Einstein: Philosopher-Scientist ("Autobiographical notes"), 1949)
   Thus, if there exists an ultimate Theory of Everything, the constants may all be derived from its logical framework, leaving no room for any contingency. Having namedropped Einstein once again, let us consider an example of this

line of thinking involving no other than Bohr, concerning the story of the 'Rydberg constant'. At the end of the 19^{th} century, it was found empirically by Johannes Rydberg (1854-1919) that the wavelengths of radiation emitted by the Hydrogen atom could be calculated through a formula containing a particular unexplained numerical constant that was measured. Perhaps this was another constant of nature. Later, however, it was found that this 'Rydberg constant' could be theoretically derived from the atomic model introduced by Bohr. Bohr had shown that the Rydberg constant was not a fundamental constant of nature, but was composed of other, already familiar, constants. The thought, then, is that progress in fundamental physics may increasingly show what we mistake for 'contingent constants of nature' to be necessary components of the fundamental theory.

While this is by no means certain, it goes to show that the existence of a space of possible outcomes, in our case a space of possible initial conditions giving rise to the necessary correlations to get superdeterminism to be compatible with experiment, is not a priori given. The fine-tuning argument then rests on an unproven assumption that need not be true.

2. Fine-tuning arguments, on top of a space of possible outcomes, typically assume a uniform probability distribution. This distribution is characterized by each element in the outcome space having equal probability of occurrence to any other. Given that the outcome space consists of all possible sets of initial conditions (conspiracy argument) or physical constants (traditional fine-tuning argument), it is then argued to be highly unlikely that we find ourselves in the small subset of outcomes conducive to systematic Bell inequality violations and the existence of life, respectively.

   Yet again, there is no a priori justification for this implicitly assumed uniform probability distribution. Perhaps the probability distribution is Gaussian, with a small width and high peak centered around the outcome we observe in our universe. One might protest, from Occam-like considerations, that this adds an extra structure that should not be assumed in the absence of evidence. But there are problems with this response. For one, the uniform probability distribution implicit in fine-tuning arguments, is itself a structure. Moreover, Occam's razor, or the principle of parsimony, is a useful guideline in everyday and often scientific reasoning, but would be misinterpreted when read as a metaphysical statement about the world with certainty being as simple as we imagine it can be. Especially given a topic such as the coming-into-being of our very universe, we should be extremely careful with making judgments on the basis of our biologically evolved intuition on something like ‘the probability distribution for the initial conditions of the universe’. The above is, thus, not say that such a Gaussian probability distribution would be likelier than a uniform one, but that there is no probability distribution ‘in sight’ at all. The problem for those advancing fine-tuning arguments, insofar these employ probability theory, is that they rely on a probability distribution over the space of possible outcomes being given.

3. The final entry can be kept brief, by stating that randomly sampling universes is simply not something we can do.

Therefore, it can be argued that this puts the advocates of the conspiracy argument in a difficult position. Its applicability range is suddenly limited by its reliance on the conditions for their legitimate use of an argument appealing to probability being there. These conditions here have significant metaphysical implications, as seen above. These metaphysical assumptions may be held by many, but are at least contestable nonetheless. For example, the idea that the initial conditions giving rise to the required correlations in superdeterminism could have been otherwise is an unverifiable statement. In summary, such statements are beyond what is knowable, so the conspiracy necessitates a commitment to metaphysical assumptions one is technically not entitled to make.

The question now is how this all impacts the conspiracy argument. This essentially comes down to how believable one thinks the metaphysical implications resulting from applying the conditions for probability theory to the superdeterministic conspiracy are. For reasons discussed before, no definitive answer is possible on this. Three comments, however, may provide some useful context on this nonetheless.

Firstly, the traditional fine-tuning problem is generally recognized as a legitimate problem, with many physicists even going so far as to propose a multiverse in an attempt to solve it[63] (De Vuyst, 2020). The argument advanced by Hossenfelder to

[63] This strategy lays bare an important difference between the common fine-tuning argument on the basis of constants of nature allowing for life, as opposed to the conspiracy argument where the fine-tuning is to be found in the correlations resulting from the initial conditions of the universe reproducing quantum statistics in all experiments. While not being an uncontested solution itself, in the former, one could argue that there are many different universes each with their own distinct values for the constants of nature. It is then no wonder that we find ourselves in a universe with constants allowing for the emergence of life, as we need to be alive to ask the question in the first place (the so-called ‘weak anthropic principle’). However, no parallel exists for the case of superdeterminism. As far as we can tell, there is no obvious reason that life should not be able to emerge in a universe where the correlations are such that it does not reproduce quantum statistics *all the time*. Given that the

rebut the conspiracy argument, however, can just as well be used to rebut the traditional fine-tuning argument on constants of nature. Nevertheless, many remain unconvinced, and it reasonably follows that most physicists who acknowledge the existence of the fine-tuning problem will likely also acknowledge the conspiracy argument. This is, of course, not an argument for its correctness or not, but it does give a hint as to what the impact of the probability-conditions argument on the perception of the conspiracy argument will be.

Secondly, the argument does away with the traditional distinction between contingent initial conditions and necessary laws of nature. If it is the case that the initial conditions of the universe could not have been otherwise, i.e., that there was no space of possible outcomes, then these are by definition no less necessary than physical laws. As was argued by Valentini and Sen as well, this is a central principle of physical theories.

Finally, I think that something like the realization of having to, for example, presuppose a space of possible outcomes in relation to the initial conditions of the universe, does not necessarily render one *fully* unable to make any meaningful comments about whether this assumption is reasonable or not. When considering the unification argument for constants of nature by Einstein, contemporary physics does not seem to obviously support it. Not only does the Standard Model of particle physics have 19 free parameters[64] in addition to the constants of nature, some proposed frameworks for a Theory of Everything, such as string theory, do not at all appear to resolve the issue (Friederich, 2021). This reduces the strength of a proactive argument in opposition to the existence of a space of possible outcomes for constants of nature. When it comes to the initial conditions of the universe, most relevant to superdeterminism, not much can be said about it in the first place. However, I would argue that from an epistemic point of view it may make more sense to believe that the initial conditions, like all other initial conditions of systems considered in physics, could have been otherwise. As a general principle, we tend to

initial conditions could conceivably be such that the correlations obey quantum statistics up until the universe is 14 billion years old, only then diverging, this actually seems rather unlikely. The multiverse – weak anthropic principle combination is thus not a viable solution to the conspiracy argument.

[64] A free parameter is a number that is needed for a physical model to work, the number itself not being determined by the model. They can, instead, be found experimentally. Examples in the standard model are the masses of elementary particles and the coupling strengths of fundamental forces. Whether these parameters will ultimately be explained in terms of a final theory and more fundamental constants of nature, or whether they themselves will ultimately appear to be brute facts like constants of nature seem to be, is as of yet unknown.

assume options are open unless we have sufficient reason to rule out others. The claim that there is no space of possible outcomes for what the initial conditions of the universe can be, even though having that space is usually how initial conditions are characterized, imposes a restriction to the ‘freedom of the universe to choose its initial conditions’. In the absence of a sufficient reason for imposing such a restriction, one could make an epistemic argument that we should not exclude the possibility of alternative initial conditions out of hand. Considering the conspiracy argument, this would be my *belief*, but it is nevertheless good to acknowledge that this is still an assumption one has to make now to run the conspiracy argument. Similar thoughts can be had on the issue of the lack of a defined probability distribution. Demanding this clearly makes sense in the case that there are many universes. But suppose there is only our universe. One may then wonder if we must necessarily demand a probability distribution in this situation. Even if we do, we know at least that there are more conceivable probability distributions for which our universe is unlikely in the way described by the conspiracy argument (and traditional fine-tuning argument), then ones for which that is not the case[65]. Nevertheless, an argument which relies on this observation is guilty of again implicitly assuming a probability distribution, only now of second order. It makes the mistake of assuming

[65] While not providing a hard proof, the outlines of one can be argued for in the following way. This footnote can be skipped if one is willing to accept the claim it is attached to.
Rather than taking the initial conditions of the universe, for simplicity, let us just imagine the value of one physical constant for this example. Suppose the value of this constant in the actual universe is *c*. Now imagine a graph with the domain of possible values for this physical constant on the x-axis, and the probability density of any particular value being realized when a universe is created on the y-axis. Knowing the significant orders of magnitude that constants can differ from each other with in the real universe, e.g., in the example of elementary particle masses, it is reasonable to let the domain of the x-axis encompass many more orders of magnitude than that of *c.* That is, we assume that the value of any physical constant could have been otherwise to high degree. Yet, we have seen that there are many physical constants whose value does not need to change *that* much compared to the width of this domain, for the world to be significantly different in some relevant way, e.g., allowing life or not. Then, for *c* to be unlikely, the probability distribution should be such that the probability density is low near *c.* But when *c* is likely, in which case we have a probability distribution for which the fine-tuning argument would fail, it would be the case that the probability distribution peaks significantly around *c.* Yet, especially given the many orders of magnitude encompassing the domain of the physical variable, and even putting uniform-like distributions aside, there are many more probability distributions that peak around values sufficiently removed from *c*. Here, sufficient removal entails leaving the world in a significantly different state as mentioned before. Therefore, we can see that there exist many more probability distributions for which it is unlikely that we find ourselves in a world with *c*, than those for which it is likely. In the example traditional fine-tuning argument, we can compound the unlikeliness of this particular physical constant with all of the others that also must inhabit a relatively small range to allow for life. The likelihood of then occupying a universe that allows for us to exist becomes vanishingly small to even higher degree. And this unlikeliness is precisely what fine-tuning arguments have questions about.

the probability distribution, specifically the second order one describing the likelihood with which any of the many possible probability distributions that could govern the selection of initial conditions for universes is the real one, is uniform.

While supporters and detractors of superdeterminism can thus go back and forth on such matters, the reality is that knowledge of whether the conditions needed for making probabilistic arguments in relation to the subject at hand is beyond what we can know. In conclusion, I do not believe the probability-condition argument is fatal to the argument, nor does it dismantle the central conspiratorial intuition. Yet, it does arguably make it depend on a number of metaphysical assumptions that follow from the conditions for the application of probability theory. While I have advanced reasons for suspecting that many will, nevertheless, be inclined to accept these assumptions, the dependence thereon inevitably takes away some overall strength of the conspiracy argument. This makes the probability-conditions argument, in my view, the strongest response to the conspiracy argument.

Response 4: the time-independence argument

This argument will focus on the interpretation of statistical dependence in the context of Bell tests. According to Hossenfelder, statistical dependence only needs to hold “at the time of measurement”. This, she takes to mean that it is not relevant what “the exact way in which these settings came about” was, but only what the detector settings are when the measurement happens (Hossenfelder S. , Superdeterminism: A Guide for the Perplexed, 2020, pp. 3-4). Therefore, fears of ubiquitous conspiratorial correlations between, e.g., quasar light and the spin state of an atom measured on Earth are ungrounded.

Hossenfelder is indeed correct that one only needs to specify the measurement settings, and not the way these came about, to determine the results of a Bell test in superdeterminism. The above quote by Hossenfelder is a reply to what she considers to be a “confusion” by those believing in the conspiracy argument. However, I would say that she thereby misunderstands the conspiracy argument. It is not that they deny that only the actual settings need to be specified for the calculation of the outcome. They merely take determinism to its logical consequence by showing the inherent connection between the settings and the setting-mechanism. The core of determinism is that if one specifies a given state *at any time*, its evolution can be determined for all past and present times. But that also means that the correlation between settings and the hidden variables must already have been encoded through the correlation between what determined the measurement settings and the hidden variables. This was also pointed out by Valentini and Sen (Sen & Valentini, Superdeterministic hidden-variables models I: nonequilibrium and signalling, 2020, p. 2), who also pointed out that fine-tuning is required for the choice of setting mechanism not to make a difference. The correlations between settings and the hidden variables do not come out of nowhere but are embedded in the past, and this embeddedness implies a conspiracy. Most critics of superdeterminism are thus perfectly aware that "Bell's theorem makes statements about the *outcomes* of measurements" (Hossenfelder S. , Superdeterminism: A Guide for the Perplexed, 2020, pp. 3-4).

In addition, Hossenfelder states that conspiracy argument proponents are involved in a confusion that "comes from mistaking a correlation for a causation" (Hossenfelder S. , 2020, p. 3). I would argue that this is also misunderstands of the conspiracy argument. The argument never refers to a causal influence of measurement settings on the hidden variables of quantum systems measured. The argument recognizes that superdeterminism introduces pre-existing correlations between these two through a type I or type II model, but for the reason laid out in the above paragraph this is the source of the conspiracy.

Finally, it could be argued that Hossenfelder herself contradicts her reasoning for ubiquitous conspiratorial correlations not having to be present earlier in the paper. She writes:

"What does it mean to violate Statistical Independence? It means that fundamentally everything in the universe is connected with everything else, if subtly so." (Hossenfelder S. , 2020, p. 3)

Yet, it is precisely because of the infinity of conceivable setting mechanisms that everything "has to be connected with everything else". The pixels of my TV would not have to take part in superdeterministic correlations if statistical dependence was truly limited to only the settings and the state of a system at the time of measurement. For these reasons, I do not think Hossenfelder is justified in her claim that the way settings came about does not matter in superdeterminism.

Response 5: the positivistic argument

Almost all of section 9 of Hossenfelder's paper addressing critics of superdeterminism does not focus on the probability-condition argument, but on what I will refer to as the 'positivistic argument'. I will briefly summarize her argument, but then go on to argue that is has no bearing on the conspiracy argument as laid out in this thesis, because she rejects the metaphysics of it and utilizes a different concept of fine-tuning.

Hossenfelder claims that “a model is fine-tuned if it lacks explanatory power” (Hossenfelder S. , 2020, p. 17). She then goes on to explain that models like the Standard Model are very successful in this way, despite having many free parameters. Therefore, they are not fine-tuned. She subsequently states that the fine-tuning established by Valentini and Sen says nothing about the explanatory power of a model, and this means there “isn’t much to learn from this”. A final key point is that in an earlier paper she explains that she understands the ’explanatory power’ of a model in terms of whether model “allows one to calculate measurement outcomes in a way that is computationally simpler than just collection the data” (Hossenfelder & Palmer, 2020, p. 11). She concludes that she believes superdeterministic models can do this and refers to a toy model[66] (Donadi & Hossenfelder, A Toy Model for Local and Deterministic Wave-function Collapse, 2022).

It is far from obvious that superdeterministic models have, in this sense, more explanatory power than other interpretations of or alternatives to quantum theory. The real problem here, however, is not necessarily the logical steps of the argument, but rather, the definition of fine-tuning on which it is based. Fine-tuning is generally defined as something having to take on very specific values to be compatible with experiment or a particular state of affairs (Landsman, 2015) (De Vuyst, 2020) (Friederich, 2021). This central property is what is emphasized by proponents of fine-tuning arguments. As an example for the traditional fine-tuning problem, we noticed that had the gravitational constant been somewhat, the structures enabling the emergence of life could never have formed[67]. Similarly, in superdeterminism, the initial conditions of the universe need to be just so that it implants correlations between every quantum system and potential measurement setting mechanism that will ever lead to a measurement sometime, somewhere in the universe. A conspiracy

[66] A ‘toy model’ is a simplified description of something, such as a physical mechanism or process, leaving out difficult details that distract from the main (often didactive) point of the model. The downside is that, while useful for many practical purposes, a model like this is not fully representative of the real world, and neither are the hidden variables contained therein. For now, superdeterminism is mostly confined to such toy models.

[67] Moreover, the universal gravitational constant is by far not the most fine-tuned one (Adams, 2019, p. 140).

of that magnitude is likely to be compatible with only a vanishingly small range[68] of possible values for the initial conditions of the universe.

Hossenfelder's definition of fine-tuning therefore circumvents these considerations, rather than addressing them. This probably is no coincidence, as near the beginning of her chapter on the conspiracy argument she notes:

"These other notions of finetuning, however, are meta-physical, meaning they have no empirical justification. One can certainly define these other notions of finetuning, but there is no reason to think that a theory finetuned in such a way is problematic." (Hossenfelder S. , 2020, p. 14)

It is true that the conspiracy argument as laid out in this thesis falls in the metaphysical category. Thus, Hossenfelder is correct that if one adopts a positivistic attitude towards physics where all metaphysical considerations are rejected or the belief that science should only concern itself with empirical adequacy, the conspiracy argument does not do much *directly*. An extensive discussion of the role of metaphysics in physics would leave superdeterminism as a topic and rather be concerned with fundamental topics in the philosophy of science. Therefore, I will just point out a number of relevant factors in response to this view.

Firstly, this view necessitates that superdeterminism must be empirically distinguishable from quantum theory. It cannot just be an interpretation. If this were the case, it would primarily be a metaphysical project, aimed at finding the correct interpretation of physical reality, without empirical justification. Hossenfelder thinks this empirical distinguishability does apply to her model and we will return to this the next chapter. However, some other authors, such as 't Hooft, do not think superdeterminism must be empirically distinguishable from quantum theory. For them, this would therefore not be a consistent line of attack with which to discard the conspiracy argument.

Secondly, one may question whether Hossenfelder is not contradicting herself here. In her previous paper on superdeterminism she states that "if one is, by contrast, willing to accept the consequences of realism, reductionism, and determinism, one is led to a theory in which the prepared state of an experiment is never independent of the detector settings" (Hossenfelder & Palmer, 2020, p. 2). This seems to imply realism, reductionism, and determinism to be motivations for adopting

[68] This does of course presuppose the conditions of probability theory to hold, but this was discussed previously, and we need not concern ourselves with this aspect of the issue now.

superdeterminism. Yet, all three of these are metaphysical views. They are views on what physical reality is ultimately like that cannot be directly (dis)proven by experiment. Another example would be that in using the probability-conditions argument, Hossenfelder will likely at least be inclined to say that the initial conditions of the universe, as an example, could not have been otherwise. This is true for anyone who rejects the conspiracy argument on the basis of it presupposing a space of possible outcomes. Yet this is also a metaphysical conviction, as we already discussed goes far beyond what can be known. Therefore, Hossenfelder seems to be motivated by metaphysical principles, while simultaneously rejecting arguments within that category.

