Studies in History and Philosophy of Modern Physics 72 (2020) 138–149

Contents lists available at ScienceDirect

Studies in History and Philosophy of Modern Physics

journal homepage: www.elsevier.com/locate/shpsb

Separating Einstein's separability

Sebastián Murgueitio Ramírez

History and Philosophy of Science, University of Notre Dame, 453, Geddes Hall, South Bend, IN 46556, United States

A R T I C L E I N F O

Article history:
Received 8 August 2019
Received in revised form
18 October 2020
Accepted 27 October 2020
Available online 12 November 2020

Keywords:
Separability
Incompleteness argument
Einstein
Quantum mechanics
Supervenience

A B S T R A C T

In this paper, I accomplish a conceptual task and a historical task. The conceptual task is to argue that (1) Einstein's Principle of Separability (henceforth "separability") is not a supervenience principle and that (2) separability and entanglement are compatible. I support (1) by showing that the conclusion of Einstein's incompleteness argument would still follow even if one assumes that the state of a composite system does not supervene on the states of the subsystems, and by showing that what Einstein says in "Quantum Mechanics and Reality" (1948) strongly suggests that separability is not a principle about how subsystem states relate to the state of composite systems. I support (2) by showing that if separability was incompatible with entanglement, then Einstein's argument would be incoherent in a trivial way. Thus, by arguing for (1) and (2) I directly challenge what has been, and still is, a very common reading of separability. The historical task is to offer the first detailed review of the different ways in which separability has been defined by physicists and philosophers in the last 60 years. Among other things, such a review distinguishes three different definitions of the principle, and shows that since the 1990s and up until the present date, it became standard to take separability (as presented by Einstein) to be a supervenience principle.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Few topics in the foundations of quantum mechanics have been given as much attention by physicists and philosophers as Einstein's incompleteness argument. Naturally, a good part of such attention has been dedicated to elucidating the structure of the argument (i.e., Howard (1985), Fine (1986), Ryckman (2017)), and this itself has led to discussions on the interpretation of the so-called "principle of separability" (henceforth "separability") introduced by Einstein therein. Understanding what separability amounts to is crucial because, as is well-established, it works as a premise in the argument (more on this in section 3.1).

After the 1980s, it became common to interpret separability as a supervenience principle, according to which states of the subsystems determine the states of the composite systems. Indeed, as I will show, many authors have connected what Einstein says about separability in "Quantum Mechanics and Reality" (1948) to Lewis' Humean Supervenience thesis, as if in this text Einstein was anticipating some version of this metaphysical doctrine (see section 2.3). Furthermore, this view of separability is often accompanied by the widespread belief that separability fails in quantum mechanics due to the occurrence of entanglement.

I will accomplish two main tasks, one which is primarily conceptual, and the other one which is mostly historical. The conceptual task is to argue, contrary to the common supervenience interpretation of Einstein's principle that emerged after the 1980s, that (1) separability, as used in Einstein's incompleteness argument, is not a supervenience principle, and that (2) separability is actually compatible with entanglement. In order to justify (1), I will show that the supervenience principle in question could not be a premise of Einstein's incompleteness argument because the conclusion of the argument follows even if the state of composite systems does not supervene on the states of the subsystems. Furthermore, some of the claims that Einstein makes in places such as "Quantum mechanics and Reality" (1948) and in his "Autobiographical Notes" (published in Schilpp's (1949)) strongly suggest that separability has nothing to do with how subsystem states relate to the states of composite systems. In order to justify (2), I will show that if separability and entanglement were incompatible, then Einstein's argument would be straightforwardly incoherent.

The historical task is to offer the first detailed review of the distinct ways that separability itself has been understood and defined by physicists and philosophers commenting on Einstein's arguments (for a review of the different ways Einstein himself

E-mail address: smurguei@nd.edu.

https://doi.org/10.1016/j.shpsb.2020.10.001
1355-2198/© 2020 Elsevier Ltd. All rights reserved.

understood separability, see Howard (1985)). I will be focusing, in particular, on the period between the time that the Bell inequalities were published in 1964, to, roughly, the present day. Among other things, such a review will reveal that there were at least four very different interpretations of separability in the 1970s, and that during this time, some authors attempted to distinguish separability from locality (this last point is particularly interesting because it is widely believed that Howard's (1985) paper was the first place where such a distinction was made). Furthermore, the review will help us see how the supervenience interpretation of separability that emerged in the late 1980s was quickly adopted by the community during the 1990s, and how that interpretation is still widely endorsed today.

The structure of the paper is as follows. In section 2, I present the historical review of the different uses of separability in the last 60 years. In section 3, by investigating Einstein's incompleteness argument, I will argue that separability is not a supervenience principle (and it does not entail one), that quantum phenomena are compatible with separability even if entanglement occurs, and that this is crucial to the coherence of Einstein's argument. Finally, in section 4, I raise and answer some possible objections.

1.1. Three preliminary remarks

Before delving into the main content of the paper, let me clarify three points. First, in this paper I will not be discussing the question of whether or not separability plays a role in the derivation of the Bell inequalities, or the related question of what formal definition of separability (if any) allows us to escape the implications of these inequalities (see Winsberg and Fine (2003), Fogel (2007) and Henson (2013) for these kinds of discussions). And related to this point, I will not be discussing the question of whether or not Humean supervenience is compatible with the presence of entanglement in the context of theories such as Bohmian mechanics (for some recent discussions around this issue, see Esfeld (2014), Bhogal (2017) and Bhogal and Perry (2017)). My goal here is to elucidate and discuss what Einstein took separability to be in the context of his incompleteness argument and of his "Quantum mechanics and Reality" paper in particular. Before trying to connect the principle with the literature on the Bell inequalities and with more recent discussions on the foundations and metaphysics of quantum mechanics, it is necessary to first understand the principle in its original context, as I intend to do with this paper.

Second, it is important to emphasize that this is not the first paper suggesting that separability is not a supervenience principle. For example, Ryckman (2017), Fine (2017) and even Henson (2013) do not read separability in terms of supervenience. However, besides its detailed historical review, what is novel about this paper compared to these other works is that it explicitly explains what the problems with a supervenience interpretation of separability are in the context of Einstein's incompleteness argument. In particular, in this paper I will offer several arguments showing why reading separability in terms of supervenience is problematic (section 3.2), and by doing so I challenge what has been, and is still, a common reading of separability (i.e. Maudlin (2011, 193).[1]

Third, someone might object that when I say that separability is not a supervenience principle, I have conflated two different notions of separability. Some scholars use the term to refer to a basic principle introduced by Einstein in his incompleteness argument, although the exact interpretation of that principle is a matter of dispute. And others −the objection goes− use "separability" to refer to some sort of supervenience principle (usually a principle connected to Lewis' idea of Humean Supervenience). Hence, to say that separability is not a supervenience principle is either misleading or simply wrong. Separability is not a supervenience principle if by "separability" one means the principle introduced by Einstein in his incompleteness argument; it is one (trivially) if one means the supervenience principle connected to Lewis' idea of Humean Supervenience.

The problem with this objection, however, is that it overlooks an important fact: after the 1990s, a significant number of authors in the foundations of quantum mechanics that use the term "separability" as a supervenience principle have claimed that Einstein introduced such a principle (or a very similar one) in his incompleteness argument (I show this in section 2.3). So by "separability" these authors are also referring to the principle introduced by Einstein −they just think that such a principle is a supervenience principle. Hence, when I say that I will show that separability is not a supervenience principle, I mean to say that, contrary to what many scholars have suggested, Einstein is not introducing a supervenience principle in his incompleteness argument. And this is a good opportunity to emphasize that when I use the terms "separability," "Einstein's separability" or "principle of separability," I am always referring to the principle introduced by Einstein in his incompleteness argument. Unless I note it, the scholars I will be engaging with in this paper are also referring to that principle.

2. Separability in the last 60 years

In this section, I will show that in the mid-1960s up until the mid-1970s, separability was taken to be a principle that was synonymous with locality or at least tightly related to locality. Then, during the late 1970s up until the early 1980s, it was treated as a general principle of isolation of systems that was identified with locality only in some cases. During the mid-1980s, it was (i) explicitly distinguished from locality and (ii) taken to be primarily about the possession of physical states of systems that are spatially apart. Finally, after the 1990s, it was taken to be a principle of supervenience according to which subsystem states determine the states of composite systems.

