Dirac and the dispensability of mathematics

Otávio Bueno

Department of Philosophy, University of South Carolina, Columbia, SC 29208, USA

Received 31 May 2004; received in revised form 13 March 2005; accepted 30 March 2005

Abstract

In this paper, I examine the role of the delta function in Dirac’s formulation of quantum mechanics (QM), and I discuss, more generally, the role of mathematics in theory construction. It has been argued that mathematical theories play an indispensable role in physics, particularly in QM [Colyvan, M. (2001). The indispensability of mathematics. Oxford University Press: Oxford]. As I argue here, at least in the case of the delta function, Dirac was very clear about its dispensability. I first discuss the significance of the delta function in Dirac’s work, and explore the strategy that he devised to overcome its use. I then argue that even if mathematical theories turned out to be indispensable, this wouldn’t justify the commitment to the existence of mathematical entities. In fact, even in successful uses of mathematics, such as in Dirac’s discovery of antimatter, there’s no need to believe in the existence of the corresponding mathematical entities. An interesting picture about the application of mathematics emerges from a careful examination of Dirac’s work.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Dirac; Delta function; Indispensability argument; Quantum mechanics; Antimatter; Application of mathematics

1. Introduction

Among the various contributions made by Paul Dirac, one, in particular, is fascinating for the plethora of philosophical issues it raises: his introduction of the delta function in 1930 (see Dirac, 1958¹). In a time when the physics community was deeply
concerned about establishing the equivalence between matrix and wave mechanics (see Muller, 1997), Dirac provided the first successful attempt to prove the equivalence result. But Dirac’s proposal relied, in an important way, on the delta function.

As is well known, the delta function has peculiar properties (among them, the fact that it’s inconsistent!), and so it’s not surprising that Dirac would be suspicious about its use. (To avoid the use of this function, von Neumann would later introduce the Hilbert space formalism; see von Neumann, 1932.) Strictly speaking, the delta function is not even a function. And Dirac was, of course, aware of this. He called it an “improper function”, since it doesn’t have a definite value for each point in its domain (Dirac, 1958, p. 58). Moreover, as Dirac also points out, the delta function is ultimately dispensable. After all, it’s “possible to rewrite the theory (i.e. quantum mechanics) in a form in which the improper functions appear all through only in integrands. One could then eliminate the improper functions altogether” (Dirac, 1958, p. 59).

In this paper, I examine the role of the delta function in Dirac’s formulation of quantum mechanics (QM), which provides an interesting opportunity to examine, more generally, the role of mathematics in theory construction. It has been argued that mathematical theories play an indispensable role in physics, particularly in QM, given that it’s not possible even to express the relevant physical principles without the use of mathematics. And given that mathematical entities are indispensable to our best theories of the world, the argument goes, we ought to believe in their existence (for a thorough defense of this view, see Colyvan, 2001). As I argue here, at least in the case of the delta function, Dirac was very clear about its dispensability.

After discussing the significance of the delta function in Dirac’s work, and exploring the strategy that he devised to overcome its use, I examine the importance of this strategy to the use of mathematics in physics. I then argue that even if mathematical theories were indispensable, this wouldn’t justify the commitment to the existence of mathematical entities. To illustrate this point, I examine an additional and particularly successful use of mathematics by Dirac: the one that eventually led to the discovery of antimatter. As we will see, even in this case, there’s no reason to believe in the existence of the corresponding mathematical entities. An interesting new picture about the application of mathematics emerges from the careful examination of Dirac’s work.

2. Dirac and the delta function

2.1. Introducing the delta function

Dirac’s (1958) work on the foundations of QM is one in a series of exceptional treatises produced on the subject between the late 1920s and the early 1930s,

---

2Schrödinger attempted to prove the equivalence before, but as he acknowledged, only one direction of the equivalence was established (see Schrödinger, 1926, and Muller, 1997 for a discussion).

3For example, the delta function entails that differential operators are integral operators, but this is never the case (for an illuminating discussion of this point, see von Neumann, 1932, pp. 17–27; especially pp. 23–27).
including Weyl (1928), Wigner (1931), and von Neumann (1932). What distinguishes Dirac’s book, particularly since its second edition, is its informal, concise style, and the explicit attempt to “keep the physics to the forefront”, examining “the physical meaning underlying the formalism wherever possible” (Dirac, 1958, p. viii). That Weyl, Wigner and von Neumann had very definite and distinct theoretical agendas is clear enough. Weyl used group-theoretic techniques to provide a systematic foundation for QM (Weyl, 1928). Wigner also explored group theory, but his concern focused in particular on the applications of quantum theory (Wigner, 1931). Von Neumann, in turn, introduced a completely different conceptual framework—in terms of Hilbert spaces—to provide a mathematically coherent formulation of QM, one in which the probabilistic character of the theory naturally emerged from the formalism (von Neumann, 1932).