Thirdly, I would argue against the view that metaphysical arguments, that is, arguments about what physical reality is like but cannot be directly empirically justified, have no bearing on physics. There are multiple reasons for this, and I will name a few. One is that our physical theories are grounded in metaphysical principles. These principles, therefore, guide the theories we are likely to conceive and investigate. One of many examples may be the history of action at a distance, the conception and perceived (im)possibility of which has changed throughout the history of physics (De Regt, Understanding Gravitation, November 2021). Because our theories are grounded upon central ideas about what nature is like, even though they can never be fully verified, these principles will inevitably be reasons to reject some theories over others. Another reason may be that theories producing seemingly bizarre metaphysical consequences are sometimes rejected on these grounds. While being ‘conservative’ on other principles, as seeing earlier in this thesis, Einstein did do away with the absoluteness of space and time. Many initially rejected relativity because of this, but he turned out to be correct. While this may seem to be a descriptive claim that, if anything, supports the idea of loosening metaphysical commitments, I would argue this famed example should be understood as the exception to the norm. When considering our web of beliefs about physics, it is hardly common or sensible practice to immediately reject the ‘hard core’ (in Lakatosian terms) in the face of evidence conflicting with a theory. While sometimes the accumulation of evidence against a well-established physical theory becomes so great that it is abandoned to eventually be replaced by a better one, in practice it occurs far more frequently that seeming counterevidence is due to experimental errors, an incomplete understanding of the theory or some datum previously unknown but perfectly compatible with the theory. A famous example of the latter is the mathematical prediction of Neptune fixing the apparent anomalies in Uranus’s orbit. Returning to the main point, frameworks such as multiverse theory and the many-worlds interpretation both produce enormous ontological baggage. One may arguably do well to investigate alternative explanations carefully before taking this on and so profoundly change our view of what nature is like. There are, in fact, plenty of physicists rejecting multiverse theory, and it indicates something that it is not empirical grounds on which this is done. Lastly, metaphysical arguments may very well cause problems for theoretical frameworks in physics down the line. After all, historically, many topics assigned to the domain of metaphysics have eventually been moved to the domain of physics. Whether the universe is static or expanding used to be considered a metaphysical question, but by the 20th century, it turned out physics could determine which of these options was true. The universe is expanding, one of many examples of the permeability of the supposed hard line between physics and metaphysics. Thus, such debates should not be discarded and viewed as inconsequential to physics[69].

Ultimately, I do not judge Hossenfelder's positivistic argument to hold up against the conspiracy argument. She uses a different definition of fine-tuning from what is used by those supporting the conspiracy argument, thereby circumventing it. Moreover, she appears to do this out of a rejection of metaphysical arguments in their entirety, which I have argued does not hold ground.

Response 6: the argument from counterfactual analysis

Philosophers of physics Giacomo Andreoletti and Louis Vervoort argue in a recent paper that the series of coincidences required by superdeterminism are not so "strange" after all. They do this using counterfactual analysis, invoking analogies with philosophical literature on time travel (Andreoletti & Vervoort, 2022, pp. 7-13). In this section, I will lay out this response to the conspiracy argument, and explain why I do not think it works.

The authors start by briefly explaining Bell's theorem and how superdeterminism gets around it by rejecting statistical independence. They go on to claim that the apparent absurdity resulting from the framework as expressed by the conspiracy argument can be represented by the following counterfactual:

"If the set-up B had been chosen, then the sub-collection b would have been selected." (Andreoletti & Vervoort, 2022, p. 8)

[69] And finally, if all else fails, one should remember that this is a philosophy thesis. Naturally, the metaphysical implications of superdeterminism are then subject of investigation!

Here, *B* refers to the detector settings, such as '12' or '32' in the terminology used in chapter 3. The *sub-collection b* then refers to the particular grouping of atoms, all with specified λ, whose spin ends up being measured in the experiment.

It is indeed the case that this counterfactual is true in a superdeterministic world. If the required correlations to reproduce quantum statistics are to appear, a setup cannot just be switched for another without also exposing the detectors to different spin states to measure, and vice versa. For a concrete example, one can test this out in the table that resulted from Mermin's experimental setup. Nature according to superdeterminism is such that it is not possible to select a sub-collection of atoms, each with a well-defined hidden-variable state determining their spin, in a way independent of the measurement settings. The consequence of this counterfactual is then the conspiracy laid out before.

However, say Andreoletti and Vervoort, this is not the relevant counterfactual to consider. They state that this point can be illuminated by using arguments that also appear in the philosophical literature on time travel. In this specific literature they reference, time travel to the past is conceived but it is demanded not to allow for changes in the past[70]. Killing your past self, as an example, would after all ensure contradictions. The only way to guarantee the absence of these contradictions given the possibility of time travel to the past is then through the appearance of particular events rendering the time traveler causally inefficacious in the past. This could be expressed through the following (paraphrased) counterfactual:

"If a time traveler P were to attempt to kill their younger self, they would slip on a banana peel, or their nerves would fail, or…" (Andreoletti & Vervoort, 2022, p. 10)

This counterfactual then must be true to prevent contradictions arising due to time travel to the past. With every possible attempt to change the past, 'coincidences' are bound to happen that stop it from working. Andreoletti and Vervoort recognize that this counterfactual seems very unlikely, but their claim is that the difficulty with its believability lies in the antecedent, not the consequent. They elaborate on this using another example, contrasting the sentences "If P were to throw a stone at the window, P would slip on a banana peel or hit a passing bird or…" and "If P were to throw a stone at the window but the window did not subsequently break, then P

[70] Some other literature proposes time travelers to create or journey to other universes, as a kind of many-worlds interpretation of time travel. However, the authors' point is not the logical necessity of this demand, but rather the arguments that will ensue from this metaphysical perspective on time travel.

would slip on a banana peel or hit a passing bird or…”. While the former sentence seems strange, the latter sheds light on the issue. Because the antecedent circumstance of throwing a stone at a window yet it not breaking describes some state of affairs recognizably difficult to make true, the coincidences in the consequent suddenly seem far more appropriate. The thing difficult to realize is the antecedent, and given its rarity the consequent logically follows. It is then argued that the time traveling counterfactual is like the latter sentence, in the sense that a world in which this model of time travel applies and where there is an agent perfectly able to go to the past and carry out this mission yet fails, is the part containing the difficulty. The ‘coincidences’ happening that prevent the success of this assassination then turn out to be a reasonable consequent.

The authors now intend to show that, using this reasoning from the philosophical debate on time travel, the conspiracy argument can be rebutted in the same way. They propose to modify the original counterfactual in the following way:

“If the set-up B had been chosen and nature is local, then the sub-collection b would have been selected.” (Andreoletti & Vervoort, 2022, p. 11)

Andreoletti and Vervoort now argue that with respect to the antecedent, “those described circumstances *require* unlikely coincidences to be true”. The difficulty is with the demand of locality. It is for this reason that the conspiracy argument is nullified, as it is argued that the apparent highly coincidental correlations it requires are very much sensible given our demands.

The immediate reply here is that in drawing this conclusion, Andreoletti and Vervoort simply assume what they set out to argue: that nature is local in the PL1 sense and statistical independence does not hold. If this is imposed on the universe, the apparent coincidences are just the result of that. However, the point of the whole debate is that we *do not* know with certainty what interpretation of (or theoretical alternative to) quantum theory is correct. One way to evaluate their strengths and weaknesses is precisely to critically examine the acceptability of their physical, metaphysical and epistemological consequences. In the case of superdeterminism, we would find extreme fine-tuning. By already assuming the antecedent, are Andreoletti and Vervoort not putting the cart before the horse?

The authors reply to the charge that they are begging the question as follows. I will provide the entire quote in order to represent it as accurately as possible.

"Once the worry about the 'coincidental' nature of superdeterminism is spelled out, i.e., once we realize that the relevant counterfactuals within superdeterminism are those such as the latter type and that there is nothing remarkable or surprising about these being true, then there should be no resistance in accepting these types of counterfactuals as true. The coincidental appearance is just what is required by the fact that it is difficult to make the antecedent true. If so, the option of rejecting statistical independence (superdeterminism) and the option of rejecting locality (other interpretations of quantum mechanics) are again on a par and further inquiry is needed to settle the issue." (Andreoletti & Vervoort, 2022, p. 12)

I do not believe that the above successfully refutes the considerations it attempts to address. The claim is essentially that while the first counterfactual seemed false, the reformulation carries with it the same logical structure as the stone throwing example and thus is true. However, this appears to both confuse logical validity with factual accuracy and deny the validity of all *reductio ad absurdum* arguments, used ubiquitously and successfully in philosophy and science alike.

The nature of this 'confusion' can be illuminated using an example from Aristotelean syllogisms. Here, there is a distinction between the logical validity of the syllogism on the one hand and the factual accuracy of its premises on the other hand. A syllogism may be logically sound, yet its premises may not hold in reality. If all dogs are birds, and all birds lay eggs, then all dogs lay eggs. This also applies to the stone-throwing and time travel counterfactuals before. The stone throwing counterfactual makes sense *precisely because* it was presented as an example where the antecedent was absolutely given, rather than being completely open to many different accounts. This and the superdeterministic counterfactual are equivalent in their logical validity, but there is simply no reason to assume the antecedent in the latter case. And if it is meant to be implied that the consequents of distinct antecedents cannot be used to evaluate the likeliness of the distinct antecedents being true, we get to the second objection.

One important form of argument in logic is the *reductio ad absurdum*. It works by showing that a given proposition P leads to a contradiction or absurdity, such that we may conclude that ¬P must be true. Arguments of this form are ubiquitous in science, philosophy, and even everyday life. They tend to have a good track record in producing fruitful ideas or filtering out bad ones. That is to say, that it is commonplace to consider existing hypotheses with which one attempts to explain something, and seeing what they would imply by thinking them through. If one hypothesis produces absurd consequences, this is then taken as a reason to reject it in favor of others. This is exactly what we are doing now, with distinct interpretations of and alternatives to quantum mechanics leading to very different ways of viewing nature. Contradictions and absurdities resulting from these frameworks are then often given as reasons to discard them in favor of others. Yet, while this (and the last) paragraph may seem very obvious, I am unsuccessful in interpreting Andreoletti's and Vervoort's statements in any other way than an attack on the logical argument of *reductio ad absurdum*.

Therefore, I do not think this argument shows that rejecting locality and rejecting statistical independence are on par, even if we somehow do consider these to be the only two options. One may think them to be on par, but this discussion can be had without sweeping the consequents of these propositions under the carpet of a question-begging antecedent.

Nevertheless, one last point of the authors is the claim that from an 'on par situation', superdeterminism might arguably be more likely, since other interpretations, contradicting relativity, reject locality (Andreoletti & Vervoort, 2022, p. 12). If one does not highly value the arguments against superdeterminism, this could work for pilot wave theory. But other than that, it ignores that most other interpretations of quantum theory still accept locality in the sense of PL2, the type required by relativity[71]. Thus, there are not necessarily any problems here. Lastly, and perhaps most interestingly, one can wonder whether the authors' argument cannot be turned against them, also shielding, for example, pilot wave theory from particular critiques. As explained before, pilot wave theory assumes the existence of a guiding wave permeating the entire universe. Some argue this is a strange and unnecessary piece of extra ontology. However, I can construct a counterfactual whereby this consequent is free from any implied absurdity, since the difficulty is in the anteceding assumptions of pilot wave theory and the particular physical situation one may describe in the counterfactual. Pilot wave theory is now on par with superdeterminism. Furthermore, this narrative does not end at pilot wave theory, since it can easily also be applied to the 'strange consequence' of many worlds existing in the Everett interpretation.

In conclusion, I do not believe the argument from counterfactual analysis holds up against the conspiracy argument.

[71] While this thesis does not cover informational interpretations of quantum theory (such as QBism), these are also relatively popular and also claim to be consistent with locality in all forms.

Response 7: the ubiquitous fine-tuning argument

A brief final counter to the conspiracy argument that will be covered here can be represented as follows:

"Fine-tuning problems are ubiquitous in physics. As shown earlier, this includes the standard model of particle physics and cosmology. Yet, we accept these theories as valid. Then why should we be worried by a little extra fine tuning in superdeterminism?"

While this argument is not tied to any specific author, it is fairly intuitive and has appeared in conversations I have had about this subject. For the sake of completion, I therefore nevertheless aim to address it.

I have three reasons why I do not think this line of argument is convincing against the conspiracy argument. These are as follows.

1. *Two wrongs don't make a right*. It is indeed true that because of its empirical adequacy and predictive power, the Standard Model of particle physics is well-accepted. At the same time, it is also true that there is broad consensus that physics needs to go beyond it. In general, two reasons are cited for this. On the one hand, there are phenomena like dark matter that are unaccounted for in the standard model. On the other hand, the Standard Model contains plenty of (finely tuned) free parameters that are experimentally determined rather than following from the theory. The latter issue is, thus, very closely related to the fine-tuning problem, and a reason given to go beyond the Standard Model (Westhoff, 2023). The same argument can be made for the standard model of cosmology.
   Therefore, the fact that these standard models are also plagued by fine-tuning problems is not a reason not to take these seriously in the case of superdeterminism. At the same time, if these problems were identical, it would at least show superdeterminism to be acceptable given the same degree of empirical success. The latter is not the case. But, more importantly for us, it can also be argued that the fine-tuning problems facing superdeterminism are worse both in *quantity* and in *quality*, which is the subject of the following two entries in this list respectively.
2. *Cumulative fine-tuning*. The fine-tuning typically associated with superdeterminism is that of the very specific correlations necessary to always yield quantum statistics in Bell tests, whatever set of atoms I choose to measure and whatever setting mechanism I adopt. Nevertheless, the

‘traditional’ fine-tuning problems in contemporary physics are not absent from the framework either. Superdeterminists do not claim that their theory explains why the electron mass is what it is, or that the universal gravitational constant *necessarily* follows from their theoretical construct. Consequently, these two types of fine-tuning add up. Even if one, for some reason, were not to consider the conspiratorial fine-tuning to be as bad as the traditional parameter fine-tuning, the sum of these problems is nevertheless greater than they are individually. This is also one of the points that is quantitatively backed up by Valentini and Sen, who argue that superdeterminism requires *more* fine-tuning than, for example, pilot wave theory (Sen & Valentini, 2020, p. 10). Finally, when looking from a purely quantitative lens, one may note that while there are a finite number of parameters in the standard models, the number of correlations between entities in the universe required to make superdeterminism work is practically infinite.

3. *Distinct fine-tuning*. The conspiracy argument is not just ‘more of the same’, introducing new unpredictable parameters that, if they had been somewhat different, would have left us a lifeless universe. The fine-tuning is here more so to reproduce the statistics of quantum theory, for which significant correlations between many remote and independent entities in the universe is required. Much has been said about this conspiratorial fine-tuning, but it is clear that the two cannot be equated. Both types of fine-tuning require very specific values to be compatible with empirical reality, but that is about where the similarities end. The fine-tuning problems arise in different ways and the tuning itself works towards different ends. Another distinction is that while the traditional parameter fine-tuning problem is model-independent, being present whenever the existence of physical parameters is acknowledged. Conspiratorial fine-tuning is not forced on us in this way, but rather a choice of theory. Moreover, as we have seen, this also affects how one can deal with the problem at hand. While the multiverse in combination with the weak anthropic principle was described to be a way to tackle parameter fine-tuning, this does not work for superdeterministic fine-tuning[72]. One may argue that to equate these fine-tuning problems and add

[72] Similarly, the argument from design as a response to the fine-tuning problem also becomes less convincing in the superdeterministic case. While, from a theological perspective, it is quite reasonable that a God or Gods would want to fine-tune their universe for life, it is less obvious why they want to make us believe the laws of quantum theory specifically. It does not help that even for the traditional

> them is a mistake, as the origin, magnitude and solution to them may be very much distinct. Perhaps a multiplication of problems, rather than mere addition, may then be a more fitting way of referring to the situation.

In conclusion, the conspiracy argument, and the fine-tuning it implies for superdeterminism, cannot be dissolved by pointing out that it is not the only problem with fine-tuning in physics.

In consideration of all of the above, I am inclined to view the conspiracy argument as a serious problem for superdeterminism. Ultimately, I do not think the counterarguments manage to rebut the problematic conspiratorial character of the framework. Nevertheless, two superdeterministic responses were acknowledged to add some asterisks to the argument, even though these could be commented on as well. Firstly, as 't Hooft points out, a type I model that would succeed in providing a convincing physical law acting locally on systems in the early universe, and thereby creating the correlations required by superdeterminism, could challenge the notion of it being fine-tuned. Secondly, one should be aware that the argument rests upon whether one is justified to apply probability theory. These conditions on which this application rests imply a number of non-trivial metaphysical assumptions that due to their 'extra baggage' serve to somewhat weaken the argument. In spite of this, I have attempted to evaluate the impact of this to conclude that, as I see it, the conspiracy argument still stands.

The science-invalidation argument

While the conspiracy argument is primarily taken to be metaphysical in nature, the science-invalidation argument takes an epistemological approach instead. Authors advancing this argument stress that the assumption of statistical independence is necessary for our ability to do science in the first place. Superdeterminists tend to respond that statistical dependence only holds in the quantum realm. This subchapter will therefore start by laying out the science-invalidation argument, after which the superdeterministic response will be considered. Finally, I will evaluate the debate as it stands, arguing that while superdeterminists can in some ways fairly defend the framework, this comes at a (too great) cost.

---

fine-tuning problem there are already not that many proposed categories of solutions, each of them highly controversial.

The argument: why the scientific method relies on statistical independence

That rejecting statistical independence would render science impossible was the first argument made against superdeterminism, as it was already brought to the forefront by Shimony, Horne and Clauser, who first explicated it as an assumption of Bell's theorem. They wrote on superdeterminism that:

"We cannot deny such a possibility. But we feel that it is wrong on methodological grounds to worry seriously about it if no specific causal linkage is proposed. In any scientific experiment in which two or more variables are supposed to be randomly selected, one can always conjecture that some factor in the overlap of the backward light cones has controlled the presumably random choices. But, we maintain, skepticism of this sort will essentially dismiss all results of scientific experimentation. Unless we proceed under the assumption that hidden conspiracies of this sort do not occur, we have abandoned in advance the whole enterprise of discovering the laws of nature by experimentation." (Shimony, Horne, & Clauser, 1976)

The critique above is clearly distinct from the conspiracy argument, even though it mentions the existence of conspiracies. It targets the epistemological consequences of superdeterminism, claiming that the inability to randomly select variables in experiments will invalidate experimental discovery. Much of science, after all, depends on this assumption (Dattani, 2022). To see why this is the case in a more concrete manner, we consider the hamburger experiment example for which the groundworks have been laid in chapter 6. This example is heavily inspired by the aforementioned philosopher of physics Tim Maudlin, who often uses a similar scenario to dismiss superdeterminism.

In chapter 6, a situation was sketched wherein a scientist wants to test whether eating many hamburgers a day leads to an increased probability of developing heart disease. This was then done in the following way. The scientist has a very large group of people that they intend to *randomly subdivide uniformly* into an experimental group of people who will eat 10 hamburgers a day, and a control group that will not. Many years after the experiment, the scientist checks whether a statistically significant difference in the frequency of heart disease can be found between the groups. Assuming a statistically significant difference is structurally found, we conclude that eating a large number of hamburgers a day causes greater risk of heart disease.

The legitimacy of the scientific research above hinges, among other things, on the uniformly random subdivision of the group into an experimental group and a control group. As an example, consider that people differ genetically in their likelihood of developing heart disease. Moreover, there are more lifestyle choices eating hamburgers that can contribute to developing heart disease, such as a lack of physical activity. If the group subdivision were *not* to be random with respect to these variables, there could be a situation where all or at least a higher number of the people with severe genetic predispositions towards heart disease or insufficient physical activity end up in the experimental group. If we then find a much higher frequency of heart disease in this group, how are we to know whether this is due to the hamburger eating, or just the genetics and physical inactivity? Our scientist ends up with potentially many interfering variables that obstruct any hope of establishing causal relationships between our variables of interest. Here, the interference between variables is defined as an effect where when an experiment is set up in which only one variable is changed to study its effect on the other, this inevitably changes other variables as well, making it impossible to know what variable(s) to attribute causal influence to and to what degree. This is effectively what happens in the example.