2.1. Separability from 1964 to 1979

During the late-1960s and early-1970s, the term "separability" was often used as a synonym of "locality" (in the Special Relativity sense[2]). For instance, in the introduction of the paper "On the Einstein-Podolsky-Rosen Paradox," after briefly referring to the concept of *elements of reality* from the EPR paper (1935), Bell says that "These additional variables [hidden variables] were to restore to the theory causality and locality" (1964, 195). Then he adds that "there have been attempts to show that even without such a separability or locality requirement no 'hidden variable' interpretation of quantum mechanics is possible" (1964, 195). Furthermore, in section "Locality and separability" of "On the problem of hidden variables in quantum mechanics," Bell says that "it would therefore

[1] In section 3.4 of his (2013) paper, Henson briefly points out why it seems inadequate to read separability as supervenience in the light of Einstein's own texts, and in that respect his project is similar to mine. However, given that he is more interested in understanding separability in the context of the Bell inequalities, he does not examine the incompleteness argument, which is my main focus here.

[2] Locality, just as separability, has been defined and understood in different ways (i.e., as the condition that there cannot be faster-than-light propagations, or the condition according to which information cannot be transmitted between space-like events, or the condition that events that are space-like separated are causally isolated from one another, etc.). For the purposes of my argument, it does not matter which of these definitions of locality is preferred.

be interesting … to pursue some further ‘impossibility proofs’, replacing the arbitrary axioms objected to above by some condition of locality or of separability of distant systems” (1966, 452). Given that in these articles Bell does not offer a definition of separability, but very carefully defines locality, it seems natural to assume that the reason he says “locality or separability” as opposed to just “locality” is that he is taking the terms to be interchangeable.

It is important to point out that, in these works, Bell's ideas about separability seem to be coming from both the EPR paper (as the title of the first paper reveals), and from Einstein's “Autobiographical Notes” (henceforth “AN”) found in Schilpp's *Albert Einstein, Philosopher Scientist* (1970). Not only does he cite the AN in both of these papers, but in the first paragraph of his 1964 paper, right after the first appearance of the term “locality,” Bell inserts a footnote that contains the following passage by Einstein (it will be convenient to refer to this passage by “P1”):

> If S_1 and S_2 are two systems that have interacted in the past but are now arbitrarily distant, the real, factual situation of system S_2 does not depend on what is done with system S_1, which is spatially separated from the former. (Einstein (1949, 85))

Here Bell seems to be reading P1 as saying that if two systems S_1 and S_2 are arbitrarily distant so that the measurement of one is space-like relative to the measurement of the other one, then, because of Special Relativity, one measurement should not affect the other.

Bell was not the only physicist using ‘locality’ and ‘separability’ as synonyms. About ten years later, Alain Aspect tells us that for other physicists “the words ‘locality’ and ‘separability’ … are sometimes taken as synonymous” (1976, footnote 9). For Aspect himself, the terms are not quite synonymous but they are tightly related to one another. In particular, consider Aspect's definition of what he calls “principle of separability” in that same paper,[3] which sounds pretty much like the definition of what we nowadays call “parameter independence”: “the setting of a measuring device at a certain time (event *A*) does not influence the result obtained with another measuring device (event *B*) if the event *B* is not in the forward light cone of event *A*” (1976, 14). By defining separability in terms of light cones, Aspect seems to be saying that separability is, essentially, a concrete version of locality. In particular, he seems to be taking separability to be a locality principle that is directly applicable to the experimental set-up required to test the Bell inequalities.

In 1975, just a year before Aspect's paper was published, Bernard d'Espagnat quotes P1 in order to introduce what he then calls “principle of separability” (1975, 1433). Importantly, however, in contrast with Bell and Aspect, d'Espagnat recognizes for the first time that separability and locality need not be the same thing for Einstein (this is important because it is usually assumed that Howard (1985) was the first scholar to explicitly distinguish these terms).[4] Immediately after quoting Einstein, he offers two readings of the principle, a broad one according to which no finite influences can propagate arbitrarily far away and a narrow one in which there is no space-like propagation of influences (1975, 1434). As I explain in greater detail now, only the narrow reading identifies locality and separability.

Although d'Espagnat does not say more about the broad reading in question, we can gain some insights on this point if we pay attention to what he says in the second edition of his *Conceptual Foundations of Quantum Mechanics* (1976). There, he does not refer to the principle with the name “principle of separability” nor does he offer two readings of it. Rather, he uses the term “Principle of the Separability of Mechanically Isolated Systems” and defines that principle in the following way:

> If a physical system remains during a certain time, mechanically (including electromagnetically, etc) isolated from other systems, then the evolution of its properties during this whole time interval cannot be influenced by operations carried out on other systems (1976, 81).

Notice here that separability is being understood in terms of isolation; separability is, according to this passage, the principle according to which the properties of isolated systems cannot be affected by other systems. Now, being in different space-like regions is a way that systems can be isolated from one another, and so it is a way by which, according to the principle in question, the properties of one system could not be influenced by the properties of another one. But this is not the only way of achieving isolation. In many circumstances in physics, if a system *A* is appropriately far away from other systems, yet is still (during some time) within their future light cones, it can be counted as being isolated. The reason is that the intensity of forces in our universe decays with distance, and so a given body *B* cannot significantly affect another body *A* that is far away (of course, how far away these bodies have to be from one another depends on the problem at hand and the degree of precision we are considering). Hence, we have good reasons to believe that d'Espagnat's narrow and broad readings in his 1975 paper correspond to these two ways by which systems can be isolated.

While discussing separability, other authors writing in the same period cite P1. Yet, they seem to follow d'Espagnat and not Bell in that they do not identify locality with separability. Let me give two examples. In 1976, M. Mugur-Schächter wrote an interesting piece commenting on d'Espagnat's (1975) article, in which he pointed out that the Bell inequalities not only suggest a conflict between locality and quantum mechanics but, more importantly to him, they also suggest a conflict between the idea or concept of a *physical system* and separability understood as an isolation condition (1976, 8). It is out of the scope of this paper to discuss such a purported conflict in more detail, but the key point is that Mugur-Schächter distinguished locality and separability. This is also clear from the way he described the goal of the paper: “The physical content of Bell's inequality is expressed directly in terms of isolation and separability, instead of locality” (1976, 1).

In 1978, Franco Selleri and Gino Tarozzi defined what they called “Einstein's locality” by means of the following passage: “the results of measurements on atomic systems are determined by “elements of reality” (sometimes called “hidden variables”), associated to the measured systems and/or to the measuring apparatus, which remain unaffected by measurements on other distant atomic systems” (Selleri & Tarozzi, 1978, 4). Despite the name, they notice that *Einstein's locality* is not locality in the relativity sense but rather *a more detailed version* of what, according to them, is called “the principle of physical separability,” according to which “all physical interactions become decreasingly important with increasing distance between the two considered bodies so that the properties of one object depend very little … on the properties and interactions of the other object, if the distance between them is large enough” (Selleri & Tarozzi, 1978, 2). They add that this principle “has played a fundamental role in the foundations and development of classical physics” and has been recognized by figures such as Bacon, Galileo,

[3] He says that this terminology comes from d'Espagnat' (see footnote 10 of the same paper).

[4] We will come back to Howard's paper later.

and Einstein (Selleri & Tarozzi, 1978, 2–3). As the reader probably noticed, *Einstein's locality* matches very well d'Espagnat's broad reading of separability, discussed above.

2.2. Separability during the 1980s

It should be clear then that at the end of the 1970s, some scholars working on the foundations of quantum mechanics started to follow d'Espagnat by distinguishing the concepts of locality and separability. However, their understanding of separability was still rather limited because they were paying too much attention to the EPR paper while ignoring other texts by Einstein where the structure of the incompleteness argument is clearer. This starts to change with Fine's (1981) paper (reprinted in *The Shaky Game* (1986)), in which he presents his discovery that in a letter to Schrödinger dated June 19th of 1935, Einstein reveals that he does not like the structure and argumentation of the EPR paper and that for reasons of language, the paper was written by Podolsky (Fine (1986, 35)). Not only is the incompleteness argument found in this letter clearer than the one found in the EPR paper,[5] but it is also here where Einstein introduces the separability principle ("*Trennungsprinzip*") for the first time. In the letter, Einstein says that because of Trennungsprinzip, "Now the real state of *B* cannot depend on what sort of measurement I undertake on *A* [where *A* and *B* are arbitrarily distant]" (translated by Fine (1986, 50)).[6]

Reading this passage as well as the whole letter, gives the impression that separability is logically connected to locality. For example, it follows from what Einstein is saying here that if the measurement on *A* is space-like relative to the region where *B* is at a given time, then the state of *B* during that time cannot be sensitive to the kind of measurement that we carry out on *A*. And to say this is simply to restate the locality condition according to which there could not be faster-than-light propagations. This explains why Fine's general definition of the separability principle in these works, which he names "Einstein-locality," expresses a particular type of locality:

> Einstein-Locality: The real state of one system is not immediately influenced by the kinds of measurements directly made on a second system, which is sufficiently spatially separated from the first (Fine (1986), 61, 63).

For the purposes of this paper, the most important points to take away from this particular work of Fine are his discoveries that (1) the EPR paper is not the most adequate place to look for Einstein's incompleteness argument, and that (2) the incompleteness argument is meant to show a conflict between separability and the completeness of quantum mechanics that has nothing to do with the existence of incompatible observables nor with the uncertainty principle (as the EPR paper seems to suggest).

Fine is certainly right to say that the incompleteness argument purports to show such a conflict between separability and the completeness of quantum mechanics, yet his understanding of separability as a kind of locality is not adequate in the light of the paper "Quantum Mechanics and reality" (QR henceforth) or in the light of Einstein's AN, both of which are some of Einstein's latest texts on the subject. And Fine is aware of this problem because in his "The Einstein-Podolsky-Rosen Argument in Quantum Theory," he carefully distinguishes locality from separability (2017).

Don Howard's paper "Einstein on Locality and Separability" (1985) is the first place where we find a thorough discussion of QR and an examination of the ways in which, in this and other texts, Einstein distinguishes locality from separability.[7] In particular, Howard notes that in the incompleteness argument presented in AN, Einstein says that there are two ways of blocking his incompleteness argument, one is to reject locality and the *other one* is to reject separability (1985, 186). And he also notes that in the following (now infamous) passage from QR, this distinction appears again:

> Further, it appears to be essential for this arrangement of the things introduced in physics that, at a specific time, these things claim an existence independent of one another, insofar as these things 'lie in different parts of space'. Without such an assumption of the mutually independent existence (the 'being-thus') of spatially distant things, an assumption which originates in everyday thought, physical thought in the sense familiar to us would not be possible. Nor does one see how physical laws could be formulated and tested without such a clean separation.
>
> For the relative independence of spatially distant things (A and B), this idea is characteristic: an external influence on A has no immediate effect on B; this is known as the 'principle of local action', which is applied consistently only in field theory. The complete suspension of this basic principle would make impossible the idea of the existence of (quasi-) closed systems and, thereby, the establishment of empirically testable laws in the sense familiar to us (Einstein (1948, 320); translated by Howard (1985, 187–188)).

There are at least three important points that we can draw from this passage (the second one also explicitly acknowledged by Howard). First of all, based on the first paragraph, notice that whatever our preferred definition of separability is, it should capture or make precise the sense in which (a) separability is a principle having to do with the independent existence of spatially apart objects; (b) it originates in everyday thought; (c) it is crucial to "physical thought" in the sense familiar to us; and (d) without it physical laws could not be formulated and tested (we will come back to these points in section 3.2).

Secondly, notice that separability does not seem to be the same thing as locality for at least three reasons. First, the definition of separability, in terms of "the mutually independent existence of spatially distant things" does not seem to be a definition of locality (for one, it makes no reference to light cones, or space-like events, etc). Second, notice that there is a subtle difference between the lines in the passage where Einstein explains why rejecting locality is problematic for doing physics and those where he states that separability is crucial for doing physics. While in the first paragraph he says that the assumption of "the mutually independent existence of spatially distant things" originates in everyday thought and without that assumption physical thought would not be

[5] For example, in the letter, it is clear that the reference made in the EPR paper to incompatible observables and to the uncertainty principle are inessential (for more on this point, see Fine (1986) and Howard (1985)).

[6] As we will see in section 4, Healey argues that in this letter Einstein is taking separability as a supervenience principle.

[7] Recall from section 2.1 that d'Espagnat had already proposed that there is a reading in which separability is not the same thing as locality. But in contrast to Howard, d'Espagnat did not point out that Einstein himself made such a clear distinction.

possible, in the second paragraph he says that locality is a requisite for the "the existence of (quasi-) closed systems and, thereby, the establishment of empirically testable laws" (notice that neither locality (in the Special Relativity sense) nor the existence of closed systems seem to be ideas originated in everyday thought). And third, as Howard (1985, 188) points out, a natural way of reading Einstein when saying that locality is *characteristic* of separability is by saying that he is actually distinguishing these principles (after all, it would be strange to say that a locality principle is characteristic of itself).

So, let's follow Howard and accept, as it is now commonly understood, that locality and separability are distinguished by Einstein. The obvious question then is "what is separability?" Based on QR and in the AN, Howard goes on to offer the following definition:

> Einstein's separability: spatially separated systems always possess separate real states (1985, 173).

Notice that this definition distinguishes separability and locality because it is one thing to assert that spatially separated systems always have states (separability, as here understood) and quite another to say that the state of one system cannot be affected by the state of another system if the systems are in space-like regions (locality). Furthermore, notice that separability is more basic than locality in the sense that for there to be a local (non-local) process, there have to be systems in different regions of space (or space-time) that are changing their states, and so they have to have states to begin with.

I will argue in section 3.1 that a very slight variation of this definition is actually the best definition of separability in the literature, but somewhat ironically, this is not Howard's preferred and only definition.[8] In 1989, inspired by discussions of the Bell inequalities of scholars such as Jon Jarrett (1984) and Paul Teller (1986), Howard modifies his original definition by adding a second condition: he says that two systems are separable if

1. Each possess distinct real states.
2. The joint state of the two systems is wholly determined by these separate states (Howard (1986, 226)).

Interestingly, ever since then a good number of scholars seem to have followed Howard in shifting their attention from EPR (and other texts) to QR and understanding the separability principle as a thesis having to do with the determination of the properties of the composite state by those of the spatially separated parts (this is condition (2) in the above definition).[9] In the next section, I will consider several representative examples of some scholars who follow Howard in this respect.[10] But before going to the next section, I invite the reader to see Table 1 for a quick summary of the different uses of separability discussed so far.

2.3. Separability after the 90s

In his paper "Holism and nonseparability," Healey offers a thorough discussion of different senses of holism and separability in the light of physical theories. But what matters for our purposes is that he reads the separability principle proposed by Einstein in QR as a supervenience principle. He says that

> A physical process is said to be spatiotemporally separable in R (where R is a space-time region) if and only if it is supervenient upon an assignment of qualitative, intrinsic physical properties space-time points in R. *Spatiotemporal separability*: Any physical process occurring in a region R of space-time is spatiotemporally separable in R (1991, 406).

A couple of lines later, he adds that "he [Einstein] himself formulated principles closely related to the principle of spatiotemporal separability" (1991, 407) and goes on to quote the passage from QR that we already discussed in section 2.2.

Consider now the following passage by Maudlin from the third edition of his popular "Quantum Non-Locality and Relativity" book:

> Einstein's world-view held that each region of space-time has its own intrinsic physical state, and that the entire and complete physical state of the universe is specified once one has determined the intrinsic state of each small region. This ontological doctrine may be called 'separability' (2011, 193).

When saying that "each region of space-time has its own intrinsic physical state" Maudlin seems to be talking of condition (1) of Howard's (1989) definition, and when saying that "the entire and complete physical state of the universe is specified once one has determined the intrinsic state of each small region" he is talking of (a more general version of) condition (2).

Let me mention in passing that in the same book, Maudlin explicitly argues against Howard's suggestion of understanding separability in terms of *outcome independence* (roughly, in the context of an EPR experiment, outcome independence is the assumption that the probabilities for outcomes in one wing are unaltered by the outcomes in the other wing) (Maudlin (2011, 89)). However, even if he disagrees with Howard in this respect (and he is not the only one, see footnote 9), Maudlin does follow him in taking separability as involving a supervenience thesis, as it is clear from the passage cited above. Or to give another example, when presenting a situation showing that separability is not equivalent to outcome independence, Maudlin says that (my own emphasis) "This is a completely separable theory: the photons and the tachyon all have perfectly determinate intrinsic states at all times, *and the joint physical state of two distinct regions or systems is just the sum of their individual states*" (Maudlin (2011, 90)). Again, we see Maudlin endorsing conditions (1) and (2) of Howard's definition.