Finally, as was also explained in chapter 6, the above is just a more macroscopic example of the rejection of statistical independence. In the hamburger experiment, we desire the statistical independence of, for example, the degree of genetic predisposition to heart disease of a person and the group they are assigned to[73]. In the Mermin experiment we desire the statistical independence of, respectively, the hidden variables and the axis of orientation the spin is measured against.

Since most scientific experiments across disciplines involve the risk of interfering variables when subjects cannot be subdivided in groups statistically independently from properties of these subjects, rejecting statistical independence appears to leave us with a situation where we are no longer able to derive causal relationships from observed correlations. After all, statistics generally warns us to ideally only change one variable in every experiment so that if we find a correlation between them, we are able to conclude that this relationship is causal. Since finding causal relationships between variables of interest is one key interest of the scientific enterprise, the absence of statistical independence assumptions of this kind can then be concluded to invalidate all supposed knowledge of such matters acquired through the scientific

[73] If it feels 'strange' that nature could be such that these variables could not be statistically dependent, that is the result of seeing the conspiratorial nature of superdeterminism being played out in a more everyday situation. For now, let us accept that it cannot, and investigate the epistemological consequences of this.

method. Therefore, we must adopt it as a necessary principle of the scientific method, just as we do for inductive reasoning.

This is the science-invalidation argument, or at least my interpretation of how it can be worked out. Given the high stakes involved, superdeterminists have not ignored this argument. We will now consider their responses.

The response: statistical dependence as a verifiable quantum property

The science-invalidation argument states that statistical independence is a necessary condition for the acquisition of scientific knowledge. The superdeterminstic response consists of a two-step approach. First, they reconceptualize statistical independence as an empirically verifiable property of systems, rather than a necessary principle for doing empirical science in the first place. This move can be found both from Hossenfelder & Palmer as well as Andreoletti & Vervoort. The second step consists of the claim that while statistical independence is a useful property that works for classical systems, it does not always apply to quantum systems (Hossenfelder & Palmer, 2020, pp. 14-15) (Hossenfelder S. , 2020, pp. 8-9) (Andreoletti & Vervoort, 2022, pp. 15-18). I will start out by explaining where these steps come from, to then move on and critically evaluate these in the next part.

On the epistemological status of statistical independence, the authors claim the following:

"[The science-invalidation argument makes a mistake][74], which is the idea that we can infer from the observation that Statistical Independence is ***useful*** *to understand the properties of classical systems, that it must also hold for quantum systems."* (bold added) (Hossenfelder & Palmer, 2020, p. 14)

"First, it is clear that statistical independence is ubiquitous in nature, and an ***experimentally well-confirmed***[75] *assumption in countless experimental situations."* (bold added) (Andreoletti & Vervoort, 2022, p. 16).

Clearly, this is a view of statistical independence that dethrones the concept as a principle, prerequisite for distilling knowledge from empirical science. Instead, it is viewed as a property that holds for some systems but may not do so for others.

[74] The part in the square brackets is a paraphrased addition to make sense of the relevant part of the citation.
[75] The bold parts were added to accentuate what is commented on in the following paragraph.

Whether this is the case for any particular system is then itself subject to empirical study.

An example given by Andreoletti and Vervoort is as follows. First, they mathematically formulate statistical independence between a variable *x* and *y* as $P(x|y) = P(x)$. In chapter 3 and 6, such conditional probabilities for the case of Mermin's experiment have already been provided. As a brief reminder, the equation states that the probability distribution of a variable *x* does not change when another variable *y* is already given. Thus, they are independent. The authors now claim that this property can generally be tested to hold. For example, suppose I do an experiment where I ask many people to take a marble out of a sack at any time they feel like it. I then use this data to construct a function for the probability distribution *P* that someone takes a particular time *t* to pick a marble, i.e., *P(t)*. Now I change the experiment slightly, where in another room I have a person who can set a button to either *A* or *B*. The button variable is called *X* and can take on the aforementioned two values. The button does absolutely nothing. Yet now I can test if whether the button is turned to A or B comes with any change in the probability distribution *P(t)*. This is, of course, a rather ridiculous notion. Naturally, we shall find that $P(t) = P(t|X = A) = P(t|X = B)$. Therefore, the time variable *t* and the button variable *X* are uncorrelated, i.e., statistically independent. Thus, Andreoletti and Vervoort conclude, it is in principle always possible to test whether statistical independence holds between two particular variables or not. Bell tests are just one of these contexts where, perhaps against expectations, it does not. Whether it is unambiguously true that statistical (in)dependence can always be tested this way will be questioned later. First, the whole line of reasoning of superdeterminists regarding the science-invalidation argument will be laid out.

Hossenfelder and Palmer make a slightly different argument from Andreoletti and Vervoort, emphasizing that in classical models, assuming statistical independence generally grants your scientific model far more explanatory power. The latter is defined in the computational sense of "fitting a lot of data with few assumptions". This may seem like an odd move, as superdeterminism, relinquishing statistical independence, will not 'explain' (in the aforementioned sense) more Bell test data than standard quantum mechanics. On the contrary, it may need to include many specific parameters for the initial conditions to fit a lot of data. I suspect, however, that their claim originates in their suggestion that superdeterminism can, in other experimental contexts, (eventually) explain more data than its competitors. Concrete examples would be the measurement problem and even the allusion to a Theory of

Everything. Nevertheless, their view ultimately comes down to a pragmatic, post hoc justification of statistical independence. That is, it is justified through its data-fitting qualities in experimental contexts.

Both duos then go on to explain that statistical independence is thus a property of classical models in particular. It works in that domain. However, this is not the case in particular quantum systems. Hossenfelder notes that:

"In fact, it is extremely implausible that a quantum effect would persist for macroscopic objects because we know empirically that this is not the case for any other quantum-typical behavior that we have ever seen." (Hossenfelder S. , 2020, p. 9)

Remembering the correspondence principle, it is indeed true that there are many quantum behaviors that were seen to wash away at the macroscopic level. Hossenfelder and other superdeterminists ask us whether statistical dependence is not just another one of these types of phenomena.

This is no unreasonable strategy, but it does require the following. We have a very clear physical mechanism or reason, including an accompanying mathematical demonstration, for why certain quantum effects are not observed in the classical domain. Superdeterminists will therefore also need to provide this to explain that statistical dependence can occur on the quantum scale, but not on the classical scale. Whether they succeed in doing this is of fundamental importance to the superdeterministic project. One of the main cited reasons for its support in the first place, is its provision of an answer to the measurement problem of quantum mechanics. The measurement problem is precisely the sort of problem whereby one is confronted with a seemingly insurmountable difference between the classic world and the world of the quantum, referred to by Bell as a 'shifty split'. Superdeterminism can explain (away) wave function collapse, but the question is whether this is of much help if it trades one problem for a very similar other one. Therefore, to prevent the formation of a 'superdeterministic shifty split', a sensible physical mechanism and mathematical formalism to explain why we observe statistical dependence in Bell tests, but not seemingly anywhere else in science, must be provided. The authors are aware of this, and both have their answer. I will now lay out these answers, after which they will be evaluated.

Andreoletti and Vervoort argue that superdeterminists can demonstrate what a solution looks like using two "ingredients". The first of these, they explain, is that superdeterminism involves a “highly specific class of variables” that are part of a hypothetical Theory of Everything. Since these are special variables part of a theory describing the universe from the big bang onwards, all later variables (such as measuring angles) are correlated with them. This, then, would imply that all variables in the universe, including macroscopic non-fundamental ones, are correlated by their common cause. Here, the authors invoke their second ‘ingredient’, which is the idea that the correlations between such variables are so small that they are unmeasurable in practice, the measurement error being larger than the magnitude of the correlations themselves. Since it is also impossible to control for the variables of the fundamental theory as they appeared shortly after the big bang, these small correlations have no way of being detected by any experiment either. Finally, to explain why we do not see correlations between macroscopic variables such as the two (supposedly) independently set measuring angles of a Bell test, but we do see correlations between these measurement settings and the fundamental real state of a particle we measure, Andreoletti and Vervoort argue that the fundamental, as of yet hidden, variables we are concerned with have a “stronger tendency to remain detectably correlated” (Andreoletti & Vervoort, 2022, pp. 15-18).

The authors conclude that the correlations between, in principle, all variables in the universe, can only be observed when “at least one of the correlated variables belongs to a fundamental theory *and* if all variables can be actualized in experiments” (Andreoletti & Vervoort, 2022, p. 18). Here, one can imagine that in a Mermin-like experiment, it is the hidden variables giving away the state of an atom that belong to a fundamental theory, and we can observe the value of the atom’s spin as well as the value of the measuring angles used to measure it in an experiment. Both conditions are met.

While such a framework is not logically inconsistent, one may rightfully demand superdeterminists to provide convincing reasons for us to accept this. After all, their explanation involves a number of (arguably ad hoc) assumptions about nature, the results of which are mostly unmeasurable. Moreover, known quantum mechanical mechanisms suppressing quantum mechanical effects at the macroscopic scale are unavailable to superdeterminists. One such example is quantum decoherence. However, decoherence makes use of a quantum mechanical ontology that superdeterminists want to do away with in the first place. It is then not clear how it can be applied to statistical independence.

Hossenfelder likewise views statistical independence as a strictly quantum effect and disagrees with the above criticism. She claims Ciepielewski, Okon and Sudarsky ('COS') have already shown how violations of statistical independence can be "effectively erased" within a superdeterminstic model (Hossenfelder S. , 2020, p. 9). These physicists from the National Autonomous University of Mexico have created a superdeterministic model which, they claim, fares significantly better against some common critiques than may be expected (Ciepielewski, Okon, & Sudarsky, 2020, p. 19). For our current purposes, the claim that the 'superdeterministic shifty split' can be explained will be the question kept in mind when considering their model. Regardless of the model's perceived success in some key areas, the authors nevertheless reject superdeterminism as an option on other grounds that will be discussed below. To understand this, I will sketch the idea of their model insofar required for our purposes, in a qualitative manner extracted from its mathematics.

The authors call their model a "local pilot-wave model". It starts out from the mathematical and ontological picture of pilot wave theory. This picture is then modified in a way that makes the nonlocality associated with the framework disappear, trading it for statistical dependence. As was discussed in chapter 5, the pilot wave ontology consists of particles and a guiding wave. This made for a mutually dependent nonlocal system, as the guiding wave determines the motion of the particles, but the position of the particles *instantaneously* affects the guiding wave. In the model of COS it is proposed that instead, every point in space contains its own complete pilot wave system. In other words, rather than there being one guiding wave whose numerical value depends on time and the distribution of particles in space, each point in space can be said to be its own guiding wave with just as many degrees of freedom[76]: a guiding 'wave field'. The state of the system can then be specified by a guiding wave field and a position field, these two entities making up the ontology of the model under consideration[77]. If the initial conditions

[76] A degree of freedom is an independent parameter which, when all taken together, specify the state of a physical system. For example, a simple pendulum in a plane has only one degree of freedom: the angle between its rope and the normal (to the Earth's surface). A single point particle has three degrees of freedom in space: its *x*, *y* and *z* coordinate. Finally, the motion of two particles connected by a spring can be completely specified by six degrees of freedom: three translational motions, two rotational ones and one vibrational one. The local pilot-wave model described above then has at each point (3N+1) degrees of freedom. Here, N is the number of particles in the system, each having three degrees of freedom as shown by the second example. The final degree of freedom can be attributed to the time parameter.

[77] The ontological elements of a physical theory are sometimes referred to as 'beables'. This term was introduced by Bell, in order to be contrasted with the more popular term 'observable'.

of the model are now taken to be homogeneous[78], the model will not only evolve locally but will be behaviorally identical to nonlocal pilot wave theory. Homogeneity is a mathematical condition. It means that, for example, at the beginning of time, the guiding wave field and position field are the same at every spatial location. On the ontological picture of a guiding wave field and position field, COS state the following: “It is as if every point in space could be thought of as a supercomputer that is running a perfectly detailed simulation of the entire universe. As a result, every point in space ‘knows’ everything there is to know about the whole history of the universe” (Ciepielewski, Okon, & Sudarsky, 2020, p. 22). The authors compare their model to Leibniz’s monadology, where monads behave in a similar way according to a pre-established harmony. For this reason, Chen refers to the model of COS as “Leibnizian quantum mechanics (LQM)” (Chen, 2020, p. 17).

The model of COS is discussed here not only because Hossenfelder views it as an answer to the superdeterministic shifty split. It was also chosen because the authors argue that the model can be used to evaluate others as well, due to them all sharing some key features on which COS base their later critiques. Comparing their model specifically to ‘t Hooft’s cellular automaton interpretation, they claim that existing differences can be removed through simple modifications of their own, for example by switching out points in space with extended cells (Ciepielewski, Okon, & Sudarsky, 2020, pp. 30-31). They conclude:

“It seems clear, then, that our model can be seen as a full realization of ‘t Hooft’s ideas. Moreover, it appears that any model capable of doing so would have to share with our model the (extreme) feature of having at each cell full information of the whole system (or something very similar).” (Ciepielewski, Okon, & Sudarsky, 2020, p. 31).

This makes it so that conclusion we ultimately draw about these authors’ model, are not unique just to it, but can be used to say something about the superdeterministic framework as a whole, which is our primary concern.

The authors go on to claim that empirical equivalence between quantum mechanics and LQM is achieved with generic initial conditions, given the homogeneity

[78] Due to their homogeneity, the initial conditions are now quite easy to characterize. For the mathematically-minded, the field gradients can be demanded to be zero, $\nabla_{\boldsymbol{x}}\Phi_{\boldsymbol{x}} = 0$ and $\nabla_{\boldsymbol{x}}Z_{\boldsymbol{x}} = 0$, where Φ is the guiding wave field and Z the position field. Generic initial conditions are then sufficient to reproduce nonlocal pilot wave theory, thus specification of the initial conditions becomes surprisingly simple.

condition. Due to this, they claim the initial conditions, despite possibly being fine-tuned[79], are at least not at all overly complex, despite the complicated correlations resulting from them. That is, they can be written down. More importantly for our purposes is their analysis of the science-invalidation argument with their model. This can be interpreted as a concrete realization of the recipe provided by Andreoletti and Vervoort. The authors claim that under their model, “everything is correlated with everything” but that “these correlations manifest themselves at the at the mass density level, only when there is entanglement” (Ciepielewski, Okon, & Sudarsky, 2020, p. 29). The same, they explain, holds for the manifestation of nonlocal behavior in pilot wave theory. Quantum entanglement, however, plays no role in macroscopic experiments researching heart disease, because even though *in principle* entanglement is ubiquitous, it “washes out” at that scale.

In conclusion, superdeterminists have risen to the shifty split challenge and provided at least one model where they show that it is at least not inconsistent for statistical independence to hold in some contexts, while not in others. Nevertheless, models such as this are not without criticisms. In the next part, I will critically evaluate both the superdeterminist reconceptualization of statistical independence as an empirically verifiable property instead of a necessary epistemological principle, and their belief that, from there, one can provide a viable solution to the superdeterministic shifty split.

The evaluation: the (unaffordable) expense of saving the scientific method

The debate on the science-invalidation argument is prone to people talking past one another. To avoid any such confusion and work towards a reasonable evaluation of the debate, clarity with regards to ‘statistical (in)dependence’ in this context is of importance. The next two paragraphs aim to achieve just that.

First of all, it is important to remember that statistical (in)dependence always holds (or not) with respect to two variables. Common statements in the literature such as ‘there is no statistical independence’ or ‘the universe is statistically (in)dependent’

[79] In truth, the authors make a distinction between a few different possible models. In one of these, the condition of homogeneity is demanded to be a constraint internal to the theory, where consequently the fine-tuning problem is diminished because any of the relatively small set of remaining initial conditions to choose from will now do. However, the authors themselves admit this may be viewed as a bit of cop-out, as there is no “independent motivation” to do so (Ciepielewski, Okon, & Sudarsky, 2020, pp. 27-28). Therefore, while their model can be said to challenge the idea that the initial conditions that would need to hold for superdeterminism to be true are too complicated to even write down, I would not say this rebuts the conspiracy argument. This is true all the more given the price of enormous ontological expansion, which will be commented on in a moment.

are therefore, technically ill-defined. Usually, the interpretation of these statements is rather straightforward if one knows what is meant by them. But this is not always self-evident and can cause confusion between authors. The above statements can be taken to refer to our typical relevant variables in Bell tests, but can also be taken to say that every existing physical system is correlated with every other one. The conspiracy argument shows that the correlation between the hidden variables and the measurement setting mechanisms more or less implies the latter, but superdeterministic authors claim these are invisible for all practical purposes outside of experiments relating such a variable to fundamental hidden variables. On the other side of the fence, it should be noted that even critics of superdeterminism will not claim that there are no statistically dependent variables in the world. For example, variables with a causal relationship or correlation due to a common cause violate statistical independence, even though ubiquitously accepted relationships of this kind are everywhere. Thus, to all parties involved, it is a matter of what type of dependence-relationships in what contexts are legitimate and which ones are not. The key difference between the parties is that the statistical dependence required by superdeterminism, is the kind that invokes a class of correlations that supposedly renders knowledge from experimental science impossible because the systematic and unavoidable interference it introduces while one tries to establish causal relationships. The superdeterminists, however, respond by accentuating that there is no problem as it is only with respect to the two relevant variables in Bell tests that statistical independence relations do not hold.

This leads to a second point that can be a source of 'talking past each other'. It concerns the now familiar issue of whether one views statistical independence as an epistemological requirement to gain scientific knowledge, or whether one views statistical independence as a (meta)physical property that holds for only particular systems in nature. While authors like Maudlin seem to focus on this first aspect, superdeterminists like Hossenfelder seem to react in spirit of the second. I would, however, say that examples such as the hamburger experiment are not supposed to show that statistical independence is a property that holds everywhere all the time, but simply that in any experimental situation where it does not hold we cannot reliably derive scientific knowledge.

At the same time, these two ways of looking at the problem are not mutually exclusive. To summarize this and create oversight, consider the table below. The first column considers the (meta)physical question: *'is it possible for statistical* *independence of the (conspiratorial) kind under consideration*[80] *not to be present in some experimental contexts*[81]*, in a way that successfully restricts this behavior only to these contexts?'* Here, a successful restriction is defined as the provision of a *realistic* physical model and accompanying mathematical formalism consistent with the rest of physics. One sensible necessary condition for this (non-philosophical) realism is that it decreases with the quantity and sensibleness of the entities introduced. To give a clear example: if it is proposed that statistical independence is violated only in Bell tests because there exist demons who use magic to achieve this because they like to confuse physicists, this is not realistic. It may seem very far-fetched to introduce such obvious standards for an explanation, but some would argue that the ontological structures superdeterminists need to introduce to solve their shifty split problem, is not so far removed from the absurdity of these demons.

The first row then asks the epistemological question: *'is statistical independence necessary for deriving scientific knowledge of causal relationships from experimental contexts?'* The table describes the four possible interactions of answers to these two questions, and the positions they entail.

[80] That is, the kind that implies pre-existing correlations if rejected, and not already familiar causal relationships or common-cause variables that do not lead to classically unexplainable correlations.
[81] These are, of course, Bell tests.

Epistemological / Metaphysical	Yes	No
Yes	Statistical independence does not hold in Bell-settings and this behavior can be confined to them. However, the fact that this is the way the world is, is bad news for the pursuit of science. It means many facts about nature are simply inaccessible to us in principle.	This is where we find most superdeterminists. Only in certain experimental settings, statistical independence does not hold. Luckily, this does not prevent us to learn facts of nature concerning these systems. Superdeterminism opens exciting new possibilities for science.
No	We need statistical independence to hold in the area of contention in order to do science there, and luckily, it also does. Superdeterminism, namely, cannot in a consistent and realistic way invoke selective violations of statistical independence. Superdeterminism is wrong on both counts.	Superdeterminism is wrong since we have convincing reasons to reject the existence of selective and conspiratorial statistical dependence. However, had this been otherwise, we would still have been able to derive scientific facts from such experimental settings.

The position I will defend in the following, is a 'no' to the first question given the criterium for 'successfulness'. With regard to the second question, I will argue 'yes' in general, but 'no' when considering the order of events in superdeterminism. This may seem contradictory, but this will be argued for in the paragraph after the next one. Therefore, we first turn our attention to the epistemological question, and see if superdeterminists are justified in their characterization of statistical independence.

One possible reason against this is the seeming impossibility of empirically verifying statistical dependence as proposed by many superdeterminists. In the paraphrased marble sack and button example inspired by Andreoletti and Vervoort, where it is measured that $P(t) = P(t|X = A) = P(t|X = B)$, one insight is that this inference can technically only be made through the prior assumption that these variables are not conspiratorially correlated to other physical variables and systems. Otherwise, the non-correlation of the marble picking time and the button could, conceivably, be due to common cause correlations with the big bang, misguiding us in believing the aforementioned equations. Thus, "verifying statistical independence" in any experimental context, as set out by the authors (Andreoletti & Vervoort, 2022, p. 16), may itself necessitate the prior assumption thereof for all other contexts. The embeddedness of any experiment attempting to demonstrate this in a universe already having set the precedent to be full of conspiratorial correlations (as demonstrated in the conspiracy argument), may rob any such effort of certainty. Moreover, the results of Bell tests, the experimental data that superdeterminists 'explain' with statistical dependence, can also be 'explained' by nonlocality or by rejecting hidden-variable theories. Thus, what superdeterminists take to be an empirical demonstration of statistical dependence, is in reality only so when embedded in a theoretical framework that is itself not forced upon us by the data. Therefore, while Andreoletti and Vervoort claim that it is in general possible to test statistical dependence, the only concrete example they have of this is one that will, by most, not be accepted as a real test thereof. The absence of solid proof for the belief that statistical (in)dependence is empirically verifiable, may be taken as supportive for the position that it, like induction, should be viewed as an epistemological principle of doing science. However, even if we accept that statistical independence cannot be proven in this way, that does not necessarily mean that rejecting it in the specific context of Bell tests invalidates all of science.

On this question of whether statistical independence is necessary for the pursuit of science I would answer 'yes' in general and 'no' when confined to the specific case where it is invoked by superdeterminists. As this paragraph will attempt to explain, this is the case because the science-invalidation argument, invoking hamburger-like experimental settings, is not actually an accurate comparison to Bell-like ones. In general, rejecting statistical independence between, broadly, 'settings' and 'states' leads to interfering variables of the type that render knowledge of causal relationships impossible, such as with the hamburger vs. genetic predisposition example. However, in the case of Bell tests there is no interference in this way if we let statistical independence go. This is simply due to the existence and knowledge of Bell's inequality, which constrains the statistically allowed outcome given local hidden variables. To have the statistics conform with measurements that lie outside this constraint, statistical independence is dropped, rather than not having it before even doing the measurements as in the hamburger experiment. This dropping quantitatively affects theoretical predictions in superdeterministic models, i.e., it fits them to the statistical data. This change is, therefore, *unambiguously* attributable to the pre-existing correlations, rather than to the hidden states. No hamburger vs. genetic predisposition interference applies. After all, a superdeterminist introduces these pre-existing correlations precisely because the known measurement statistics of hidden states will not agree with what is experimentally found. In other words, there is no room for confusion as to whether it is the distribution of hidden variables or the correlations between it and the measurement settings employed that lead to the observed statistics. Rather than being a given complication muddying the waters, statistical dependence is introduced for this sake, and thus is not an interfering influence in this context. It would be as if we could already derive a constraint on the relation between heart disease frequency and hamburger eating prior to the experiment. If the experiment then turns out to violate this constraint, we reject statistical independence to render the constraint illegitimate. In conclusion, while statistical independence is necessary to do science in general, the hamburger-like macroscopic comparisons to Bell-type tests are not accurate. They assume that we start without statistical independence, and subsequently get data that is 'interfered with'. But the reality of the situation is the other way around. We first get the data, and only then drop the statistical independence assumption in response to it. Its rejection then need not inhibit the acquisition of data or the attribution of quantifiable causal roles *when confined to the specific experimental context in which it is dropped*. While superdeterminists do not spell this out explicitly, their answer is effectively equivalent as they go on to attempt to justify this required confinement. If they do not succeed in doing so, and other experimental settings *start* without being able to assume statistical independence, we are left with the general case where science is, indeed, invalidated.

But even in the case where this does succeed, yet we do not believe this allows superdeterminists to derive any scientific knowledge from Bell-type tests, they could still have replied that the fact that humanity does not have epistemic access to all facts of the universe according to a theory, does not guarantee the falsehood of that theory. In fact, the Copenhagen interpretation of quantum mechanics itself greatly limited what can be known about certain physical systems, but this is generally accepted nowadays too.

Thus, while the criticism that statistical independence is an epistemological principle necessary for doing science is correct in general, superdeterminists implicitly know they can get away with it if they can restrict the impact thereof to the Bell tests where dependence is invoked. In other words, they must solve their shifty split problem. This then leads to the second issue of whether it is (meta)physically possible for a framework to successfully only allow for statistical dependence in the relevant sense in particular quantum experiments. For this, we return to the model of Ciepielewski, Okon and Sudarsky (COS), and Chen's reply to it.

Neither COS nor Chen are superdeterminists, but both their papers aim to evaluate the theory. To do this in a more concrete manner, COS have developed the model, similar to 't Hooft's, that was earlier referred to as 'Leibnizian Quantum Mechanics' (LQM). This name was concocted by Chen in response to the similarity of the model with Leibniz's monadology, already commented on by COS. After all, the model is such that in every point in space contains itself a simulated universe. This 'infinity of universes' is fundamental to LQM and, as we saw was argued by COS, to all superdeterministic models. The heavy ontological load this entails is a main point of critique of the model from both COS and Chen.

Nevertheless, both attribute some merits to LQM. Firstly, it shows that the initial conditions for a superdeterministic model need not necessarily be extremely complicated to write down. This is, however, only true given these are demanded to be homogeneous, which in the words of the authors is "imposed in a completely ad hoc manner, with no independent motivation" (Ciepielewski, Okon, & Sudarsky, 2020, p. 28). Moreover, as Chen points out, the model moves the complexity from the initial conditions to the ontology (Chen, 2020, p. 18). But secondly, and most importantly for our purposes, the model succeeds in providing a consistent account whereby statistical independence is violated in the desired way in Bell-type settings, while otherwise not (Ciepielewski, Okon, & Sudarsky, 2020, pp. 28-30).

This, then, shows that superdeterminism may indeed be able to defend itself against the science-invalidation argument. Moreover, LQM is an example of a model that shows this confinement is realizable in a consistent manner. However, superdeterminists appear to only be able to avoid the science-invalidation argument in this way at a great cost that may be argued to undermine the superdeterministic program anyway. This challenges the 'realism criterium' as defined above the earlier table. These 'costs' will now be considered and evaluated on their impact. I will cover three categories of them, largely inspired by the analysis of COS and Chen on the matter, but in addition adding my own arguments and commentary.

Firstly, it is argued by COS that LQM implies the existence of an absolute frame of reference (Ciepielewski, Okon, & Sudarsky, 2020, p. 31). This means that one can define a privileged coordinate system that is always at rest, and all objects in the universe can have their position and momentum defined relative to it. However, giving up on notions of absolute space and the existence of such a privileged frame of reference is the cornerstone of special relativity, and thereby quantum field theory. If some opt for superdeterminism over pilot wave theory in defense of locality, often with reference to its importance in modern physics, it would be a problem if their alternative were to likewise violate such as important principle. The argument of COS follows from the fact that at each point in space, the guiding wave field and position field, which are the fundamental degrees of freedom of LQM, are completely autonomous. The latter refers to the fact that no point in space interacts with any other, but already itself simulates the entire universe. The idea is now that we require an absolute rest frame to be able to refer to a particular point $\boldsymbol{x}$ (for 't Hooft: a particular cell) in space throughout time. Without an absolute rest frame and the privileged position it carries with it, there is no way to get observers to agree on the evolution of the degrees of freedom associated with a specific point. In relativity spatial distance itself is, in fact, relative to one's frame of reference. Consequently, the relativistic spacetime of modern physics does not respect the autonomy of individual points in space, and is therefore incompatible with LQM. When homogeneous initial conditions are enforced, the theory becomes empirically indistinguishable from pilot wave theory and we have no access to this privileged frame, but it exists nevertheless. Thus, while due to the initial conditions this incompatibility is not empirically accessible, the fundamental metaphysical assumptions about the nature of spacetime in relativity and LQM strongly contradict one another.

Next, some metaphysical and epistemological problems present themselves. The first one consists in the idea that any observer in LQM would not be able to acquire information outside of the points in physical space this observer exists in. This is argued by COS because all information for the time evolution of the degrees of freedom at a point $\boldsymbol{x}$, is completely contained in $\boldsymbol{x}$. A human being, 'occupying' a collection of points in space, would then not be able to gain any information outside of themselves. This would imply a "profoundly solipsistic scenario" according to the authors (Ciepielewski, Okon, & Sudarsky, 2020, p. 32). I would agree that LQM bears out a solipsistic picture in the sense of one not influencing or being influenced by anything external to the entity, and that this is a serious objection insofar solipsism in this sense is perceived as a no-go. I would, however, add another potential objection to this, which may suitably be called ‘the combination problem of LQM’. If every single point in space contains all information for the evolution of the state, and conscious entities such as humans occupy a collection of these points, then how does the collection of autonomous points give rise to a single, unified stream of conscious experience? In other words, the metaphysics of LQM may act to worsen existing problems in the philosophy of mind. By taking away the possibility of interaction, problems such as these could become even more difficult to coherently solve.

Two more objections in this category are presented by Chen. For the first one, he asks us to imagine each point in space as a 3-dimensional ‘internal space’ with N particles whose dynamics follow that of pilot wave theory. This should not be taken to mean that this internal space exists in our dimension, but the points of our real, physical space can be visualized as such in accordance with the ontology and the dynamics of LQM. Thus, there are an infinite number of copies of the world, next to the real, physical world. Now he imagines that one lives in a LQM universe. Because the theory contains only one real, physical space, but an infinite[82] number of internal spaces, it is far more likely to find oneself in an internal space than in the real space. However, the *internal* spaces in LQM do not respect locality. This, I would say, then creates another de facto fine-tuning problem where we should be overwhelmingly likely to find ourselves in a nonlocal world. While I think this is a clever argument from Chen, I am not convinced that it is necessary to ascribe existence to entities in internal spaces that seems to be required for the argument to work. This might be another debate about the model and its implications. Continuing, Chen has a second argument in which he asks us what the fundamental space in LQM is. Since the internal spaces are ‘smaller’ in a sense, they could be viewed as the most fundamental spaces, rather than our real, physical space. This would then imply that the dynamics are fundamentally still nonlocal, as that is what they are in the internal spaces. This is an interesting observation, but I am inclined to believe that most physicists and philosophers are primarily interested in whether our physical world does or does not obey locality. Moreover, what is viewed as the most ‘fundamental

[82] Or at least practically infinite. In fact, this reveals, I would say, another issue for LQM. They must be very precise in their mathematical conception of a ‘point’, since these are generally not extended, yet they are to make up space in the theory.

space' may be somewhat arbitrary, and the argument could conceivably go either way. Therefore, I do not think this latter argument is sufficiently strong against LQM (Chen, 2020, p. 17).

Finally, and most prominently to both authors, LQM results in an immense ontological expansion. It has been stated before that the model is such that every point in space becomes a sort of 'supercomputer', simulating the entire world. Unlike the Everett interpretation of quantum mechanics, these worlds do not come into existence due to a specific process, but they are of fundamental status. Extreme ontological enlargement can itself be viewed as problematic. For Chen, it comes down to the observation that the complexity formerly present in the laws and initial conditions of superdeterminism, is in the LQM framework simply moved over to its ontology. Moreover, all natural phenomena can be explained away by postulating many unobservable entities that apparently cause them, but this generally not seen as a legitimate move in science. This goes beyond just Occam's razor. Even before Newton conceived of his universal law of gravitation to compare it to, most natural philosophers at the time would probably not have accepted the theory that all phenomena we now relate to gravitation were actually due to invisible dwarfs lifting and pushing objects. That is not to say that new successful physical theories cannot be accompanied with new ontology, such as the 'invisible' electric and magnetic fields of classical electromagnetism and, possibly, weakly interacting massive particles as dark matter. However, the new ontologies in these theories are justified by typical epistemic virtues of scientific theories, such as empirical adequacy, simplicity, explanatory power, etc. Yet, the authors claim that LQM seems unable to present such virtues through which it can justify its ontological expansion. The locality LQM has over pilot wave theory and that grants it compatibility with relativity, may be self-defeating if it needs to assume an absolute rest frame in place of that. In addition, the model itself and its metaphysical and epistemological objections show that when it comes to clarity, simplicity and explanatory power it is also hard-pressed to be judged as superior to a contender such as pilot wave theory. Both Chen and COS therefore go on to claim that LQM is disqualified based on an unjustifiable enlargement of the ontology (Chen, 2020, p. 18) (Ciepielewski, Okon, & Sudarsky, 2020, pp. 33-34). I think both authors are correct in their argumentation, since the 'parallelist quantum mechanics' of this model indeed introduces a radically unintuitive and ontologically bloated picture of the universe with, crucially, little to show for it in the way of epistemic virtues.

We now have an overview of the cost of defending superdeterminism against the science-invalidation argument. This argument showed that the structural interference effect that comes with the rejection of statistical independence could potentially destroy the pursuit of science. Superdeterminists deny this because they do not see statistical independence as a prerequisite principle for science and argue that it can be consistently modelled to hold for Bell tests while not in other experimental settings in nature. While I have argued that I do not think statistical dependence destroys science *given* it is realistically confined to Bell settings in a consistent manner, it was shown that those willing to walk this route and to attempt to build models rising to this challenge must pay a high price. This cost lies in an undermining of the foundations of relativity, some considered metaphysical and epistemological problems, and an enormous expansion of the ontology of the natural world. Therefore, if one accepts these criticisms, the LQM answer to the superdeterministic shifty split is neither consistent (e.g., due to clashing with core relativistic tenets), nor realistic in the sense defined for the earlier table (e.g., due to the ontological expansion). Superdeterminists are then left with a dilemma. Either they reject statistical independence entirely, which invalidates science, or they claim it is only for Bell-like tests that it does not hold, which was just argued to be unconvincing. The model of COS is relatively general in its approach, and any assessment regarding full information of the universe being present at each point was argued by the authors to carry over to other superdeterministic models. Therefore, it seems unlikely that any superdeterministic model other than LQM would be up to the task. In conclusion, I think the science-invalidation argument, like the conspiracy argument, is a significant challenge to superdeterminism.

I will not argue this to be the case for the next argument, one that is commonly invoked in debates about superdeterminism. This is the argument about free will, and it will be the subject of the next section.

The free will argument

As we have seen before, Bell already levied the conspiracy argument against superdeterminism in his paper on the topic. However, in a 1985 BBC radio interview, he had also made some comments on the topic. The transcript, as provided by Davies and Brown, reads as follows:

"There is a way to escape the inference of superluminal speeds and spooky action at a distance. But it involves absolute determinism in the universe, the complete absence of free will. Suppose the world is super-deterministic, with not just

inanimate nature running on behind-the-scenes clockwork, but with our behavior, including our belief that we are free to choose to do one experiment rather than another, absolutely predetermined, including the 'decision' by the experimenter to carry out one set of measurements rather than another, the difficulty disappears. There is no need for a faster-than-light signal to tell particle A what measurement has been carried out on particle B, because the universe, including particle A, already 'knows' what that measurement, and its outcome, will be." (Davies & Brown, 1993)

As one may note, Bell's emphasis here seemingly lies more so on superdeterminism doing away with free will than the conspiracy argument. Or at least, it can reasonably be interpreted that way. Consequently, superdeterminism has often been attacked on the grounds that it is a framework that is incompatible with free will. In reality, this was not quite what Bell intended, as even Hossenfelder and Palmer admit (Hossenfelder & Palmer, 2020, p. 12). Maudlin attributes this to Bell's strength of clarity lying more so with his writing than his speaking (Maudlin, 2022, p. 0:52:15).

Nevertheless, the association has stuck around[83], despite both proponents and the major voices on the opposing side claiming that the free will debate is mostly irrelevant to superdeterminism (Hossenfelder & Palmer, 2020, p. 12) ('t Hooft, 2017) (Maudlin, 2022, p. 0:53:18) (Sen & Valentini, 2020, p. 4). I think all of these authors are correct in this conclusion, but I will attempt to provide a much broader and more specific line of argumentation as to why I consider this to be the case. The persistence of the free will objection can be attributed either to the idea that free will must necessarily exist for experimentation to make sense or to a strong metaphysical belief in its existence. I will now attempt to show that the former group's argument is equivalent to the science-invalidation, and that the latter group has nothing to fear from superdeterminism, or at least not from superdeterminism *specifically*.

On whether free will is a necessity for doing science

Starting with the group taking free will to be necessary for science, one could possibly interpret the following quote by Zeilinger to be representative of their way of thinking:

"We always implicitly assume the freedom of the experimentalist... This fundamental assumption is essential to doing science. If this were not true, then, I suggest, it

[83] For example, when searching for 'superdeterminism' on web forums, the complaint that 'we would not have free will' is often the first to come up.

would make no sense at all to ask nature questions in an experiment, since then nature could determine what our questions are, and that could guide our questions such that we arrive at a false picture of nature." (Zeilinger, 2010, p. 266)

While the lack of "freedom of the experimentalist" in superdeterminism is advanced as an argument against superdeterminism in this quote, upon careful inspection one finds this particular critique to be equivalent to the science-invalidation argument. When Zeilinger argues superdeterminism lacks this experimental freedom, he cannot just be referencing the fact that it is deterministic. If determinism would take away the sort of freedom he has in mind, then his critique would apply to all of classical physics rather than specifically superdeterminism. The trouble then supposedly comes with the addition of statistical dependence. When considering the universe as a whole as a physical system, determinism implies that (given any initial conditions) the system evolves through a single grand causal chain. Superdeterminism works in the same way, except that in addition, it effectively preselects a very small subset of initial conditions that encode the correlations that cause systematic violations of Bell's inequality. Thus, it is not that humans are (un)freer to make choices outside this causal chain. It is, however, true that the initial conditions in these superdeterministic universes result in an inability to do statistically independent experiments. This holds not just for human choice with regard to detector settings, but also to Bell-like tests using π-digits or happening spontaneously in nature without any human involvement at all. And it is this more general inability, not just 'unfree' human choices, that muddies the waters for "doing science". After all, it is due to the general inability that the science-invalidation argument applies in a way that we can be 'led' to a "false picture of nature". Zeilinger's demand then, I would say, is not that free will must necessarily exist for experimentation to make sense, but that statistical independence should, with free will just being one of multiple means that we would 'normally' consider to guarantee us the ability to do statistically independent research. Therefore, the lack of free will in isolation is not the problem here.