Furthermore, in his book *The Metaphysics Within Physics*, Maudlin defines "separability" as "the complete physical state of the world is determined by (supervenes on) the intrinsic physical state of each spacetime point … and the spatio-temporal relations

[8] His preferred definition is the one offered in his 1989 paper, that I discuss next (personal conversations). However, based on his (1996) and (2017) works, one gets the impression that sometimes Howard is working with two different notions of separability. One is the notion defined in his 1985 paper (that I just explained) and which he takes to be a basic principle of individuation of objects. The other notion is the one he defines in his 1989 paper, and that he ties to some sort of supervenience thesis.

[9] If one pays attention to footnote 38 of Howard's (1985) paper, one can realize that, despite the definition offered in that paper, Howard was already entertaining the thought that separability is (or entails) the thesis that the state of the system is given by the product of the states of the subsystems (1985, 191).

[10] Let me mention in passing that there has been controversy about whether separability, as defined by Howard here, is indeed equivalent to *factorizability*, as Howard himself suggests (1989, 226) (*factorizability* is, roughly, the thesis that the joint distribution for measurements on a composite system equals the product of the distributions of measurements on the individual subsystems). See, for example, Winsberg and Fine (2003) and Fogel (2007) for two opposite views around this question. A discussion of these more technical disagreements on the issue of how to formalize the concept of separability is out of the scope of this paper, but let me point out that even Fine, Winsberg, and Fogel seem to follow Howard in assuming that separability is (or entails) a principle about the determination of the state of the composite system by the state of the subsystems (see Winsberg and Fine (2003, 86–87) and Fogel (2007, 922)).

Table 1
A summary of some of the different ways in which separability has been understood in the last 60 years.

	1964–1974	1974–1980	1980–1984	1984–1989	1989-present
Representative figure(s)	Bell Aspect	d'Espagnat Selleri Tarozzi	Fine	Howard	Howard Healey Maudlin
Exegetical locus Classicus	EPR, AN	EPR, AN	Einstein's letter to Schrödinger (June 19th of 1935), AN	QR, AN	QR
What does Einstein mean by separability?	Locality	Isolation	Locality	That spatially separated systems always have real states	That the states of subsystems determine the states of composite systems

between those points" (2007, 51). Importantly, in both of these books Maudlin uses QR in order to argue that what he is calling "separability" is precisely the principle Einstein is introducing in this text.

Michael Esfeld's "Quantum Entanglement and a metaphysics of relations" is yet another place where we find the separability principle presented by Einstein in QR formulated in terms of supervenience. Esfeld says that "Albert Einstein based his criticism of quantum theory on the principle of separability" (Esfeld (2004, 608)). Then, after briefly discussing Howard's (1989) definition, Esfeld goes on to formulate the following principle (which he takes to be an improvement over Howard's (1989) definition):

> Non-separability: The states of two or more systems are non-separable if and only if it is only the joint state of the whole that completely determines the state-dependent properties of each system and the correlations among these systems (to the extent that these are determined at all). According to this characterization, any case of quantum entanglement is a case of non-separability, and non-separability is the reason why quantum entanglement is a sort of holism (2004, 608).

In this definition of non-separability, Esfeld makes no reference to separation in space or space-time because, for him, what is essential to the principle is that the state of the whole is prior to (or determines) the state of the parts (and this is true independently of whether or not these parts are separated in space or space-time or in a different more abstract space). One might argue, against Esfeld, that reference to space-time is crucial to Einstein because the fields of classical field theories are defined on space-time, and according to Einstein, these theories are paradigmatic examples of theories satisfying separability. But in any case, even if Einstein would not approve of Esfeld's definition of non-separability, the important point here is that Esfeld takes QR as a place where Einstein puts forward some sort of supervenience principle.

Another example of an author who seems to associate separability with supervenience is Wayne Myrvold (2011). In his paper, Myrvold offers an insightful study of the contrast between the non-separability found in quantum mechanics and the one found in electromagnetism (manifested, for example, in the Aharonov-Bohm effect). But Myrvold also implies that the separability principle is (or motivates) a supervenience principle. After citing some fragments of QR where Einstein introduces the separability principle, Myrvold goes on to explain the concept of quantum non-separability in the following terms: "If two quantum systems *A*, *B*, are in an entangled state, the reduced state of *A* and *B* … do not suffice to determine the state of the combined system" (Myrvold (2011, 420)). Although he does not explicitly attribute a supervenience principle to Einstein, by defining quantum non-separability in terms of supervenience, and by introducing quantum non-separability just after citing some fragments from the QR paper, Myrvold is clearly implying that the principle introduced by Einstein in QR strongly motivates a supervenience principle.

Let me add here that in the texts by Healey, Maudlin, Esfeld and Myrvold just discussed, all the authors also discuss Lewis' doctrine of Humean Supervenience, as if this doctrine was tightly related to Einstein's separability principle (i.e., Maudlin goes as far as saying that in QR, Einstein is endorsing (some version of) Lewis'doctrine (Maudlin (2007), ch 2)). This further stresses my point that these scholars take it to be natural to read separability as a supervenience principle.

Even more recently, in the *ontological models* literature in quantum foundations, one finds authors formalizing the separability principle presented in QR in terms of Cartesian products between subsystem states. In particular, Harrigan and Spekkens say that "An ontological model is separable if the ontic space [read "hidden variables"] of a region *R* is given by the Cartesian product of the ontic space of the subregions" (2010) and Leifer says that:

> Implicit in this [concept of separability found in Einstein] is the idea that there are no inherently global joint properties of the composite system that are not determined by the properties of the individual systems. In the language of ontological models, this means that ontic state spaces [for product states] should compose according to the Cartesian product (Leifer (2014, 104)).

In other words, if there is a composite ontic state λ_T that cannot be written as a Cartesian product of the ontic states of the subsystems (i.e., $\lambda_T \neq \lambda_1 \times \lambda_2$), then that indicates that the ontic state of the composite system is not determined by the ontic state of the parts, and so this state does not satisfy the separability principle (as understood by Leifer).

Incidentally, that the state of the composite system (of a product state) is given by the Cartesian product of the states of the subsystems is an assumption required in the proof of the PBR theorem (i.e., Leifer (2014, 101)), and so this is an interesting case where the reading of separability as a supervenience principle plays an important role in the justification of a recent theorem in the foundations of quantum mechanics.

I have thus shown evidence of how, after 1989 and up until very recent years, it became rather common for scholars working on the foundations of quantum mechanics to read the separability principle (presented by Einstein in QR) as some sort of supervenience principle.[11] I will now argue that this is an inadequate understanding of the principle and that we should go back to Howard's (1985) definition. Let me mention in passing that in section 4.1, I

[11] See Timpson and Brown (2002, 7), Henson (2013), Fine (2017) and Ryckman (2017, Ch 4) for four authors who do not seem to link separability to supervenience. These authors, however, do not explain what would fail in the incompleteness argument if one were to read separability in terms of supervenience, which is precisely what I will do next.

will also offer an explanation for why it became natural to understand the separability principle presented by Einstein with a supervenience principle, despite the fact that, as I will show here, separability is not a supervenience principle.

3. Separability in the incompleteness argument

3.1. The structure of the incompleteness argument

Given that he never defined the term very explicitly, the best way of figuring out what Einstein meant by 'separability' is to find out what role the principle plays in his argument that quantum mechanics is an incomplete theory. In particular, we should pay attention to the role that the principle plays in some of the later versions of Einstein's incompleteness argument such as the one found in QR or the one found in the AN because it is here where the separability principle is presented more explicitly (as Howard (1985), Fine (2017) and Ryckman (2017, 145) have pointed out). In what follows, I will focus on the version of the argument found in the AN (which is a very similar version to the one found in QR).[12] In contrast to other discussions of the argument (i.e., the ones in Howard (1985), Fine (2017) or Ryckman (2017, Ch 4)), mine will focus on an issue that has not been addressed by previous presentations, namely, explaining in detail why we could not make sense of the argument if we were to take separability to be some sort of supervenience principle. But in order to appreciate the problems of the supervenience interpretation, I need to start by presenting the structure of the incompleteness argument itself.

As its name indicates, the incompleteness argument intends to show that quantum mechanics is an incomplete theory. For Einstein, quantum mechanics is incomplete in case the wave-function "is not an exhaustive description of the real situation of the system" (Schilpp (1949), 83). For example, if particles always have a definite position even before a position measurement is made, then one can say that quantum mechanics is incomplete in that there are no states of the theory (no pure quantum states) representing these definitive positions of particles *before* a measurement takes place (there is nothing special about position, the same point applies to any other property such as spin, momentum, etc). On the other hand, if the individual system has no definitive position value before a measurement, but only acquires one after a measurement is performed, then the (position) wave-function is indeed a complete description of the real state of the system because there is a definite wave-function corresponding to the system after the measurement takes place (Schilpp (1949, 83–85)). Hence, in order to show that quantum mechanics is incomplete, Einstein needs to show that the states of the theory (i.e. the wave-functions) do not offer a complete description of the real state of systems. And to do that, he attempts to show that wave-functions corresponding to different pure states (henceforth "incompatible wave-functions") stand in a many-to-one relation to the physical state of the system (these wave-functions are incompatible precisely because they represent different measurement statistics). If incompatible states of the theory are all assigned to the same physical state of the system, then the states of the theory are not adequately tracking what is going on with the physical states of the system. In Einstein's words, "it would be impossible that two different types of ψ functions could be coordinated with the identical factual situation of S_2" (Schilpp (1949, 87)). For Einstein, this suggests that one should adopt the view that the wave-function merely gives us information about *ensembles* of systems (as opposed to giving us exhaustive information about the individual systems).

Here, then, is the argument (I will follow to a large extent Howard's (1985) presentation). Consider a typical EPR case of two spatially separated systems, S_1 and S_2, that interacted in the past so that they are now entangled. Standard quantum mechanics tell us that the wave function ψ_2 of S_2 will be sensitive to the kind of measurement performed on S_1.[13] Given this set-up, Einstein says:

> But now the real state of S_2 must be independent of what happens to S_1. Thus, different wave-functions can be found (depending on the choice of measurement on S_1) for the same real state of S_2. (One can only avoid this conclusion either by assuming that the measurement on S_1 changes (telepathically) the real state of S_2, or by generally denying independent real states to things which are spatially separated from one another. Both alternatives appear to me entirely unacceptable).
>
> If now the physicists A and B accept this reasoning as sound, then B will have to give up his position that the ψ-function is a complete description of a real situation. For in this case it would be impossible that two different types of ψ-functions could be correlated with the same situation (of S_2) (Translated by Howard (1985, 186)).

From this very short argument, it follows that quantum mechanics is incomplete (in the sense mentioned above): since the real state of S_2 (call that state "λ") is described or represented (at the same time) by multiple pure wave-functions, then these wave-functions cannot be offering exhaustive descriptions of the real state of the system. That is, the theory is incomplete in that different pure states of the theory (states corresponding to different measurement statistics) are all mapped into the same real state λ of a system.

At this point, a clarification is in order. We know that after the measurement on S_1 has taken place, one can claim that S_2 has a well-defined state which is adequately represented by a certain wave-function. The key point for Einstein is that if after the measurement on S_1 we can attribute a wave-function ψ to S_2, then one should be able to attribute ψ to S_2 even an instant before the measurement on S_1 occurs or even if no measurement is performed on S_1 at all (and the same is true for any other possible wave-function corresponding to any other possible outcome).[14] This is due to the fact that what happens with S_1 at time t should not affect (given locality) the state of S_2 at time t nor at an instant earlier (while entanglement still holds). In short, then, what worries Einstein is not just what happens after the measurement of one of the systems, but rather the fact that even an instant before any measurement on S_1 takes place, incompatible wave-functions should be assigned to S_1 (due to locality).

Now, although it is widely accepted that the separability principle is a premise of the incompleteness argument, we will need to do some work to extract a precise definition of the principle, as I will show here. In the parenthetical remarks of the first paragraph

[12] For a detailed study of the differences and similarities between Einstein's different versions of the incompleteness argument, I refer the reader to Howard (1985). Einstein was not always very clear about exactly what the separability principle is, and how it is different from locality. This is true in both the June 19th, 1935 letter to Schrödinger and in his "Physics and Reality" (1936) paper.

[13] Here Einstein is assuming Quantum Mechanics as found in standard physics textbooks, which explicitly appeals to the Collapse Postulate.

[14] Einstein says: "It follows that every statement about S_2 which we arrive at as a result of a complete measurement of S_1 has to be valid for the system S_2, even if no measurement whatsoever is carried out on S_1" (Einstein et al. (1971, 172)).

just cited, Einstein tells us that to block his argument we must either 1) defend that "the measurement on S_1 changes (telepathically) the real state of S_2" or we 2) deny "independent real states to things which are spatially separated from one another." Einstein does not explain in any detail how 1) or 2) would block his incompleteness argument, so I will explain that next.

Defending 1) amounts to denying *locality* because we allow that the measurement performed on S_1 brings about an instantaneous change on the state λ of S_2. If these instantaneous changes are possible, we no longer have a case where a single state λ is being described by two or more incompatible wave functions but a case where different states of S_2 (produced by different measurements on S_1), say λ_m and λ_p, are described by different wave functions. Thus, we no longer get a many-to-one relation between (incompatible) wave-functions and the states of S_2.

Let us see now how 2) blocks the incompleteness argument. One can "deny independent real states to things which are spatially separated from one another" by asserting one of two things (or both): 2.a) sometimes objects which are spatially separated from one another have states that are not independent; and 2.b) sometimes, things which are spatially separated from one another do not have states. For the case of two objects S_1 and S_2, 2.b) amounts to saying that sometimes, when S_1 and S_2 are spatially separated, S_1 lacks real states or (inclusive) S_2 lacks real states. Notice that 2.a) sounds a lot like the denial of locality; if sometimes the states of two objects are not independent, then, presumably, changing the state of one will change the state of the other one. If these changes of states are such that they can occur when the measurements are *space-like*, then asserting 2.a) or asserting the denial of locality amounts to essentially the same thing. Given that he takes 1) and 2) to be different ways of blocking his argument, and given that 2.a) seems to collapse into 1), it is then natural to assume that by 2) Einstein had in mind 2.b).[15] In short, then, one can block the argument by asserting that

(CLAIM) Sometimes, when S_1 and S_2 are spatially separated, S_1 lacks real states or (inclusive) S_2 lacks real states.

Einstein is saying, then, that to block his argument one must either reject locality or endorse CLAIM. It follows then that CLAIM must be the negation of separability (assuming, as Howard (1985) and Fine (2017) do, that 2.b) expresses the negation of separability). It follows from elementary logic that, for the case of two systems, separability must be the following principle (whose negation is CLAIM):

Separability (SEP): if S_1 is spatially separated from S_2, then S_1 has a physical state and S_2 has a physical state.[16]

We have thus extracted, quite transparently I think, a definition of separability (SEP) directly from the incompleteness argument. And notice, by the way, that this definition is essentially Howard's (1985) definition presented in section 2.2 (and the one defended in Fine (2017) and Ryckman (2017, ch 4)).

Given the understanding of separability just presented, the incompleteness argument can be summarized as follows:

1. Spatially separated subsystems S_1 and S_2 always have (real) states (separability).
2. Measurements on S_1 do not instantaneously affect the state of S_2 (locality).
3. If quantum mechanics attributes with certainty quantum state ψ_2 to S_2 *just* after a measurement on S_1 takes place, and if the measurement on S_1 could not affect the state of S_1 at that point in time (locality), then quantum mechanics must attribute ψ_2 to S_2 some time before the measurement on S_1 takes place, or even if no measurement on S_1 occurs.
4. The previous premise is true of any arbitrary quantum state ψ_2 that quantum mechanics could attribute to S_2 after any of the multiple types of measurements that can be performed on S_1.
5. Thus, before any measurement on S_1 is even performed (or even if no measurement is performed), quantum mechanics must simultaneously attribute multiple pure quantum states to the same (real) state of S_2.
6. The wave-function does not offer a complete description of the real states.