On whether a metaphysical belief in free will collides with superdeterminism

Having established the foregoing, we will now move on to the group who believe in the existence of free will as a metaphysical principle and see the supposed incompatibility of this with superdeterminism as an argument against the framework as such. Valentini and Sen list it as one of the most prevalent arguments against superdeterminism (Sen & Valentini, Superdeterministic hidden-variables models I:

nonequilibrium and signalling, 2020), but do not consider it to hold much ground. One example that could be interpreted as an expression of this conviction comes from the mathematician John Conway (1937-2020), one of the originators of the free will theorem. Quoted through ‘t Hooft, Conway states that “*We have to believe in free will to do anything; I Believe I am free to drink this cup of coffee, or throw it across the room. I believe I am free in choosing to have this conversation*.” ('t Hooft, The Free-Will Postulate in Quantum Mechanics, 2007, p. 4) While it might seem quite straightforward, I will first state the reason why superdeterminism and free will are viewed as incompatible as explicitly as possible, so as to steelman the conclusion.

As has been stated many times, superdeterminism requires a pre-existing correlation between the measurement settings and the state of the object being measured. While physical randomizers such as random number generators, or events in the past such as distant quasars seem intuitively unlikely to be statistically dependent on the state of atomic spins we can measure, another setting mechanism for which this intuition holds is that of freely willing agents choosing measurement settings. Bell tests where settings were chosen in this way have actually been performed, but Bell’s inequality was still violated (Collaboration, 2018). Nevertheless, on a very intuitive level, we feel that we are free to always pick any measurement setting on, for example, Mermin’s device that we want. Yet, according to superdeterminism, these pickings must be correlated with the object being measured. So, unless states instantaneously change every time I change the detector setting (pilot wave theory), given a situation where a hypothetical atom is already underway, the correlations exist before these acts are being performed. This, then, restricts the freedom an agent has to truly pick any measurement setting they want. The deep correlations present in nature necessitate that we will choose settings that will give us the ‘correct’ statistics that end up violating Bell’s inequality, and we could not have picked other settings in a Bell experiment than the ones we in fact did. Our actions thus only appear free, but in reality, the setting I choose in a Mermin-like experiment is already determined.

Does the above mean that there is no free will? This depends on the answer to a question that is central to the free will debate in the first place, namely, how free will is to be defined. Incompatibilists hold that we cannot have both determinism and free will, a conclusion they usually arrive at by conceptualizing free will as the ability to have acted otherwise than you in fact did. This notion of free will contradicts the central thesis of determinism[84]: they are incompatible. Formally this is laid out by the 'consequence argument', but a short version is as follows. Determinism states that, given that the state of any physical system is known at some time, the time evolution of the system can be uniquely determined at all past and future times. Humans are, ultimately, also physical systems. Therefore, given the initial state, the actions performed by a human being are also uniquely determined for all past and future times. In conclusion, it is not possible for a human to have acted otherwise than they actually did at that time. This position is usually referred to as 'hard determinism' (Slors, de Bruin, & Strijbos, 2015, pp. 181-182).

The conclusion uses three premises: determinism, the incompatibilist notion of free will and physicalism. Physicalism comes into play through the premise that human decision-making can also be viewed as a physical system. Consequently, if one wants to avoid the argument's conclusion, one (or more) of these premises must be rejected.

Traditionally, one option is to reject physicalism with respect to domains such as the human mind. I say traditionally, because as controversial as this strategy already is, I would argue that it becomes logically impossible once we realize the effects of accepting statistical dependence beyond just determinism, i.e., superdeterminism. The rejection of physicalism is meant to put human action outside the causal deterministic chains of physics. Human action is then not determined by prior states of the universe, but able to introduce new causes into the natural world, themselves not determined by it. These causes can set off new causal chains in the natural world, functioning, in a sense, like first causes whose origin would otherwise be laid only at the big bang. The assumption that mankind is not subject to the same physical substance and rules, however, often implies a form of dualism. This comes along with the familiar baggage of the interaction problem, unfalsifiability critiques, and others. This is the origin of the controversiality of this strategy. But for the superdeterminist, the situation is even worse. Superdeterminism relies on correlations that have existed for a long time on a cosmological time scale. For type I models, it has been shown that these correlations must at least have been created over half the age of the universe ago, and likely go even further back. For type II models, they have been there since the big bang in the form of its initial conditions. One of these is required to get superdeterminism to reproduce the results of Bell tests. Now consider a non-physical being with free will in the sense that it could

[84] Note: just determinism. For this contradiction, the 'super-' part is not required!

indeed have done otherwise for every act it has performed, its actions being undetermined by nature[85]. The being starts performing many Mermin-like experiments and to its shock finds that at least one-third of the time, the detectors flash a different light. One may recall, after all, that this is the result of Bell's inequality for hidden-variable theories where measurement axes *can* be chosen independently of the spin state of the atom measured. The being can do this because its actions do not originate from causal chains in the natural world, and it is able to introduce new and undetermined causes into the physical world[86]. It is not bound by superdeterministic correlations, and it can choose measurement settings truly freely, without being unable to 'choose' other settings than the ones so correlated with the atoms to result in violations of Bell's inequality. Its influence has therefore broken correlations required for superdeterminism to work, and these cannot in any way be re-established either due to (from a type I perspective) local influences no longer being possible. The 'deactivation' of the correlations is a no-go unique to superdeterminism since it cannot explain empirical reality otherwise, which leads to the conclusion, that unlike for regular determinism, superdeterminists cannot reject physicalism to show its detractors that free will is safe.

Another option of the three logically available ones, then, is to reject determinism. Incompatibilists who take this position are called 'libertarians' (Slors, de Bruin, & Strijbos, 2015, pp. 183-185). But, if we are of the opinion that the free will sceptics have nothing to fear from superdeterminism, which is by definition deterministic, this can trivially be seen not to be an option. It may nevertheless be noted that the rejection of determinism to achieve free will is a dubious move at best that naturally leads us to question the incompatibilist notion of free will. If we take fundamental quantum probabilism as an alternative to determinism, it is indeed true that there exist at least some physical systems that could have done otherwise[87]. However, few would state that if all my choices were, as it were, *compelled* by the flipping of a

[85] Consider this a version of Laplace's demon particularly relevant to superdeterminism.
[86] This must be presupposed by those who aim to defend free will by rejecting physicalism, as they could otherwise not explain the evident fact that human actions (which are taken to be 'free') can ever affect the physical world.
[87] One should keep in mind that the probabilistic behavior at the quantum level cannot trivially be transposed to the macroscopic world of human decision-making. This is the central thesis of 'adequate determinism', the position that due to the correspondence principle and quantum decoherence, the world still functions deterministically macroscopically, despite probabilism at the quantum level (Adequate (or Statistical) Determinism, n.d.).

coin[88], I would be acting out of free will. Fundamental, uncontrollable randomness deciding ‘what will happen’ just does not seem to agree with the image of what we *mean* or *want* when we refer to an act out of free will in the first place.

Thus, a majority of philosophers reject the incompatibilist ‘could have done otherwise’ notion of free will (Bourget & Chalmers, 2014, p. 15). This final option of the three is what can still enable superdeterminists to hang on to free will. The strategy here is to critically examine how free will can best be defined. It is then argued that there are better ways of doing so than the ‘could have done otherwise’ version of it. These new ways of understanding the concept then need not contradict determinism. Hence, this approach often leads to ‘compatibilism’: the position that determinism and free will can coexist. The compatibilist notion of free will stresses that agential control is what is at the heart of free will. While they do not all carry precisely equivalent definitions beyond this, one such conception that can be shown to allow for coexistence with (super)determinism is that of philosopher A.J. Ayer (1910-1989). In his essay “Freedom and Necessity” Ayer argues that the antipode of freedom is not causality, but rather, constraint. Making a choice by accident is not what we consider to be a freely willed act, thus it would be strange to contrast freedom to causation in general. It is then not, for Ayer, that we should demand the ability to have done otherwise, but rather, that a) we should have been able to do otherwise *if we had chosen to do otherwise*. After all, if I after long deliberation choose to eat Italian food over Indian food, it would be strange to say that I could only have *freely* chosen to make pasta if, after my choice for Italian, I might as well have prepared red curry. However, the conditional analysis, in which I make the choice for Italian freely *given that* had I chosen to eat Indian food I would have been able to prepare curry, is argued to make much more sense, and all we should really want. This analysis, moreover, does not contradict determinism. After all, it is perfectly reasonable within a deterministic framework to assume that had I made another choice, this would have impacted the causal chains going forward. Whether I have the ability to have made another choice than I did itself then does not matter. After this modification, Ayer moves on to make his positive case for contrasting freedom with constraint. He argues that acts out of free will should not result from b) being coerced to choose as we did by external factors and c) inner compulsions to our behavior such as kleptomania, Tourette syndrome or the trembling of the hand of a Parkinson’s patient. Here, the emphasis on agential control is seen to come to the

[88] Where, to avoid misinterpretation of the metaphor, it is also *not* the case that the agent has any *control* over when to flip that coin.

forefront. Ayer concludes that the combination of the above three conditions a), b) and c) allows for an act out of free will, and that these are perfectly compatible with our actions being causally explainable in terms of determinism (Ayer, 1972). This reasoning naturally extends to superdeterminism, as the additional assumption of statistical dependence as an experimental property has no further bearing on Ayer's three factors.

In conclusion, if one accepts that the agential compatibilist definition of free will is at the heart of the concept, rather than the 'could have done otherwise' incompatibilist definition, then superdeterminism does not contradict free will. Most philosophers appear to agree on this semantic matter (Bourget & Chalmers, 2014, p. 15). If one is inclined to agree with them, the argument on the basis of free will is left nullified.

Two last comments must be made on this issue. Firstly, while I have attempted to evaluate the argument of free will being a necessary metaphysical demand on its own merits, it should be recognized that it is essentially an argument against all of classical physics and general relativity. The one exception was for the dualist position, that was explained to be incompatible with superdeterminism while not necessarily being so with determinism. Nevertheless, few side with this camp these days, and it follows then that all the issues that have been raised based on metaphysical belief in free will are already at play, and have been widely discussed, with regard to the classical determinism we have known for ages. Secondly, a superdeterminist disagreeing with compatibilist notions of free will and confronted with the free will critique can also choose to just bite the bullet and deny free will. The goal of the above was to show that if one values free will, superdeterminism need not pose a threat, at least not uniquely. It was not to argue that if superdeterminism would somehow not leave room for incompatibilist free will, this would make it an unviable framework. For people with a metaphysical commitment to this version of free will, superdeterminism not accommodating it might be viewed as a weakness of the theory and be counted as metaphysical baggage in its disfavor, like the infinity of worlds in the Everett interpretation. But that does not constitute a logical rebuttal of superdeterminism (or classical physics and relativity) in any way. In fact, many people have found this conclusion entirely satisfactory, Einstein being one example among them (Einstein A. , The World As I See It, 2010).

In conclusion, I do not think arguments invoking free will pose a challenge to superdeterminism. To the compatibilist, they are wrong. To the hard determinist, they do not matter. To the libertarian, these arguments are not unique to superdeterminism, as the determinism of classical physics and relativity are already incompatible with free will. Moreover, given the arguably unsatisfying alternative of randomness in this regard, one may wonder whether the libertarian point of view holds up in the first place. This then concludes the three considered arguments against superdeterminism. Two less discussed angles will, for the sake of completeness, very briefly be introduced now.

The nomological argument

The nomological argument concerns the effect superdeterminism has on quantum mechanical laws of nature. It is advanced by Augustin Baas and Baptiste Le Bihan, two philosophers of physics from the university of Geneva. They argue that from the perspective of superdeterminism, quantum mechanical laws are contingent and ontologically dependent on the initial conditions of the universe. Therefore, it is compatible only with a neo-Humean account of the laws of nature (Baas & Baptiste, 2020, pp. 11-14).

The above logically follows from superdeterminism. According to it, what we see as quantum mechanical lawful behavior is displayed only because the initial conditions were the way they were. Without the correlations that exist due to them, quantum mechanical predictions in Bell tests would be structurally violated as per Bell's inequality. Therefore, if we count the laws of quantum mechanics among the laws of nature, it is no longer possible to account for all of them in the necessitarian framework where they are viewed as necessary and ontologically independent from initial conditions. The neo-Humean picture, which rejects these views and instead emphasizes inductive observations of regularities, then is all that remains.

However, I think superdeterminism can be safeguarded against this argument. Firstly, the authors speak of dependence on the universe's initial conditions, so the argument is specifically directed at type II superdeterministic models. More importantly, though, superdeterminists do not claim that quantum mechanics is a complete and final theory. In the framework, it turns out that we were just wrong about quantum mechanical laws being laws, and this need not have further bearing on all other physical laws, including ones pertaining to the hidden variables of the theory. Quantum mechanics, after all, can in the superdeterministic framework be reduced to a hidden-variable theory that provides a more fundamental description of nature and its laws. These laws need not be viewed as neo-Humean per se, as it is only the apparent 'laws' of quantum mechanics that explicitly arise from systematic correlations between the probability distribution of hidden variables and measurement settings, not the hidden variables themselves. What we refer to as the 'laws of quantum mechanics' are indeed not laws, but an, as the conspiracy argument lays out, admittedly peculiar result from the initial conditions of our universe. Yet, since this has no bearing on the other laws of nature, a necessitarian conception of those is consistent with this state of affairs. It is therefore an interesting argument, but its conclusions are seemingly easily avoided by superdeterminists even if one does not believe in a neo-Humean account of the laws of nature.

The inexplicability argument

The inexplicability argument focuses on the effect superdeterminism has on the notion of a good scientific explanation. Consider a type II superdeterministic model. When doing a Bell test, what explanation can be given for Bell's equality being violated? The reason for this is result is the systematic correlations superdeterminism postulates to exist between the measurement settings and the hidden states of the measured atoms. But this correlation is entirely the result of how the initial conditions of the universe happened to be. Therefore, explanations of quantum mechanical measurement results such as these seemingly come down to 'because the initial conditions of the universe are the way they are'. Yet one would be hard-pressed to call this a 'scientific explanation' compared to the standards we usually hold these to (Ciepielewski, Okon, & Sudarsky, 2020, pp. 17-18). Thus, in a superdeterministic world, how are we to think about important concepts in the philosophy of science such as a scientific explanation, understanding in science and the intelligibility of the theory (De Regt, Understanding Scientific Understanding, 2017)?

The above argument is not of the metaphysical kind, nor the epistemological, when the latter is defined in the narrow sense of how much raw data we can access. It is, however, crucial to the philosophy of science and to how we understand the natural world. Therefore, as interesting and important as these issues are, they do not directly concern whether superdeterminism is viable as a framework in light of the truth-pursuing criteria that were defined for this thesis. For this reason, the issue will not be explored extensively here. There is nevertheless much to explore here for further research, such as the distinctions between models of superdeterminism (what about a type I model?), whether a similar 'it is that way because it is that way' wall can be found elsewhere in physics (why are the Standard Model, or the electron mass, the way they are?) and how the nature of this wall affects its intelligibility (how do, and should, we deal with this?).

Chapter 9: The future of superdeterminism

So far, much has been said about the arguments in favor of and against superdeterminism. These have mostly been confined to the realm of theory, with the angle primarily being philosophical. Given the ongoing debate on these matters, one may then wonder if superdetermism can ever be verified or falsified in any way, and how these debates can be expected to develop in the future. Historically, there has been some precedent regarding constraining, or even outright rejecting, some of interpretations of or alternatives to quantum mechanics, namely through the mathematical derivation of no-go theorems and the verdict of experimental physics.

To see to what degree these methods can be applied to the case of superdeterminism, we must briefly remind ourselves of a key element in the interpretation-theory distinction that was introduced in chapter 4. Simply put, an interpretation of quantum theory has the same empirical content but provides a particular ontology over QM_0, telling us how to understand what these equations say about how nature is. An alternative theory to quantum mechanics also comes with different predictions with regards to natural phenomena. Of specific importance for this chapter will be whether a given superdeterministic framework leaves the Born rule intact or not. A framework that does this is the aforementioned superdeterministic Cellular Automaton *interpretation* of ‘t Hooft, while one that does not is the Future-Bound Path Integral *theory* of Hossenfelder. Taking these as examples, this chapter will now say a word on how we can envision the future of these two approaches to superdeterministic model building. The former will be analyzed through the mathematical lens of a relevant no-go theorem, while the latter through that of experimental physics.

‘t Hooft’s interpretation and Landsman’s no-go theorem

Without getting to deep into the details of ‘t Hooft’s superdeterministic interpretation, let us briefly restate his overall approach. In ‘t Hooft’s model, physical reality is ultimately seen as classical and deterministic. On top of this, he rejects statistical independence, which he views as ‘going all the way’ with determinism as he includes the observers themselves as determined systems. He assumes that there is a set of states that the universe can, in fact, be in, and calls these ‘ontological states’. This very assumption is central to the interpretation. Due to the model’s determinism, the universe evolves through time by moving between these ontological states, but it always ‘picks’ and finds itself in one with a probability of 100%. He realizes this by introducing a cellular grid wherein the cells around each one change discretely and deterministically through time. Doing this all, ‘t Hooft aims to use his model to reproduce the Born rule, solve foundational issues such as the measurement problem and pave the way for developments in quantum gravity ('t Hooft, The Cellular Automaton Interpretation of Quantum Mechanics, 2016).

Due to the aim to reproduce the predictions of quantum mechanics, it is not possible to experimentally test the model, making it an interpretation in our nomenclature. However, if his interpretation turns out to be the right one, then his proposed search for ontological bases for quantum mechanics may provide the language in which a consistent and correct theory of quantum gravity could be formulated. In his own words:

“Finding quantum theories that have an ontological basis will be an important and difficult exercise. Our hope is that this exercise might lead to new theories that could help elementary particle physics and quantum gravity theories to further develop.” ('t Hooft, 2016, p. 45)

Given ‘t Hooft’s extensive work on creating this new ‘vocabulary’, as well as the lack of significant breakthroughs on the quantum gravity front over the last decades, this could be a welcome new endeavor that may one day establish its own correctness in this way. The interpretation may lead to theoretical and eventually, through quantum gravitational predictions, empirical progress.

Whether this will be the case is, of course, quite uncertain, with serious counterarguments to the fruitfulness of such a research program having been discussed in the last two chapters. However, this ambition further comes with its own unique problem. It may be that all truly deterministic interpretations of quantum mechanics as ‘t Hooft’s are incompatible with the Born rule. This is argued by mathematical physicist Klaas Landsman. His argument for this will be summarized below, and I will attempt to make its assumptions explicit such that its effect on superdeterministic *interpretations* of quantum mechanics can be fairly evaluated.