C. C. Quantum mechanics is incomplete.

Let's see now how CLAIM, or, equivalently, the negation of SEP, allows one to avoid the incompleteness argument.

To defend CLAIM in this context is to say that sometimes (in cases of entanglement), spatially separated systems do not have states and so it does not matter, or it does not matter in the same way, that one can attribute different inconsistent wave-functions to the (non-existent) states of a system based on measurements on the other system. According to someone defending CLAIM, intending to attribute wave functions to a system that is part of an entangled state is simply a mistake, a misuse of quantum theory. Hence, Einstein's incompleteness argument fails in the sense that it does not show that multiple wave functions are all being mapped into the same real state λ of S_2 simply because there is no such a state. In other words, SEP allows Einstein to talk about the real states of S_1 and S_2 at any time while they are separated, even before a measurement on S1 takes place (and so even during the time they are entangled, although Einstein did not use the term 'entanglement' here). And since he is aware that someone might try to resist SEP by suggesting that subsystems do not always have states, Einstein feels obligated to argue in favor of separability in places such as QR or in the AN.

3.2. Separability is not about supervenience

So far I have argued for the positive thesis that understanding separability as SEP makes sense of the incompleteness argument. However, at this point, someone can suggest that there are other ways of making sense of the argument according to which separability is indeed a supervenience thesis (after all, as we saw in section 2.3, many authors believe that separability is some sort of supervenience principle). In this section, I will offer four different reasons to support the negative thesis that separability cannot be taken to be a supervenience thesis.

First and most importantly, the conclusion of the incompleteness argument would still follow even if the (real) states of S_1 and S_2 fail at determining the (real) state of the composite system S_T (that is, even if a standard supervenience thesis fails). As long as one presupposes locality and that subsystems always have states, Einstein's conclusion that the theory assigns mutually incompatible quantum states to the same real state follows, and this is so even if the supervenience thesis just mentioned fails (see the previous section, where the derivation of this conclusion made no assumptions at all about the truth of this supervenience thesis). Naturally, if the conclusion of the argument still follows when the supervenience thesis in question fails, then such a thesis cannot be a premise of the argument. And since separability is a premise of the

[15] For a good discussion of the prospects and limitations of different ways of making sense of the probabilities gathered from EPR-type experiments, see Pitowsky (1994).

[16] In the definition I omit any mention of the term "independence" because it can mislead us to think that Einstein is talking of something like locality. That is, by stressing the concept of independence, we might think that all Einstein is saying is that one state cannot affect the other state if the measurements in question are space-like (notice that when defining separability, Howard (1985) and Fine (2017) also avoid mentions of the term "independence").

argument, then separability cannot be identified with such a supervenience thesis (nor can it be taken to entail one).[17]

Second, on the face of it, to talk of the "mutually independent existence of spatially distant things," as he does in QR, is not to talk of how the state of a composite system is (or is not) completely determined by the states of the subsystems. And indeed, nowhere in QR does Einstein explicitly mention the state of a composite system when he is talking of the separability principle! In particular, in multiple occasions he talks of "the independent existence of the real state of affairs existing in two separate parts of space" or of "the independent existence of the physical reality present in different parts of space" (Einstein, Born and Born (1971, 172)), yet he does not talk of the given relation between the states of the subsystems and the states of the composite system. On the face of it, what is common to all of these locutions is clearly not something having to do with how the states of parts (or subsystems) relate to the states of wholes (or composite systems), but rather Einstein's emphasis is that objects or systems separated in space must have their own (real) states, as SEP explicitly says.[18]

Third, recall that in QR Einstein says this: "Without such an assumption of the mutually independent existence … of spatially distant things, an assumption which originates in everyday thought, physical thought in the sense familiar to us would not be possible." Now, it is not clear why physical thought "in the sense familiar to us" would not be possible if sometimes the state of the whole was not determined by the state of the parts. That is, it is not clear why the possibility of, for example, emergent properties would make physical thought impossible in any interesting sense. To be sure, emergent properties might be puzzling (and maybe more so for a fundamental theory such as quantum mechanics), but to claim that they would make physical thought so unfamiliar, and that we would not be able to formulate and test physical laws (as Einstein says in QR) just sounds a bit like an exaggeration.

In contrast, imagine that SEP actually fails. That would mean, from the definition of SEP, that sometimes, one (or both) of the spatially separated systems lacks a physical state. And that is certainly something surprising, for it is a rather basic physical intuition (and a common sensical one) that physical systems always have physical states.[19] What would make them "physical" if they did not?! And, furthermore, were we to reject SEP, we wouldn't be able to formulate physical laws, for how would we write a physical law for an object that lacks a physical state? Recall Einstein's words when talking about the rejection of separability: "nor one does see how physical laws could be formulated." It is then rather easy to see that Einstein would have been very puzzled about a theory that rejects SEP, that is, a theory in which physical systems lack physical states and yet (somehow) posits physical laws for these systems.[20]

Fourth and finally, recall that in QR Einstein says that locality is characteristic of separability. Hence, if separability was indeed a supervenience principle, it would follow that locality is characteristic of supervenience. But it is not clear at all what it would mean to say something such as "the determination relation that obtains between parts and wholes is mediated by the speed of light." To illustrate why this is problematic, consider the macroscopic state of a chair that (presumably) supervenes on the state of the atoms making it up. Although one can make perfect sense of what it means to say that the states of the atoms in the chair only affect one another locally, it is harder to understand what it means to say that the states of the atoms determine the macroscopic state of the chair in a local manner. For one, it is not clear what the relevant events are. Where is the macroscopic state of the chair at a given time *t*, and where is the state of the atoms at that same time? Hopefully this is enough to illustrate why saying that locality is characteristic of supervenience relations is, at the very least, a controversial claim. In contrast, the sense in which locality is characteristic of SEP is rather natural: it is characteristic of the theories that ascribe states to spatially separated subsystems (i.e. classical field theories) that these states cannot causally affect each other in a superluminal way.

Thus, for these four reasons, reading separability as a supervenience principle is quite problematic.

3.3. Entanglement and separability are compatible

Not only is separability not a supervenience thesis but, as I will now explain, it is also a (common) mistake to think that separability is incompatible with entanglement. In particular, it is a mistake to think that if there is entanglement, then separability fails (or is false).

Suppose for the sake of the argument that the following conditional is true: if there is entanglement, then separability is false. Then, by the contrapositive, the following thesis is true: if separability is true, then there is no entanglement. But separability is an assumption of Einstein's incompleteness argument, and so it follows from this that Einstein is assuming in his argument against quantum mechanics that entanglement cannot occur. But if that is true, then Einstein is extremely incoherent and his argument is trivially problematic for he would be assuming that entanglement cannot happen at the same time he uses entanglement (i.e. he considers subsystems that have previously interacted, and hence are entangled) in order to derive the conclusion that quantum mechanics is incomplete. That is, Einstein would be assuming that quantum mechanics is false in order to conclude that it is incomplete! Needless to say, it is rather implausible that someone like Einstein would commit such an elementary mistake.

What we just said is actually very general in that it does not depend on the particular definition of separability we are working with (as far as we agree that separability is an assumption of Einstein's argument). Whatever our definition of separability is, if it turns out that such a definition is falsified by the occurrence of entanglement, then this strongly suggests that our definition doesn't adequately capture what Einstein meant by that term. And it follows from this (maybe surprisingly to some) that given an appropriate definition of separability, quantum systems can be held to satisfy the separability principle at the ontic level (e.g., as described by a purportedly more fundamental theory such as Bohmian mechanics) despite the occurrence of entanglement at the level of description of quantum mechanics. And incidentally, notice that if one thought that entanglement and supervenience were incompatible (i.e., Maudlin (2007, ch. 2)), then this would be a further argument against the view that separability is a supervenience thesis (for separability is compatible with entanglement).

Something like a corollary follows from this: by "real physical state" Einstein cannot mean a state adequately represented by the wave-function, for then separability would be trivially false in cases

[17] In contrast, as it was explained in the previous section, the conclusion of the argument would not follow if SEP was rejected, which strongly indicates that SEP is indeed a premise of the argument.

[18] The only place where Einstein, very briefly, seems to talk of a supervenience principle, is in his June 19th, 1935 letter to Scrödinger. We will discuss this point in section 4.

[19] And some plausible philosophical principles seem to suggest this as well: if one defines a state to be a (non-null) subset of the physical properties of a system (i.e., pressure, volume and temperature in thermodynamics, position and momentum in classical mechanics, etc.) and we define physical systems as systems that have one or more physical properties, then, necessarily, physical systems have states.

[20] See Timpson and Brown (2002) for an insightful discussion around this point.

of entanglement (because during entanglement, all we can do is attribute a wave-function to the composite system, not to the spatially separated subsystems).[21] In other words, if separability was a thesis demanding that spatially separated systems always have states adequately represented by the wave-function (or by pure quantum states, or by rays in Hilbert space), then separability would automatically be a principle incompatible with quantum mechanics, and so Einstein's incompleteness argument would be unsuccessful in a rather naive way.[22]

4. Objections

I will now consider three objections.

4.1. First objection

The first objection is that it is almost a contradiction in terms to say that quantum mechanics can satisfy the separability principle because quantum mechanics is non-separable.

No doubt, it is strange and confusing to say that a non-separable theory such as quantum mechanics can satisfy the separability principle, and so it is helpful to say one or two things about why this is so. The confusion is easy to understand and resolve: despite appearances, "separability" in the expression "non-separability" does not refer to the separability principle. In particular, "non-separability" means that (some) composite quantum states do not factorize in terms of the quantum states of the subsystems, but the separability principle is not about the factorizability of vector states but about spatially apart subsystems having states (see section 3.1). And so it follows that a theory can be such that subsystems always have (real or "ontic") states yet the quantum states of composite systems do not factorize in terms of the quantum states of the subsystems. That is, despite appearances, a theory can be both non-separable and satisfy the separability principle introduced by Einstein.

Here it can be helpful to make a very brief reference to Bohmian mechanics. For notice that if one adopts Bohmian mechanics and one takes positions and velocities to specify the (real) physical state of particles, then one finds that entangled (and hence non-separable) but spatially separated particles always have states, and so one finds that separability (SEP) holds during entanglement.

4.2. Second objection

The second objection is that there might be some textual evidence of Einstein saying or implying that separability is indeed some sort of supervenience principle. In particular, Healey thinks that the separability principle ("*Trennungsprinzip*") formulated by Einstein in his letter written to Schrödinger in June 19th of 1935 is indeed a supervenience principle. Healey says (my emphasis):

> *Spatial separability*: The qualitative, intrinsic physical properties of a compound system are supervenient on the qualitative, intrinsic physical properties of its spatially separated component systems together with the spatial relations among these component systems. I think *we can recognize the spatial separability principle as a natural generalization of Einstein's formulation* (1991, 410).

When he says "Einstein's formulation" in the last sentence, Healey is referring to the following passage of the letter in question: "The real state of *AB* consists precisely of the real state of *A* and the real state of *B*, which states have nothing to do with one another" (I will use "P2" to refer to this passage). In "Holism and nonseparability in physics," Healey refers again to P2 as a way of defining what he calls there "Real State Separability Principle" (2016).

I will point out to two problems with this reading by Healey. The first and main problem is that if one reads the letter in question, one sees that Einstein is not intending to define separability (or "*Trennungsprinzip*") by means of P2. Rather, in the letter Einstein says that the 'separation principle' is the thesis that "the second box [system], along with everything having to do with its contents, is independent of what happens with regard to the first box (separated partial systems)" and that (emphasis in the original one).

> After the collision, the real state of (*AB*) consists precisely of the real state of *A* and the real state of *B*, which two states have nothing to do with one another. *The real state of B thus cannot depend upon the kind of measurement I carry out on A* ('Separation hypothesis' from above.) (translated by Howard (1985, 180)).

Given Einstein's own emphasis, it is clear that the 'Separation hypothesis' explains why "the real state of B thus cannot depend upon the kind of measurement I carry out on A," for *A* and *B* are spatially separated systems. And for the same reason, it is not accurate to say, as Healey suggests, that the 'Separation hypothesis' is the thesis expressed by P2.[23]

The second problem is that one does not find a sentence like P2 in AN nor in QR, which are the places where Einstein offers the clearest versions of the incompleteness argument (as argued by, for example, Howard (1985) and Fine (2017)). Indeed, as I say in the second main point of section 3.2, nowhere in QR does Einstein explicitly mention the state of composite systems when talking of the separability principle. Thus, the textual evidence linking separability to a supervenience principle is very scarce, to say the least.

4.3. Third objection

Let me end by considering a different kind of objection. The objection I am considering can be formulated as a question: why, if separability was appropriately understood in terms of SEP, have so many scholars linked separability to supervenience while also taking entanglement as a sign that separability fails?

Unsurprisingly, I think that the answer to this question is that there is indeed an interesting way by which separability leads to a supervenience principle, and an interesting way by which entanglement can be taken to falsify such a principle. However, as I will now explain, separability only leads to such a supervenience principle if one also endorses locality and makes further assumptions.

Suppose that we endorse both SEP and locality. Then, it follows from Einstein's incompleteness argument that there is more to the

[21] As Fine puts it, for Einstein the "real states" are not given by quantum observables (Fine (1986), 61).

[22] It is important to not confuse my claim here according to which separability (understood as SEP) is compatible with entanglement with the claim that separability understood as supervenience is compatible with entanglement. The latter claim has been the matter of rather recent and interesting debates in the metaphysics of science literature (i.e., Esfeld (2014), Bhogal (2017), Miller (201), Bhogal and Perry (2017)).

[23] Healey seems to recognize this problem with his interpretation in footnote 16 of his (1991, 410) paper.

quantum state of particles than what the wave-function tells us. That is, it follows that there are additional "hidden" variables that offer a more complete description of the state of quantum objects than the one achieved by the wave-function. So, there is (at least in principle) a theory that would be able to appeal to these variables in order to explain quantum phenomena. Now, given that any adequate theory that appeals to hidden variables has to be able to reproduce the predictions of quantum mechanics, such a theory should be able to reproduce, in particular, the predictions of entangled systems.

A very natural thing to do with a hidden variable theory is to try to recover the predictions of entangled systems (which are composite systems) by appealing to local interactions between the subsystems and by appealing to the hidden variables of the subsystems in question (as opposed to, for example, appealing to hidden variables for the composite system). That is, a very natural thing to do when endorsing a hidden variable theory is to endorse:

(SUPERV): The statistics associated with composite systems (such as entangled systems) are completely recovered from the hidden states of the subsystems.

(LOC): Only local interactions between these subsystems are possible.

Thus, given locality and SEP (and Einstein's argument), it is natural to think that there should be a theory that respects (SUPERV) and (LOC) and that reproduces the predictions of quantum mechanics.

As it turns out, one of the main lessons from the violation of the Bell inequalities is that one cannot account for the correlations of EPR-experiments simply by focusing on the (hidden) states of the subsystems *and* by considering local interactions between these subsystems.[24] That is, the lesson is that there cannot be a theory that endorses (SUPERV) and (LOC). Therefore, one might suspect that something has to be wrong with the assumptions of Einstein's incompleteness argument as this argument led us to believe that there could be theories endorsing (or assuming) (SUPERV) and (LOC). In particular, one might think that Einstein cannot be right about how crucial separability and locality are for physical thought (for if these were essential to physics, then one would expect that a theory endorsing something like (SUPERV) and (LOC) was possible).

This is, I think, a (very simplified) version of what actually happened: after the 1980s, scholars reading Einstein's incompleteness argument were reading it from the perspective of those who had very good reasons to believe in locality (given the success of Special Relativity) and so who had very good reasons to believe that theories endorsing (SUPERV) are impossible (because the conjunction of (SUPERV) and (LOC) was impossible according to Bell's results, and (LOC) was taken to be true). Thus, it was natural to take the apparent impossibility of theories endorsing (SUPERV) to be evidence against the premises of Einstein's argument (for if the premises hold, then quantum mechanics is incomplete and a theory endorsing (SUPERV) is expected). And given that they were not willing to reject locality, scholars in the 1980s and 1990s took the failure (or impossibility) of theories endorsing (SUPERV) to suggest a failure of one of the other premises of Einstein's argument, namely, separability. That is, they took the (apparent) failure of (SUPERV) to indicate a failure of separability.