In his paper “Randomness? What Randomness?”, Landsman starts with an analysis of what should even be understood by the term. He views ‘randomness’ as a typical example of a ‘family resemblance’, as coined by Wittgenstein. That is, randomness is a word that has many different meanings depending on the context it is used in, that are in many ways similar, yet lack the kind of universal shared element commonly sought after in such cases. Specifically, these different meanings in the case of randomness can often best be identified by looking at what they are contrasted with, i.e., their antipodes. For example, randomness can in some contexts be used in contrast with full knowledge of a physical system, as in the case of the ‘randomness’ of a classical coin flip. After all, if you know exactly the forces applied to the coin at all times, its initial conditions such as the angle it is tossed at, height from the ground etc., classical mechanics in principle allows for the exact computability of the result. Another example is how Leibniz, in his correspondence with Clarke, contrasts his use of randomness with the presence of a determining cause. In his paper, Landsman primarily makes use of a notion of randomness that is well-suited for both the ends of physics and mathematics (as well as computing science), namely 1-randomness. For the definition to make sense, one must understand what a binary string is. This is a finite array of zeros and ones, for example ‘1001110100011’. A string in this way can be given meaning. An example from classical physics could be a number of coin flips, where ‘1’ refers to ‘heads’ and ‘0’ refers to ‘tails’. In the string ‘110’ I then first got heads twice, and then got tails once. A quantum physics example could be the measurement of the spin of a particle. ‘Spin up’ could be ‘1’ while ‘spin down’ could be 0. Landsman now provides three equivalent definitions of 1-randomness. In particular, a given binary string is 1-random if it is *incompressible*, *patternless* and *unpredictable*. Here, incompressibility means that the string itself is its own shortest description. For example, if I have a string ‘1010101010’, I can ‘compress’ it in some language through describing it as, e.g., 5*’10’. However, due to no such discernible pattern existing in a string like ‘111010011000010’, there is no way to write it shorter than to just write the entire thing itself. Patternlessness refers to the fact that there is simply no pattern in the entries of the string, while unpredictability means that there exists no possible strategy with which one can reliably predict a next string entry. Thus, we now have a working definition of 1-random strings (Landsman, Randomness? What randomness?, 2019, pp. 62-69).

Landsman now explains that strings resulting from the Born rule, for example through performing many independent and identical spin up or down experiments, are almost all 1-random. With this assumption, he utilizes more information about the nature of quantum randomness than Bell does: he does not just use the probability predictions of quantum mechanical averages (‘half will be spin up and half will be spin down’), but also what the outcome sequence, i.e., an infinite string, of a measurement run looks like (‘while 1010101010… has an equal probability of occurrence as any other string, the Born rule will not produce sequences with such patterns’). There is a more extensive logical and mathematical reasoning behind this which we shall not venture into too deeply, but a key insight used is an analysis of two procedures for doing repeated independent and identical quantum mechanical experiments. In the case of, for example, spin measurements, you can apply quantum mechanics to the whole run of the resulting string or to single experiments while using probability theory to combine these. These procedures turn out to be equivalent, and given the almost guaranteed 1-randomness of an infinite run, this then also holds for the single experiment case. Finally, note that the use of an infinite run is an idealization that the argument relies on, and can be subject of further discussion (Landsman, 2019, pp. 82-96). Nevertheless, Landsman shows that, realistically, the argument still works without it (Landsman, Indeterminism and Undecidability, 2021, pp. 10-11).

Having acquired an understanding of randomness, next we should look at Landsman's more precise definition of determinism. For him, there are two things a hidden-variable theory must adhere to, to be deterministic. First is the more common association that it is in fact the case that the hidden variable value λ determines the outcome of, e.g., a Bell test. This is also used by people like Bell, Kochen and Conway for the derivation of their respective no-go theorems. Landsman, however, adds another requirement, which consequently, *in conjunction with* the 1-randomness of quantum mechanical measurement sequences, allows him to *derive a stronger result*. This requirement is that, in principle, a deterministic theory should be able to provide the value of the hidden variable in any given experiment. If not, it will still not be possible to predict the outcomes of single experiments. Perhaps more importantly though, as will be described in a moment, it would necessitate a self-defeating appeal to an external random process (Landsman, 2019, pp. 96-100).

Logically, Landsman proceeds to differentiate between two cases: either the hidden-variable theory accounts for this second deterministic requirement or it does not. The first case can be shown to lead to a contradiction. This is due to Chaitin's second incompleteness theorem, which states that (for our practical purposes[89]) of any binary sequence only finitely many digits can be computed. From the computation of the hidden variables, the outcome sequences of experiments can be computed in the deterministic theory. But as established earlier, these sequences are 1-random upon observing the Born rule, which consequently requires computing infinite digits as the

[89] Like the more well-known incompleteness theorems of Gödel, the theorem holds only for axiomatic mathematical set theories that allow for at least basic arithmetic, such as the well-known Zermelo-Fraenkel set theory. A deterministic hidden-variable model in physics will most certainly fit this description, and thus the theorem will apply to it.

sequence can, by definition, not be compressed. This is then in violation of the incompleteness theorem, and thus the hidden-variable theory cannot account for the sampling of the hidden variables. The second case renders the hidden-variable theory unable to account for the outcomes of individual quantum experiments, from which Landsman claims nothing is gained over quantum mechanics. That is all the hidden-variable theory can then provide, by reintroducing the Born measure by averaging over the hidden variables in experiments. In a similar way, when considering simple classical coin tossing, the familiar 50-50 probability of heads or tails is found by averaging over the initial conditions, i.e., by doing probabilistic sampling. The initial conditions of the coin, such as its initial velocity, distance from the ground, etc., play the same functional role as do the hidden variables in spin experiments. But the only way to retrieve the Born rule from sampling like this, is if the sampling is done truly randomly. As was explained, a deterministic theory cannot be the source of this required randomness[90]. Landsman thus concludes that hidden-variable theorists are forced to invoke some ‘random oracle’, external to the theory, which he states undermines its very purpose. The effort is self-defeating. The ultimate conclusion is a new no-go theorem by Landsman: *the Born rule with its 1-random outcome sequences on the one hand, and determinism with its two clauses as described above on the other, are incompatible* (Landsman, 2021, pp. 8-11).

The typical response of both superdeterminists and pilot wave theorists is that the source of the required randomness lies in the initial conditions of the model. In general, when considering any isolated classical physical system, the evolution of the degrees of freedom contained therein through time is known for all times given the deterministic laws of nature as well as the initial conditions of the system. These initial conditions themselves are then not specified by the deterministic system. Thus, they can be an external source of randomness that is not in conflict with the deterministic nature of the system.

When introducing determinism in earlier chapters, however, it was explained that it is not only the future evolution of a system that is determined, but also its past. When a comet falls towards the Earth, and its position and momentum are known at one point in time, then not only does classical physics tell us the future trajectory of the comet, it also tells us its past trajectory. Given this aspect of determinism, Landsman points out that within the context of the kind of deterministic hidden variable theories

[90] This also has repercussions for the classical coin tossing example. Since the deterministic character of classical physics can by the same argument not provide truly random sampling, completely fair classical coins do not exist (Landsman, Randomness? What randomness?, 2019, p. 78).

under discussion, there is no difference between the Copenhagen school placing the origin of randomness at the *outcome* of a measurement while they themselves place it at the *initial conditions*. Again, one may ask where the real difference then lies (Landsman, 2019, pp. 78-79). I would also like to add that whatever the initial conditions may be, they still serve as input for deterministic equations or maps that, by virtue of their definition, produce compressible binary sequences. Thus, I would say that it is not clear how initial conditions can be responsible for 1-random measurement output sequences within the confines of a deterministic model. Moreover, when considering the universe as a whole, everything denoted as an 'initial condition' in any performed experiment, is already a result of this grand deterministic causal chain. From this zoomed-out point of view, only the initial conditions of the entire universe could even be random to begin with.

In conclusion, I think Landsman's argument poses a critical problem for superdeterminists aiming to reproduce the Born rule in their models. This argument, of course, hinges on some factors that one will have to accept in order to arrive at this conclusion. These consist in the definitions of randomness and determinism, the sharing of the conclusion that not much is gained through hidden variable theories if one needs to invoke an external source of randomness to justify the results of measurement runs and the idealization of infinite runs. Given the above, I think Landsman shows that these definitions as well as the use of the idealization are justified, with the latter in particular being so through the fact that even without it the argument, realistically, still works. A superdeterminist could object to the notion that 'nothing is gained' through hidden variable theories if an external source of randomness is still required. While determinism as used here would indeed not be established, other things still may be. For example, realism in the sense that quantum systems can be said to possess definite values of properties prior to any measurement could be, and with that a solution to the infamous measurement problem. On the other hand, if a new, truly random ontological layer underneath the hidden-variable theory is necessary to even justify it in the first place, we would end up with an untestable double-layered ontological extension, whereby the first 'pushes down' the problems to the second. Therefore, I am inclined to agree that given Landsman's no-go theorem, there does not seem to be a fruitful future for superdeterministic *interpretations* of quantum mechanics.

While the above-described incompatibility between superdeterminism and the Born rule may be said to close one door, another is still unaffected. Superdeterminists could just contest the Born rule. This will be the topic of the next subchapter.

Hossenfelder's theory and an experimental proposal

In one of her papers on superdeterminism, Hossenfelder directly addresses the work of Landsman, stating that to her, the point of superdeterminism is to find a description underneath quantum mechanics that does not always make the same predictions as it. She compares the situation to the aforementioned one between statistical mechanics and thermodynamics. The former gives the same predictions as the other only in certain limiting cases. In the case of superdeterminism, the limit in which the hidden variables are truly random is then to recover quantum mechanical predictions (Hossenfelder S. , 2020, p. 7). Because of this, Hossenfelder claims superdeterminism might be testable. We will now consider how she expects this to work and what can be said about it.

Together with Sandro Donadi, Hossenfelder developed her own superdeterministic model, based on future-bounded path integrals. This paragraph will briefly outline the core idea, after which its empirical implications will be considered. In brief, quantum field theory can be formulated by setting up the path integral[91] of a particular system and then demanding it to obey a certain condition[92]. This is one mathematical foundation from which quantum field theory can be derived, but the

[91] While not important for understanding what comes later, this footnote will very briefly describe what one can intuitively think of when considering a 'path integral', as it is mentioned numerous times in the paragraph. Please note that the following is a very barebones 'explanation' that is just meant to give some slight intuition of what is being talked about. The reader may skip it without consequence, as after this paragraph briefly describing the key concept behind Donadi's and Hossenfelder's model, we will restrict our discussion to whether this model is testable.
Suppose I am pushing a frictionless block forward with a constant force. From this, it gains energy of motion. That energy of motion can, in this case, be calculated my multiplying the force I exerted on the block with the distance I moved the block while pushing. But it may also be the case that the force to the block changes over time in magnitude and direction, and that I am not pushing in a straight line but through a wigglier path. I can then no longer just multiply force with distance. For one, what force to use, as it has changed throughout the pushed-over trajectory. Secondly, the direction of the force and path are relevant in this calculation. What I *can* do to deal with this, is to consider a very small interval of the total path I moved the block over. I can multiply that distance interval with the force I exerted on the block when I was at that location along the path. Such as small interval of a curvy line will be approximately straight, and the force will be approximately constant when I consider it only at the associated very brief time interval. I can then slice up the total path in many small intervals and do this multiplication for each of them, summing up all these contributions. Then, I have again calculated the energy of motion for this more difficult case.
This is an example of a path integral. In quantum field theory, the path is through spacetime, and we are not considering a force, but rather, a quantity composed of different types of energy of a quantum system under consideration. But the fundamentals of the mathematical recipe remain the same.

[92] This demand is that of the principle of stationary action. The principle is incredibly important in physics, and allows one to find the correct equations of motion from the path integral that has been set up.

further details need not concern us. Donadi and Hossenfelder take this approach and amend it by taking the path integrals over a different space[93] and accentuating the upper integration time of the integral. In the model, the time-evolution of a quantum system depends on the detector settings at the time the system is measured, hence the model's name. It should be noted that by future-input dependence, Hossenfelder does not mean to refer to retrocausality, i.e., information travelling back to the past. It is meant in the sense of future developments in the deterministic causal chain enabling statements such as "if measurement X takes place, then…". The future input would then be that measurement (Hossenfelder & Palmer, 2020, pp. 8-9). The determinism comes in because instead of considering all possible paths as in quantum field theory, the integral formulated for the detector and quantum state only follows one optimal path leading to macroscopically classical states. The hidden variables in this model are not unknown features of a quantum particle itself, but are associated with the degrees of freedom of the measurement device. When averaging over these uniformly distributed hidden variables, the model retrieves the Born rule. However, if the distribution is not uniform[94], for example due to a small sample size, this then need not be the case. This introduces the possibility of testability (Donadi & Hossenfelder, A Toy Model for Local and Deterministic Wave-function Collapse, 2022) (Donadi & Hossenfelder, A path integral over Hilbert space for quantum mechanics, 2022).

Based on such frameworks, Hossenfelder proposes experimental conditions for testing a subset of superdeterministic models, specifically those for which the hidden variables, like in her model, stem from the correlation with the measurement device (as opposed to elsewhere in the universe). She describes how similar experiments were already described long ago by von Neumann, and that this is a possible realization of these ideas (Hossenfelder S. , Testing Super-Deterministic Hidden Variables Theories, 2011, pp. 9-11). According to her, carrying out these experiments is possible, but this has not yet been done due to the "obsession with Bell-type tests" of the quantum foundations community. Together with "lack of a

[93] More precisely: it is not paths in spacetime that are summed, but paths in Hilbert space. In chapter 1, the quantum mechanical wave function or quantum state vector was introduced. These vectors 'live' in a Hilbert space.

[94] This was referred to by Valentini and Sen in chapter 8 as the nonequilibrium state.

generally applicable fundamental theory[95]", she views this as the most significant problem for superdeterminism at the moment (Hossenfelder S. , 2020, p. 18).

The experimental situation she envisions is as follows. As a starting point, she wants to consider the results of measurements on identically prepared states. According to quantum mechanics, the Born rule predicts the statistical distribution of outcomes. One then expects no correlations between them. But according to a deterministic hidden-variable theory, if the hidden variables are also included in what is identical about the states, the measurements should yield the same result each time. This implies a strong correlation instead. The obvious problem is that the hidden variables are unknown, so preparing identical states seems to be an impossible task. That is why she proposes doing the same experiment consecutively on the same system, with at least some probability of the system being returned to the initial state after measurement. The idea is that in a deterministic system, properties you measure evolve through time in a continuous manner. They do not, as in the probabilistic picture, take on whatever value in the set, all with a probability described by the Born rule. Because of the continuous evolution seen with determinism, after the passage of a small increment of time, values of properties will still be quite close to one another. After each measurement, the state is returned to its initial state. The revelation of these values one after another will then carry a correlation between them (Hossenfelder S. , 2011, pp. 4-6) (Hossenfelder & Palmer, 2020, pp. 18-19).

A number of considerations are then introduced that are meant to minimize changes in the initial state. The most important ones are as follows them (Hossenfelder S. , 2011, pp. 6-8):

- *The use of small measurement devices*. Since in Hossenfelder's model, the hidden variables are associated with the detector's degrees of freedom, it is assumed that the bigger the detector the more hidden variables will be relevant. These are then likely statistically distributed which yields Born-like measurement outcomes, so this probability will decrease when less are involved, i.e., a smaller detector is used.
- *The measurements should be performed one after another as fast as possible*. Changes of the hidden variables in between them should, after all, be minimized in order to detect a correlation.

[95] There are currently various superdeterministic (toy) models, but a full-fledged, single fundamental theoretical framework is not yet available. There is, therefore, still much more theoretical work to be done on this.

- *The setting should be as cool as possible*. A lower temperature implies slower change, but it is also of specific importance for the setup she has in mind, as will be shown in a moment. This can, however, not be done by immersing the system in a cooling liquid, because this would greatly increase the number of atoms in the vicinity, increasing the amount of background noise that needs to be minimized.

The concrete experiment proposed by Hossenfelder will be described briefly in a simplified manner. The setup is shown in figure 14.

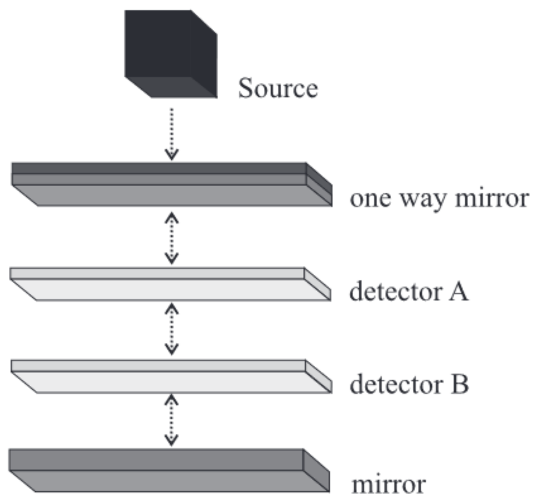


Figure 14: The experimental setup with which a subset of superdeterministic models could be tested, as contrived by Hossenfelder. The source emits a particle with a small probability of passing the mirror. The particle is then likely to travel up and down many times, being subjected to measurements of two non-commuting variables by detector A and B. Since according to quantum mechanics the measurement of one should destroy information of the other, this allows measurements on the particle to be made in rapid succession while the particle is effectively returned to its initial state each time (Hossenfelder S. , 2011, p. 7).

In the experiment, a particle, such as a photon or electron, is emitted towards a 'one way mirror'. In reality, this mirror will allow, for example, 1% of particles to pass. Given the same holds for the mirror at the other end, any particle that passes the first mirror is likely to be moving up and down at least tens of times. While doing so, two non-commuting variables of the particle are measured at detector A and B respectively. The traditional example introduced all the way back in chapter 1 was that of position and momentum. According to standard quantum mechanics, information about one of these will destroy any information on the other. One may here recall the uncertainty principle. Therefore, the setup allows for the making of measurements in rapid succession whereby each time, the particle is effectively returned to its initial state. The array of measurement outcomes of the detectors can then be studied, in particular the question of whether the results display correlations (according to some superdeterministic models) or not (according to the Born rule, and thus quantum mechanics).

Using some physical modelling, Hossenfelder ultimately approximates the duration of correlation between consecutive results to be in the order of microseconds. Present day technology allows for the measurement of correlation times of this order. The result of this is that superdeterministic hidden-variable models and quantum mechanics yield different predictions for how long a particle can be expected to remain in the mirror loop. Repeating the experiment for large ensembles of particles will then result in detectably different confinement times. Thus, while some assumptions are made throughout, and although this would only work for the subset of superdeterministic models where the hidden variables are associated with the degrees of freedom of the detector, the final result is proposed experiment that could differentiate between quantum theory and such superdeterministic hidden variable theories. While according to Hossenfelder a test like this has not yet been performed, it could be a way of putting her model to the test (Hossenfelder S. , 2011, pp. 7-10).

As far as I can tell, which, as only a Master's student not even in an experimental track is not very far, this setup appears realizable. It is in any case good scientific practice that Hossenfelder explicitly attempts to put her model up to the test. One may however wonder if such tests, intentionally or not, have not already been performed. The predictions of quantum mechanics have themselves been confirmed up to tremendous precision, including the possibility of nonlinear extensions[96] that superdeterministic models can belong to. For example, following the well-known quantum physicist Steven Weinberg's paper on recommendations of how to test such extensions, a collaboration of experimenters have shown that nonlinear corrections to the energy of a quantum system could be present up to a limit of 26 orders of

[96] This can be done by adding a term to the Schrödinger equation. As was explained in chapter 1, this is a linear equation, and that linearity leads to the possibility of superpositions. The addition of a nonlinear term is, therefore, a strategy of many models that aim to explain (away) wave function collapse. After all, the need for collapse is there because superposition itself does not explain why we always observe singular measurement outcomes.

magnitude below the binding energy of a nucleon (Weinberg, 1989) (Bollinger, Heinzen, Itano, Gilbert, & Wineland, 1992). Since the latter is typically of the order of ~10 MeV, we are then talking about energy measurements with a precision up to 10^{-19} eV, or 10^{-38} joules. Yet, any greater than this, and the corrections are incompatible with the data. These incredibly small constraints were already reached in this 1992 experiment, and the capabilities of the experimental physics community have come a long way since then. Consequently, it seems reasonable to at least ask how it can be that given the relatively modest parameters required for performing Hossenfelder's proposed experiment, no deviations from standard quantum mechanics of this kind have been noticed yet. This might count as a point for quantum mechanics, but this can of course, as argued by Hossenfelder, also be due to the specific experiments needed to establish such effects not having been done before. In that case, her experiment would make for a highly important test for the validity of a significant subset of superdeterministic models, with different experimental setups conceivably expanding that set.

Thus, while superdeterministic hidden-variable *interpretations of quantum mechanics* such as 't Hooft's cellular automaton interpretation seem to reach a dead end if one accepts Landsman's results showing an incompatibility between them and the Born rule, superdeterministic hidden-variable *theories underlying quantum mechanics* may sooner or later be put up to definitive empirical tests due to them diverging from the Born rule in some experimental contexts. This, then, provides to us an image of the possible future of superdeterminism.

Conclusion

With the input of chapter 1 up and including 9, we have now reached the conclusion. First, the research question, thesis structure and viability criteria will be briefly repeated and summarized. After this, it will be explained whether superdeterminism can measure up to the criteria. Lastly, the research question itself will be answered.

In the introduction of this thesis, the following research question was presented:

"Is superdeterminism a viable theoretical framework for addressing foundational philosophical questions in quantum theory?"

With this question and the meaning of viability in mind, each chapter had a role in ultimately being able to answer it. The first two chapters functioned to provide the background with which these foundational philosophical problems could be understood. With superdeterminism originating as one of the two hidden-variable theorists' answers to Bell's theorem, chapter 3 followed to describe this situation. In chapter 4, we considered the meaning of terms like 'theoretical framework' that appear in the research question. Chapter 5 introduced some 'competitors' of superdeterminism, which aided in understanding the frequent comparisons during evaluations and the ability to assess viability in a broader context. We then moved on to thoroughly explain what superdeterminism is in chapter 6. To study the viability of the framework, the bulkier chapter 7 and, in particular, chapter 8, consisted of an as complete as possible description and evaluation of arguments for and against superdeterminism, respectively. Finally, chapter 9 addressed the viability question through the lens of whether superdeterministic models can be subject to verification or falsification through logic and experimental science.

Having gone through all of this, we can now answer the research question by considering superdeterminism's compatibility, or lack thereof, with the defined viability criteria. All of these have been addressed either implicitly or explicitly throughout the text. The criteria are repeated below:

- *Can superdeterminism provide clear answers to the philosophical and foundational questions and problems in quantum theory?*
- *Is superdeterminism self-consistent?*
- *Is superdeterminism consistent with well-established physical theories in all measurable regimes?*
- *Does superdeterminism have the prospect of possibly being testable now or in the future?*

- *Are the metaphysical consequences of superdeterminism acceptable?*

Using the conclusions reached in the many (sub)chapters, each of these questions will now, in order, be answered in a way that aims to combine nuance and decisiveness as set out in the introduction.

- *Yes*. Regarding the list of foundational philosophical questions in quantum theory that was provided in chapter 4, it was found in chapter 6 that superdeterminism, through how the framework is defined, can clearly answer all of these questions. This was advanced as a strength in chapter 7, as this cannot trivially be said for all competitors. On the other hand, superdeterminists were argued to overstate their case when they suggest their foundational stances could pave the way for a Theory of Everything.
  One point of nuance is that this list is, as stated, nowhere near exhaustive of all questions that appear in the large subfield of the foundations of quantum mechanics. Due to it hardly being possible to cover everything, the main ones have been selected. For instance, concepts like quantum contextuality and experiments like the Elitzur-Vaidman bomb tester, particularly with regards to how superdeterminism interprets and explains these, have not gotten much attention in this thesis. I do, however, suspect that it would be able to provide a clear interpretation of such phenomena as well. Superdeterminism, in my view, is good at addressing the usual problems, but it does so, arguably, at an at least as great cost of introducing new ones such as the conspiracy.
- *Yes for theories, but no for interpretations*. Superdeterministic hidden-variable *interpretations* were shown not to be self-consistent through the theorem of Landsman discussed in chapter 9. Specifically, the property of determinism when also applying this to the sampling process of the hidden variables, was shown not be compatible with such models' claim of reproducing the Born rule. However, even if one accepts this argument, this does not expose an inconsistency in superdeterministic hidden-variable *theories*, such that the framework as a whole need not be so.
  Other than the above, we have found at least one *apparent* inconsistency, namely in the fact that superdeterminism claims to be local yet seemingly allows for superluminal communication. The resolution of this paradox can, as was explained by Valentini and Sen, be found in the fact that upon inspection, there are not truly signals present in this communication. There are only series of coincidences mimicking an actual conversation, thereby, I anything, motivating the conspiracy argument.

- *Yes if one is concerned purely with empirical compatibility, no if the foundational principles of these physical theories are also considered*. This conclusion was drawn at the end of our evaluation of the science-invalidation argument. Superdeterminism was argued to necessitate the existence of an absolute rest frame, which was explained to be fundamentally incompatible with a relativistic understanding of spacetime. Concerning foundational principles, this is a clear inconsistency. Specifically, this result was based on the particular model under consideration, but it was described that these conclusions carry over to all others. Because the model chooses initial conditions that make it empirically indistinguishable from quantum mechanics, however, this absolute rest frame would be inaccessible to us. Therefore, on a purely empirical basis, the required consistency with relativity is achieved.
  While not being structurally similar to relativity theory, it was noted that superdeterminists may invoke structural similarity with well-established physical theories as a reason for investigation. This was shown for the case of the Liouville equation in chapter 7.
- *Yes, in relation to the theories*. It was shown in chapter 9 that superdeterminists like Donadi and Hossenfelder have built a model that comes with the possibility of an experimental test against quantum theory. The latter was described to have been tested to extreme precision, constraining possible extensions thereof to much greater orders of magnitude than appear in Hossenfelder's experimental proposal. Nevertheless, these are different experiments with other parameters being involved, and hers has not yet been tested. For at least the subset of superdeterministic models considered by Hossenfelder, there is the prospect of testability.
- *Cannot be answered objectively*. Throughout the text, it was shown that superdeterminism comes with significant metaphysical baggage. Some of these are intended for and come with the definition of the framework, such as determinism and locality even in the PL1 sense. More importantly, there are metaphysical consequences that originate from superdeterminism in a corollary fashion. These were studied in chapter 8, and several have been found. Due to the controversial nature of their acceptability we note that, in line with comments in the introduction, no objective answer on this question is possible. To provide the information with which one could answer the question in this criterium, the results will be listed below, accompanied by a brief summary of their content and impact evaluation.

- The first metaphysical property of nature given superdeterminism, is the presence of *ubiquitous conspiratorial correlations*. The conspiracy argument has been discussed in great detail. It implies that nature must be fine-tuned in a very specific way. In every Bell-like test, for whichever of practically infinite ways of installing the measurement settings, there is always a correlation between the setting mechanism and the probability distribution of the hidden variables determining, e.g., the real spin state of a quantum object, such that we observe statistics in agreement with the predictions of quantum theory. Valentini and Sen go beyond just this and were shown to both quantify the degree of fine-tuning involved and lay bare the multiple sources thereof. This all has far-reaching implications for our understanding of nature that can hardly be dissociated from the intuition of an absurd conspiracy by nature. Nevertheless, superdeterminists were argued to correctly show that this line of argument itself depends on a number of metaphysical assumptions. For one, in its claim that these correlations stem from fine-tuned initial conditions of the universe, proponents of the conspiracy argument implicitly appeal to the use of probability theory. However, certain assumptions were shown to be required in order to do this. Prominent example of this are the existence of a space of possible sets of initial conditions and a probability distribution over this space. Moreover, ‘t Hooft has pointed out that at least on the surface, fine-tuning arguments flow less smoothly for type I superdeterministic models, wherein a causal law operating in the early universe when systems were still in causal contact, is responsible for the observed correlations.
  While having acknowledged that these responses have some merit, it was ultimately, for several reasons, argued that they do not undermine the conspiracy argument, especially not in the case of the latter. They, in particular the probability-conditions objection, are good to keep in mind for showing that even the conspiracy argument itself hinges on metaphysical assumptions.
- If one wishes to avoid the consequences of the science-invalidation argument, then one must commit to, among other things, the existence of *absolute space*. See the third criterium above.

- Similarly, one may have to commit to a view of nature wherein *each spatial point or cell can be thought of as an autonomous supercomputer that is running a perfectly detailed simulation of the entire universe*. The degrees of freedom describing these internal universes entail *extreme ontological enlargement* and, moreover, lead to the *solipsistic situation* of being unable to acquire information about anything outside the space one oneself occupies. At least, this is the case if the observation by Ciepielewski, Okon and Sudarsky that this holds for all superdeterministic models is accepted.
- Lastly, one must accept that *the universe does not allow for libertarian free will*. While this concept was argued to be incoherent, it is nevertheless another metaphysical consequence of superdeterminism that can be listed. However, due to the expressed harmlessness and inescapability of this conclusion even outside of superdeterminism, I would propose focusing more so on the earlier entries.

This concludes our analysis of the applicability of the viability criteria to superdeterminism. In the introduction, it was argued that not only is there not an objective method of determining the ‘right’ criteria to use, there is also none for how to combine the answers thereto in one binary output regarding the research question. For this reason, it was decided there to be both explicit and complete such that the reader is able to make up their own mind in the face of them, irrespective of my own verdict.

That being said, based on the criteria for viability, I would ultimately draw the conclusion that superdeterminism is a viable theoretical framework for addressing foundational philosophical questions in quantum theory. Superdeterministic hidden-variable theories in particular were seen to, in principle, be able to check all criteria defined to evaluate this viability question. The primary difficulties were seen to lie in the last one. I would submit that metaphysical consequences in the form of

- a world full of conspiratorial correlations between, effectively, setting mechanisms and system states, being postulated with the goal that quantum statistics are always observed in every possible Bell-like test;
- space as being made up of autonomous points or cells, containing themselves a simulation of the entire universe in a Leibnizian, parallelist fashion;

- this space, further, being absolute, thereby adhering to a philosophy of time and space that directly contradicts the foundational one of the theory of relativity;

are, at least to me, very hard to accept. Yet, none of these constitute a direct rebuttal of superdeterminism. Moreover, while insufficiently so in my view, most of these can be pushed back on in legitimate ways, such as through the nuances from the probability-conditions argument and the empirical inaccessibility of superdeterminism's absolute space. Lastly, some physicists and philosophers do not judge the understanding of nature based on these metaphysical principles to be less acceptable than the ones that follow from other, competing frameworks. Given its popularity in the community, and the serious recognition it receives, the perpetual creation of new universes after every measurement in the Everett interpretation is apparently not a reason to judge it as unviable on metaphysical grounds. Similar arguments can be made on the basis of the anti-realism, anti-reductionism and spooky action of the Copenhagen interpretation. On page 3, Heisenberg was quoted to have asked whether nature could possibly be so absurd as it seems to us in atomic experiments. Given all the interpretations thereof, perhaps it is. Then who am I to say that, e.g., the conspiracy of superdeterminism, is objectively more disqualifying than the perpetual universe creation of the many-worlds interpretation. Superdeterministic hidden-variable theories may, in fact, even have the advantage when also considering the other criteria, such as testability.

In conclusion, based the criteria of viability and the above reasoning, I am inclined to answer the research question in the affirmative.

Discussion

The discussion section will be made up of six parts. Firstly, I will briefly mention my own opinion on superdeterminism. After this, a brief 'bonus part' will follow, by picking up the question in chapter 3 as to what Einstein may have thought about superdeterminism. Thirdly, implications of this research will be discussed, after which its shortcomings will be explicated. The fifth part will concern possible suggestions in response to these shortcomings. The final part will contain brief words of gratitude and remarks on 'the making of' this thesis.

My own view on superdeterminism

In the conclusion, it was stated that superdeterministic hidden-variable theories are a viable theoretical framework for addressing foundational philosophical questions in quantum theory. In the introduction, however, it was already noted that viability ought not to be conflated with truth-value. Viability, it was described, is more so about whether the framework is deserving of serious recognition when compared to its competitors, with it at least conceivably turning out correct. Truth is, of course, about whether superdeterminism accurately describes the nature of physical reality.

While granting that superdeterminism can be a viable framework on the basis of the defined criteria, I do not personally belief it to be true. Admittedly, though, superdeterminism is often too easily dismissed. The most prominent example of this is the free will argument, which as recognized by both proponents and opponents is not a fair argument against superdeterminism. Moreover, the science-invalidation argument in the form of the one-to-one comparison with a hamburger experiment-like analogy doing the trick, was also argued to be insufficient on its own. Less extensively, it was also mentioned that superdeterminism is not retrocausal, does not allow for superluminal signaling and is not open only to a neo-Humean conception of the laws of nature.

My reasons for not believing in the framework are the conspiracy argument and the consequences of trying to restrict violations of statistical independence only to specific experimental contexts in the quantum realm. While even on these issues, superdeterminists can respond to some extent, I do not think they can refute these critiques. Ultimately, the conjuring of correlations, with nature seemingly conspiring against us to make it appear that quantum mechanical laws reign supreme, just seems too far-fetched to me. The many examples of bizarre setting mechanisms that would require absurd correlations with the hidden variables of an atom created in a lab here to measure, is a good way to draw this picture out. It leads both too extreme fine-tuning and expansion of ontological structure that is not warranted by the facts. It appears to me that while having much of the same merits argued to be present in superdeterminism, objective-collapse theories, as an example, simply do not require all of the above. Therefore, while I can understand some are drawn to the framework, and that there is more to it than sometimes assumed, I am not inclined to believe in it. As Bell wrote on the topic:

"When that theory [superdeterminism] is announced I will not refuse to listen, either on methodological or other grounds. But I will not myself try to make such a theory." (Bell J. , 1987, p. 103)

Bonus: would Einstein have been a superdeterminist?

As extensively discussed in chapter 2, Einstein was a local realist who was amiable to the hidden-variable approach. As he died before Bell's theorem was established, we will never know what he would have done in response to it. Before its publication, after all, there was no need for hidden-variable theorists to think about nonlocality or statistical dependence.

One may be inclined to belief that Einstein would have committed to superdeterminism. After all, it is the only way in which the local realism he so strongly believed in can still be attained. I have even seen some comment sections and forums online where I have stumbled upon this suggestion.

Yet, I think there are reasons to believe that Einstein may not have been willing to throw out statistical independence. One more obvious reason is the established incompatibility with the foundational principles of relativity that this leads to. But another possible reason follows from Einstein's strong commitment to the separability principle. This principle states that systems that are spatially separated can be described independently from one another. It is about the possibility of individuation of physical systems. For Einstein, it was a grounding principle for his realism. He explained this as follows:

"However, if one renounces the assumption that what is present in different parts of space has an independent, real existence, then I do not at all see what physics is supposed to describe. For what is thought to be a 'system' is, after all, just conventional, and I do not see how one is supposed to divide up the world objectively so that one can make statement about the parts." (Einstein, translated by Howard, 1985, p. 191)

Thus, Einstein accentuates that without the separability principle, physics cannot describe somehow localized physical systems if no system can be individuated from any other. His own, later version of the EPR-argument put it much more to the forefront, as he was unhappy with the backseat it took in the 1935 paper with Podolsky and Rosen (Howard, 1985).

Knowing this, I would argue that Einstein may not have liked the strategy of relinquishing statistical independence. After all, we have seen that this would imply pre-existing correlations between all physical systems. Consequently, it would no longer be the case that these can be fully described on their own. As an example, superdeterminism allows for situations where I can only fully describe the spin measurement of an atom when invoking the correlations with TV-show pixels used to determine the measurement settings for an experiment on another atom at the other side of the galaxy. Since physical systems cannot be fully described without reference to correlations with all other systems, superdeterminism can be argued to violate the separability principle. Ergo, Einstein might have been very critical of superdeterminism.

This is all not to say that I can say with certainty what Einstein would have done when confronted with Bell's theorem. But at the very least, I do not think it is obvious that he would have been a superdeterminist. Given the fundamental status he ascribed to the separation principle, in addition to the issue of absolute space, I strongly doubt that he would have.

The implications of this research

The conclusion that superdeterministic hidden-variable theories can be considered viable, could have some interesting implications. For one, superdeterminism could emerge from its relatively fringe position in the discourse surrounding the interpretations of quantum mechanics and become another more familiar name subject to widespread discussion. Secondly, the argumentation that has led to the answer to the research question might affect the content of already-held discussions on superdeterminism. The use of currently persistent but weaker arguments for or against it, such as the free will objection, could then diminish. Lastly, it might be a reason for more theoretical and experimental research into it, the possible details of which will be discussed among suggestions for further research later.

The shortcomings of this research

Throughout this thesis, some required choices can be argued to have led to shortcomings. One of these is the relatively little attention paid to the many specific superdeterministic models out there. While some models have received broader consideration, this was usually in the service of the general philosophical discussion on the framework as a whole. An analysis of the mathematical methods in these models, or the physical mechanisms postulated by them, was mostly absent. Another possible shortcoming lies in the choice of philosophical questions in quantum foundations that were considered, as well as the viability criteria that were selected. It was mentioned multiple times that in both cases, these were not exhaustive lists encompassing all possible entries. Further, in the case of the criteria, there is no general consensus on these particular ones having to be used to determine the viability of a theoretical framework in the context of interpreting or finding alternative theories to quantum mechanics. Finally, the style of the thesis may be a shortcoming in the sense that it will not appeal to all audiences. It was written with the intent of being understandable to all philosophy Master students, but that did entail an extensive deep dive in necessary context surrounding the topic at hand. Combined with the absence of mathematics to this end, and the frequent use of simplifications, this made for a long thesis. Although they were not the primary audience, this could make the text unattractive to experts in the field looking for a review on the viability of superdeterminism. There is no doubt in my mind that further shortcomings and weak points could be detected, although great effort has been put in achieving the goals that were set out, as well as presenting information and arguing in a way that is fair to both superdeterminists and those critical thereof.

Suggestions for further research

This research leaves plenty of opportunities for further investigation. I will describe five ideas about this below.

Firstly, I would suggest experimental physicists to look at the proposals made by Hossenfelder to test a significant subset of superdeterministic-hidden variable theories. As we have seen, these experiments seem realizable, and they do not require precision close to CERN-level capabilities. If Hossenfelder's model does not hold up, then alongside with superdeterministic interpretations (see chapter 9), a significant subset of superdeterministic theories is also ruled out. Even if most, including myself, will not expect it to, if it would hold up the result would be revolutionary.