This would explain not only why it became common to associate the failure of (SUPERV) with a failure of separability, but also why it became common to take entanglement and separability to be incompatible with one another (for entanglement is what the Bell inequalities exploit in order to show that theories endorsing (SUPERV) seem to fail). *Notice, however, that the Bell inequalities do not force us to reject (SUPERV), for we could also reject (LOC) (i.e., as Bohmian mechanics does).* And if we do reject (LOC), we can perfectly save the separability principle. After all, Einstein can be right about separability and its importance to physics and yet be wrong about locality and its importance to physics.

5. Conclusion

In this paper, I started by presenting a detailed review of the different ways the concept of separability has been defined and explained in the last 60 years. From this review it comes to light that after the 1990s, it became common to define separability (as presented in QR) as a supervenience principle. Then, I defended two main theses: if one goes back to the incompleteness argument, it becomes clear that separability cannot be (and does not entail) a supervenience principle, and that separability is not incompatible with entanglement.

Author statement

Sebastián Murgueitio Ramírez: Conception and design of study, Drafting the manuscript, Approval of the version of the manuscript to be published.

Acknowledgments

This paper owes a lot to Don Howard, whose work on Einstein, as well as his classes on the subject, taught me a lot about the history and philosophy of quantum mechanics. This work was presented in the Philosophy of Science Association Biennial Meeting at Seattle (2018), and in the Philosophy of Logic, Mathematics, and Physics Graduate Conference at Western University (2019). The participants' feedback at both of these events helped me improve upon my argument. I also want to thank Paddy Blanchette, Katherine Brading, Harvey Brown, Michael Cuffaro, John Dougherty, Jamee Elder, James Nguyen, Mousa Mohammadian, Pablo Ruiz de Olano, Jeremy Steeger, and the two anonymous referees for their feedback. Finally, I want to thank Geoff Hall, Ellen Lehet, and Stacy Sivinski for their editing help regarding various versions of this paper.

References

d'Espagnat, B. (1975, March). Use of inequalities for the experimental test of a general conception of the foundation of microphysics. *Physical Review D, 11*, 1424–1435.
d'Espagnat, B. (1976). *Conceptual foundations of quantum mechanics*. Mathematical physics monograph series. W. A. Benjamin, Advanced Book Program.
Aspect, A. (1976, Oct). Proposed experiment to test the nonseparability of quantum mechanics. *Physical Review D, 14*, 1944–1951.
Bell, J. (1964, Nov). On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika, 1*, 195–200.
Bell, J. S. (1966, Jul). On the problem of hidden variables in quantum mechanics. *Reviews of Modern Physics, 38*, 447–452.
Bhogal, H. (2017). Minimal anti-humeanism. *Australasian Journal of Philosophy, 95*(3), 447–460.
Bhogal, H., & Perry, Z. (2017). What the humean should say about entanglement. *Noûs, 51*(1), 74–94.
Einstein, A. (1948). Quanten-Mechanik und Wirklichkeit. *Dialectica, 2*(3/4), 320–324.

[24] Of course, there are always additional assumptions one might be willing to reject in order to account for the EPR correlations in a local way, although their rejection is not popular. To give one example, one could deny that events in the future cannot affect events in the past.

Einstein, A. (1949). Autobiographical notes. In P. Schilp (Ed.), *Albert Einstein, philosopher-scientist* (1st ed., Vol. 7). Evanston, Ill: Library of Living Philosophers.

Einstein, A., Born, M., & Born, H. (1971). *The Born-Einstein letters: Correspondence between Albert Einstein and Max and Hedwig Born from 1916-1955, with commentaries by Max Born*. Macmillan.

Einstein, A., Podolsky, B., & Rosen, N. (1935). Can quantum-mechanical description of physical reality Be considered complete? *Physical Review, 47*(10), 777–780.

Esfeld, M. (2004). Quantum entanglement and a metaphysics of relations. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics, 35*(4), 601–617.

Esfeld, M. (2014). Quantum humeanism, or: Physicalism without properties. *The Philosophical Quarterly, 64*(256), 453–470.

Fine, A. (1981). Einstein's critique of quantum theory: The roots and significance of EPR. In P. Barker, & C. G. Shugart (Eds.), *After Einstein: Proceedings of the Einstein centennial celebration at Memphis state University: 14-16 March 1979*. Memphis State University Press.

Fine, A. (1986). *The Shaky game. Science and its conceptual foundations series*. University of Chicago Press.

Fine, A. (2017). The Einstein-Podolsky-Rosen argument in quantum theory. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy (Winter 2017 ed.)*. Metaphysics Research Lab, Stanford University.

Fogel, B. (2007). Formalizing the separability condition in Bell's theorem. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics, 38*(4), 920–937.

Harrigan, N., & Spekkens, R. W. (2010). Einstein, incompleteness, and the epistemic view of quantum states. *Foundations of Physics, 40*(2), 125–157.

Healey, R. A. (1991). Holism and nonseparability. *The Journal of Philosophy, 88*(8), 393–421.

Healey, R. (2016). Holism and nonseparability in physics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of philosophy (Spring 2016 ed.)*. Metaphysics Research Lab, Stanford University.

Henson, J. (2013, Aug). Non-separability does not relieve the problem of Bell's theorem. *Foundations of Physics, 43*(8), 1008–1038.

Howard, D. (1985). Einstein on locality and separability. *Studies In History and Philosophy of Science Part A, 16*(3), 171–201.

Howard, D. A. (2017). Einstein's philosophy of science. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of philosophy (Fall 2017 ed.)*. Metaphysics Research Lab, Stanford University.

Howard, D. (1989). Holism, separability, and the metaphysical implications of the Bell experiments. In J. T. Cushing, & E. McMullin (Eds.), *Philosophical consequences of quantum theory* (pp. 224–253). University of Notre Dame Press.

Howard, D. (1996). A peek behind the Veil of Maya: Einstein, Schopenhauer, and the historical background of the conception of space as a ground for the individuation of physical systems. In J. Earman, & J. Norton (Eds.), *The cosmos of science: Essays of exploration, Pittsburgh-Konstanz series in the philosophy and history of science* (pp. 87–150). Pittsburgh, PA: University of Pittsburgh Press.

Jarrett, J. P. (1984). On the physical significance of the locality conditions in the Bell arguments. *Noûs, 18*(4), 569–589.

Leifer, M. S. (2014). Is the quantum state real? *An Extended Review of ψ-ontology Theorems, 3*(1), 89.

Maudlin, T. (2007). *The metaphysics within physics*. Oxford: Oxford University Press.

Maudlin, T. (2011). *Quantum non-locality and relativity: Metaphysical intimations of modern physics*. Wiley.

Miller, E. (2014). Quantum entanglement, Bohmian mechanics, and humean supervenience. *Australasian Journal of Philosophy, 92*(3), 567–583.

Mugur-Schächter, M. (1976, March). What is at stake in the experiments on Bell's Inequality? *Epistemological Letters, 9*, 1–10.

Myrvold, W. C. (2011). Nonseparability, classical, and quantum. *The British Journal for the Philosophy of Science, 62*(2), 417–432.

Pitowsky, I. (1994). George Boole's 'conditions of possible experience' and the quantum puzzle. *The British Journal for the Philosophy of Science, 45*(1), 95.

Ryckman, T. (2017). *Einstein*. Routledge.

Schilpp, P. A. (1949). *Albert Einstein, Philosopher-Scientist* (1st ed., Vol. 7). Evanston, Ill: Library of Living Philosophers.

Schilpp, P. A. (1970). *Albert Einstein, Philosopher-Scientist* (4th ed., Vol. 7). New York, NY: MJF Books: Library of Living Philosophers.

Selleri, F., & Tarozzi, G. (1978, October). Extension of the domain of validity of Bell's inequality. *Epistemological Letters, 21*, 1–22.

Teller, P. (1986). Relational Holism and quantum mechanics. *The British Journal for the Philosophy of Science, 37*(1), 71–81.

Timpson, C. G., & Brown, H. R. (2002, December). Entanglement and relativity. In R. Lupacchini, & V. Fano (Eds.), *Understanding physical knowledge*. CLUEB: University of Bologna.

Winsberg, E., & Fine, A. (2003). Quantum life: Interaction, entanglement, and separation. *Journal of Philosophy, 100*(2), 80–97.