Secondly, there is no fundamental theory of superdeterminism that is generally applicable to all of physics. Currently, there exist only a number of distinct superdeterministic (toy) models that still leave many questions unanswered. Proponents of superdeterminism may want to contribute to developing these, something both 't Hooft and Hossenfelder wish to inspire because few people are currently doing so. Opponents may instead further investigate the viability of individual models and whether they hold up, as this was described as a shortcoming for not having been done here.

Thirdly, some arguments in favor of and in opposition to superdeterminism could be explored further. While some, such as the conspiracy argument, got plenty of attention, this was not the case for, e.g., the inexplicability argument. In that particular case, there exist interesting research opportunities for studying how superdeterminism relates to our understanding of a scientific explanation.

Fourthly, relating to one of the shortcomings, the work done in this thesis could be presented in a shorter and more technical way for experts in field of quantum foundations. This could serve to provide them with both a general overview of superdeterminism and the tools to judge the viability of the framework themselves.

Fifthly, and finally, similar viability studies may be done for other less well-known interpretations of and alternatives to quantum theory. Doing so using the same viability criteria and methodology would enable fair comparisons between frameworks.

In conclusion, superdeterminism is a relatively little explored framework in the 'interpretation' of quantum theory, and many papers about it have only appeared in the last few years. This, therefore, leaves plenty of opportunities for further research.

Personal evaluation and acknowledgements

Doing my Master's thesis in the Philosophy of Science on superdeterminism and, more broadly, the interpretations of quantum mechanics, has brought me a lot. I sincerely hold that it has contributed more to my *understanding* of quantum mechanics than all of my Bachelor's courses in physics did combined. It has been a fascinating experience to get to know this wildly interesting field. The deep connection between physics and philosophy has also never been clearer to me. Finally, I feel like I have a much greater grasp of the relatively esoteric theoretical framework of superdeterminism, and hope to even have contributed to the debate surrounding it at one or two points.

Content aside, this thesis has also been great practice in philosophical writing, especially in trying to explain physical and philosophical concepts in an as understandable possible way. I have also gained some more insight in academic processes and research at universities.

One downside to the motivation I felt for this topic is that it turned out that the research question this translated into was too ambitious. Superdeterminism may be relatively new and somewhat niche, but the existing community has nevertheless generated quite some discourse. Further, one demand of the thesis is that it must be understandable to all Master's students of philosophy. This demand is both understandable, admirable, and has been taken up with great seriousness by me. One disadvantage, however, is that laying out quantum theory and the long history leading to superdeterminism while, furthermore, barely using any of the mathematics, also requires quite some space. The end result of these factors is a thesis that exceeds the word limit by more than 2.5 times that same amount. Scrapping this to create a version that can readily be handed in (not this one), has not necessarily been a joyful process, and it is a hard lesson I will be sure to take with me.

Given this situation, some words of gratitude are in order. First of all, I would like to thank Prof. dr. Henk de Regt for supervising this thesis, including the always pleasant discussions, critically thinking along with me on form and content, and his endless patience. Regarding the latter, I have kept you waiting longer than I intended to more often than I would have liked, but I sincerely appreciated the trust I felt you had in my ability to complete this project. I also want to thank Prof. dr. Klaas Landsman for more than once sharing his extensive knowledge on the topic of this thesis. This has led me to many new alleys of thought and papers including his own work, which has been instrumental in the advancement of my understanding, especially in chapter 9. Finally, I want to thank my girlfriend, friends and family for not throwing me out of the window after telling them this project would be done in 'just one more week' for the N^{th}-time. The scrapping process has often been compared by me to a Jenga game, where it was sometimes hard to predict that a tower would collapse when deleting one part, causing it to take more time than accounted for to rebuild. Ultimately, though, I can say that I am proud of the resulting tower.

Bibliography

Adams, F. (2019, February 11). The Degree of Fine-Tuning in our Universe -- and Others. *Physics Reports*, pp. 140-143.

Adequate (or Statistical) Determinism. (n.d.). Retrieved from The Information Philosopher: https://www.informationphilosopher.com/freedom/adequate_determinism.html

Amos, J. (2022, October 4). *Physics Nobel rewards 'spooky science' of entanglement*. Retrieved from BBC News: https://www.bbc.com/news/science-environment-63121287

Andreoletti, G., & Vervoort, L. (2022, August 24). Superdeterminism: a reappraisal. *Synthese volume 200, Article number: 361*.

Ayer, A. (1972). Freedom and Necessity. In *Philosophical Essays* (pp. 271-284). London: Palgrave Macmillan.

Baas, A., & Baptiste, L. B. (2020, September 29). What does the world look like according to superdeterminism? *(forthcoming in the) British Journal for the Philosophy of Science*, pp. 1-20.

Bacciagaluppi, G. (2020, April 21). *The Role of Decoherence in Quantum Mechanics*. Retrieved from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/qm-decoherence/

Bacciagaluppi, G. (2021, December 8). Lecture: Introducing Foundations of Quantum Mechanics. *Geschiedenis en Grondslagen van de Natuurkunde, Radboud Universiteit*, p. 41.

Bacciagaluppi, G., & Valentini, A. (2009). *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference.* Cambridge: Cambridge University Press.

Baggott, J. (2021, December 6). *Calculate but don't shut up*. Retrieved from aeon: https://aeon.co/essays/shut-up-and-calculate-does-a-disservice-to-quantum-mechanics

Bell, J. (1964, November 4). On The Einstein Podolsky Rosen Paradox. *Phys. Phys. Fiz. Vol. 1, No. 3*, pp. 195-200.

Bell, J. (1987). Free variables and local causality. In J. Bell, *Speakable and unspeakable in quantum mechanics* (pp. 100-104). Cambridge: Cambridge University Press.

Bell, J. (1990, August). Against 'measurement'. *Physical World 3 (8)*, pp. 33-40.

Berkovitz, J. (2007, January 26). *Action at a Distance in Quantum Mechanics*. Retrieved from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/qm-action-distance/

Bohm, D., & Aharonov, Y. (1957, November 15). Discussion of Experimental Proof for the Paradox of Einstein, Rosen and Podolsky. *Physical Review Vol. 108 No. 4*, pp. 1070-1076.

Bohr, N. (1935). Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Physical Review Vol. 48*, 696-702.

Bohr, N. (1949). *Discussion with Einstein on Epistemological Problems in Atomic Physics.*

Bollinger, J., Heinzen, D., Itano, W., Gilbert, S., & Wineland, D. (1992, April). ATOMIC PHYSICS TESTS OF NONLINEAR QUANTUM MECHANICS. *Foundations of Quantum Mechanics*, pp. 40-46.

Bourget, D., & Chalmers, D. (2014). What do philosophers believe? *Philosophical Studies 170*, 465-500.

Brans, C. H. (1987, August 12). Bell's Theorem Does Not Eliminate Fully Causal Hidden Variables. *International Journal of Theoretical Physics Vol. 27, No. 2*, pp. 219-226.

Budroni, C., & al., e. (2022, August 31). Kochen-Specker Contextuality. *arXiv:2102.13036v4 [quant-ph]*, p. 68.

Casado, C. (2008, May). A brief history of the mathematical equivalence between the two quantum mechanics. *Lat. Am. J. Phys. Educ. Vol. 2, No. 2*, pp. 152-155.

Chen, E. K. (2020, October 23). Bell's Theorem, Quantum Probabilities, and Superdeterminism. *Forthcoming in Eleanor Knox and Alastair Wilson (eds.), The Routledge Companion to the Philosophy of Physics.*, pp. 1-22.

Ciepielewski, G., Okon, E., & Sudarsky, D. (2020, August 3). On Superdeterministic Rejections of Settings Independence. *ArXiv*, pp. 1-38.

Collaboration, T. B. (2018, November 9). Challenging local realism with human choices. *Nature, volume 557*, pp. 212-216.

Conway, J., & Kochen, S. (2006, April 11). The Free Will Theorem. *arXiv:quant-ph/0604079v1*, p. 31.

Crull, E., & Bacciagaluppi, G. (2016). *Grete Hermann - Between Physics and Philosophy*. Berlijn: Springer.

Dalton, M. (2021, March 2). *Is our universe super-deterministic? (# 31)*. Retrieved from Medium: https://medium.com/intuition/is-our-universe-super-deterministic-31-98a2885ccbeb

Dattani, S. (2022, February 10). *Why randomized controlled trials matter and the procedures that strengthen them*. Retrieved from Our World in Data: https://ourworldindata.org/randomized-controlled-trials

Davies, P., & Brown, J. (1993). *The Ghost in the Atom: A Discussion of the Mysteries of Quantum Physics.* Cambridge: Cambridge University Press.

De Regt, H. (2017). *Understanding Scientific Understanding.* Oxford: Oxford University press.

De Regt, H. (November 2021). Understanding Gravitation. *Geschiedenis en Grondslagen van de Natuurkunde* (pp. 1-25). Nijmegen: Radboud Universiteit.

De Vuyst, J. (2020, December 10). A Natural Introduction to Fine-Tuning. *arXiv:2012.05617 [physics.hist-ph]*, pp. 1-27.

Donadi, S., & Hossenfelder, S. (2022, May). A path integral over Hilbert space for quantum mechanics. *Annals of Physics Vol. 440*, pp. 1-6.

Donadi, S., & Hossenfelder, S. (2022, August 19). A Toy Model for Local and Deterministic Wave-function Collapse. *Physical Review A 106, 2*.

Einstein, A. (1949). *Albert Einstein: Philosopher-Scientist ("Autobiographical notes").* Chicago: Open Court Press.

Einstein, A. (2010). *The World As I See It.* New York: Philosophical Library.

Einstein, A., & Born, M. (2004). *Born-Einstein Letters, 1916-1955: Friendship, Politics and Physics in Uncertain Times.* London: Palgrave Macmillan.

Einstein, A., Podolsky, B., & Rosen, N. (1935). Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Physical Review Vol. 47*, 777-780.

Erkoç, Ş. (2009, January). History of Science for Science Courses: “Spin” Example from Physics. *Latin American Journal of Physics Education Vol. 3 No. 1*, pp. 9-12.

Faye, J. (2002, May 3). *Copenhagen Interpretation of Quantum Mechanics*. Retrieved from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/qm-copenhagen/

Friederich, S. (2021, November 12). *Fine-Tuning*. Retrieved from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/fine-tuning/

Ghirardi, G., & Bassi, A. (2020, May 15). *Collapse Theories*. Retrieved from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/qm-collapse/

Goldstein, S. (2021, June 14). *Bohmian Mechanics*. Retrieved from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/qm-bohm/

Griffiths, D. J. (2014). *Introduction to Quantum Mechanics Second Edition.* Harlow: Pearson Education Limited.

Hartjes, L. (2016). Quantum technology: from research to application. In W. Schleich, K. Ranade, & C. e. Anton. Berlijn: Springer Verlag.

Heisenberg, W. (1958). *Physics and Philosophy.* Harper & Brothers: New York.

Hossenfelder, S. (2011, May 18). Testing Super-Deterministic Hidden Variables Theories. *Foundations of Physics Vol. 41*, pp. 1521–1531.

Hossenfelder, S. (2018, January 10). Screams for Explanation: Finetuning and Naturalness in the Foundations of Physics. *Synthese*, pp. 1-17.

Hossenfelder, S. (2019, September 27). *The Trouble with Many Worlds* . Retrieved from Backreaction: https://backreaction.blogspot.com/2019/09/the-trouble-with-many-worlds.html

Hossenfelder, S. (2020). *Lost in Math: How Beauty Leads Physics Astray.* New York: Basic Books.

Hossenfelder, S. (2020, October 3). *Superdeterminism: A Guide for the Perplexed.* Retrieved from ArXiv: https://arxiv.org/abs/2010.01324

Hossenfelder, S., & Palmer, T. (2020). Rethinking Superdeterminism. *Frontiers in Physics*, 8:139.

Hossenfelder, S., Palmer, T., & Price, H. (2022, May 17-20). *Superdeterminism and Retrocausality*. Retrieved from https://superdetermined.org/

Howard, D. (1985). Einstein on Locality and Separability. *Stud. Hist. Phil. Sci. Vol. 16, No. 3*, pp. 171-201.

Kaiser, D. (2004). The Postwar Suburbanization of American Physics. *American Quarterly, Vol. 56, No. 4,* 851-888.

Kaiser, D. I. (2022). Tackling Loopholes in Experimental Tests of Bell's Inequality. In D. I. Kaiser, *Oxford Handbook of the History of Quantum Interpretations* (pp. 339-370). Oxford: Oxford University Press.

Kuhlmann, M. (2020, August 10). *Quantum Field Theory*. Retrieved from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/quantum-field-theory/

Kumar, M. (2008). *Quantum: Einstein, Bohr and the Great Debate About the Nature of Reality.* New York: W. W. Norton & Company.

Landsman, K. (2015, May 21). The Fine-Tuning Argument. *ArXiv*, pp. 1-16.

Landsman, K. (2017). *Foundations of Quantum Theory.* Cham (Switzerland): Springer International Publishing.

Landsman, K. (2019, November 13). Randomness? What randomness? *arXiv:1908.07068v2 [physics.hist-ph]* , p. 32.

Landsman, K. (2021, March 24). Indeterminism and Undecidability. *arXiv:2003.03554v3 [math-ph]*, p. 24.

Landsman, K. (2021). The Fine-Tuning Argument: "What-if" Physics. *Lecture 'Geschiedenis en Grondslagen van de Natuurkunde' 9th of November* (p. 16). Nijmegen: Radboud University.

Maudlin, T. (2022, October 10). Tim Maudlin on Quantum Realism, Bell's Theorem, Pilot Wave, and Interpretations of Quantum Mechanics. (C. Jaimungal, Interviewer)

Mermin, N. D. (1981). Brining home the atomic world: Quantum mysteries for anybody. *American Journal of Physics 49,* pp. 940-943.

Muller, D. (2016, November 2). *Is This What Quantum Mechanics Looks Like?* Retrieved from Youtube: https://www.youtube.com/watch?v=WIyTZDHuarQ&t

Muller, F. (2014, June 21). Circumveiloped by Obscuritads. The nature of interpretation in quantum mechanics, hermeneutic circles and physical reality, with cameos of James Joyce and Jacques Derrida. *Festschrift for P.C. Suppes*, p. 26.

Myrvold, W., Genovese, M., & Shimony, A. (2019, March 13). *Bell's Theorem - 8.2: QM and Relativity*. Retrieved from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/bell-theorem/#QuanMechRela

Pais, A. (2000). *The Genius of Science: A portrait gallery of twentieth-century physicists.* New York: Oxford University Press.

Palmer, T. (2016, September 24). *p-adic Distance, Finite Precision and Emergent Superdeterminism: A Number-Theoretic Consistent-Histories Approach to Local Quantum Realism.* Retrieved from ArXiv: https://arxiv.org/abs/1609.08148

Peat, F. D. (1997). *Infinite Potential: The Life and Times of David Bohm.* New York: Basic Books.

Rauch et al., D. (2018, August 17). Cosmic Bell Test using Random Measurement Settings from High-Redshift Quasars. *Phys. Rev. Lett. 121, 080403*, p. 9.

Rizzi, A. (2018, November 2). Does the PBR Theorem Rule Out a Statistical Understanding of Quantum Mechanics? *Found. Phys. 48 No. 12*, pp. 1770-1793.

Schlosshauer, M., Kofler, J., & Zeilinger, A. (2013). A Snapshot of Foundational Attitudes Toward Quantum Mechanics. *Studies in History and Philosophy of Modern Physics Vol. 44,* 222-230.

Sen, I., & Valentini, A. (2020, December 2). Superdeterministic hidden-variables models I: nonequilibrium and signalling. *Proc. R. Soc. A. 476,* p. 12.

Sen, I., & Valentini, A. (2020, December 2). Superdeterministic hidden-variables models II: conspiracy. *Proceedings of the Royal Society A Vol. 476,* p. 11.

Shimony, A., Horne, M., & Clauser, J. (1976, October). Comment on "The Theory of Local Beables". *Epistemological Letters,* pp. 1-8.

Sivasundaram, S., & Nielsen, K. H. (2016, December 2). Surveying the Attitudes of Physicists Concerning Foundational Issues of Quantum Mechanics. *ArXiv*, pp. 1-20.

Slors, M., de Bruin, L., & Strijbos, D. (2015). *Philosophy of Mind, Brain and Behaviour.* Amsterdam: Boom uitgevers.

Spin Quantum Number. (2022). Retrieved from Chemistry Learner: https://www.chemistrylearner.com/spin-quantum-number.html

't Hooft, G. (2007, January 15). The Free-Will Postulate in Quantum Mechanics. *ITP-UU-07/4, SPIN-07/4*, pp. 1-9.

't Hooft, G. (2012, July 16). *Discreteness and Determinism in Superstrings.* Retrieved from ArXiv: https://arxiv.org/abs/1207.3612

't Hooft, G. (2013). The Fate of the Quantum. *Conference on Time and Matter* (pp. 1-14). Venice: ITP-UU-13/22, SPIN-13/15.

't Hooft, G. (2016). *The Cellular Automaton Interpretation of Quantum Mechanics.* Berlin: Springer.

't Hooft, G. (2017, September 19). *Free Will in the Theory of Everything*. Retrieved from ArXiv: https://arxiv.org/abs/1709.02874v2

Twice, T. (2017, September 30). *Double pendulum | Chaos | Butterfly effect | Computer simulation* . Retrieved from Youtube: https://www.youtube.com/watch?v=d0Z8wLLPNE0

User:Michael32710. (2020, May 1). *Solvay conference 1927 (group photograph)*. Retrieved from Wikimedia Commons: https://commons.wikimedia.org/wiki/File:Solvay_conference_1927_%28group_photograph%29_bw.jpg

Vaidman, L. (2021, August 5). *Many-Worlds Interpretation of Quantum Mechanics*. Retrieved from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/qm-manyworlds/

Weinberg, S. (1989, September). Testing quantum mechanics. *Annals of Physics Vol. 194, Issue 2*, pp. 336-386.

Werner, F. (1964). The Foundations of Quantum Mechanics. *Physics Today, 17 (1)*, pp. 53-60.

Westhoff, S. (2023). The Standard Model at its Limits. *The Standard Model and Beyond - Course for Master Students in Physics and Astronomy* (p. 3). Nijmegen: Radboud University.

Zeilinger, A. (2010). *Dance of the Photons.* New York: Farrar Straus & Giroux.

Zurek, W. (2003, June 19). Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys. 75, 715*.

Summary

Superdeterminism is a subclass of deterministic hidden-variable models underlying quantum theory. Crucially, it rejects that in all experimental contexts, the settings of a measuring device are independent of the state of the measured object. While relatively little-known, it has recently gained some traction. Proponents argue it could solve philosophical and foundational problems stuck to quantum theory, while opponents hold that superdeterminism may involve a ubiquitous conspiracy of nature, the invalidation of science and the absence of free will. I critically examine whether superdeterminism is a viable theoretical framework to address the philosophical problems of quantum theory with. While facing serious metaphysical challenges, I ultimately argue that although I do not consider it correct, it is viable given the defined criteria.