Dirac tried something different. Just as Weyl and Wigner, he explored group-theoretic techniques; just as von Neumann, he introduced new representations. But as opposed to Weyl, Wigner and von Neumman, with Dirac mathematics was never in the forefront. He used mathematics, to be sure. But, he insisted, “the mathematics is only a tool and one should learn to hold the physical ideas in one’s mind without reference to the mathematical form” (Dirac, 1958, p. viii). One of the goals of this paper is to understand what exactly this means.

In the first few chapters of The principles of quantum mechanics, Dirac provides an algebraic framework to formulate the fundamental laws of QM (see Dirac, 1958, pp. 1–52). The framework is developed in terms of the now familiar bra vectors, ket vectors, and linear operators.4 Although the framework has a number of attractive features, to be able to solve certain problems, it’s useful to have an additional representation in terms of numbers. As Dirac notes:

For some purposes it is more convenient to replace the abstract quantities by sets of numbers with analogous mathematical properties and to work in terms of these sets of numbers. The procedure is similar to using coordinates in geometry, and has the advantage of giving one greater mathematical power for the solving of particular problems. (Dirac, 1958, p. 53)

Using this numerical representation, Dirac is then able to establish four main results (Dirac, 1958, p. 57):

(a) The basic bras of an orthogonal representation are simultaneously eigenbras of a complete set of commuting observables.
(b) Given a complete set of commuting observables, there is an orthogonal representation in which the basic bras are simultaneously eigenbras of this complete set.

4The introduction of bra and ket vectors is actually an innovation of the second edition of The Principles of Quantum Mechanics (see Dirac, 1958). This innovation was preserved in all subsequent editions. In many ways, the most substantial revisions in the Principles were made in the second edition, which adopted a less formal style than the one developed in the first edition—a feature that many people praised, including Heisenberg.
(c) Any set of commuting observables can be made into a complete commuting set by adding certain observables to it.

(d) A convenient way of labeling the basic bras of an orthogonal representation is by means of the eigenvalues of the complete set of commuting observables of which the basic bras are simultaneously eigenbras.

For the purposes of this paper, the details of this representation are not important. What is important is to note, as Dirac does, the representation’s power and generality. However, despite these features, the representation is still limited. How could we characterize the lengths of the basic vectors invoked throughout the discussion? As is well known, if we were considering only orthogonal representations, it would be natural to normalize the basic vectors. The trouble, however, is that vectors can be normalized only if the parameters that label them have discrete values. If the parameters have continuous values, it’s impossible to normalize the vectors. As a result, in this case the lengths of the vectors cannot be defined. Dirac makes the point very clearly:

We have not yet considered the lengths of the basic vectors. With an orthogonal representation, the natural thing to do is to normalize the basic vectors, rather than leave their lengths arbitrary [...]. However, it is possible to normalize them only if the parameters which label them all take on discrete values. If any of these parameters are continuous variables that can take on all values in a range, the basic vectors are eigenvectors of some observable belonging to eigenvalues in a range and are of infinite length [...]. (Dirac, 1958, pp. 57–58)

Given this difficulty, an entirely new approach is required. Dirac is explicit about what is demanded:

Some other procedure is then needed to fix the numerical factors by which the basic vectors may be multiplied. To get a convenient method of handling this question, a new mathematical notation is required [...]. (Dirac, 1958, p. 58; italics added.)

This “convenient method” and “new mathematical notation” is the delta function.

Before examining the details of the function, note the way in which Dirac describes the latter. It is characterized as a “convenient method”, which highlights the function's pragmatic character. Moreover, the function is only thought of as a “mathematical notation”, which downplays any ontological significance the function may have. As we will see shortly, Dirac has a justification for this way of describing the delta function — or, at least, he tries to provide such a justification, which has to do with the strategy he devised to dispense with the function altogether.

But what exactly is the delta function? As Dirac points out, it is a quantity \( \delta(x) \) that depends on a parameter \( x \) and satisfies the following conditions:

\[
\int_{-\infty}^{+\infty} \delta(x) \, dx = 1,
\]

\[
\delta(x) = 0 \text{ for } x \neq 0.
\]
In order to “visualize” the behavior of the delta function, consider a function that is identical with 0 at every point except 0, and in the immediate neighborhood of 0, although the value of the function is not exactly defined, it is so large that its integral adds up to 1. A function behaving in this way would be a delta function. In Dirac’s own words:

To get a picture of $\delta(x)$, take a function of the real variable $x$ which vanishes everywhere except inside a small domain, of length $\varepsilon$ say, surrounding the origin $x = 0$, and which is so large inside this domain that its integral over this domain is unity. (Dirac, 1958, p. 58)

Now, it’s perfectly natural to ask whether any object satisfies the conditions above. And Dirac immediately acknowledges that there is something peculiar about the delta function. First, it is not exactly a function, given that at $x = 0$, it doesn’t have a sharply defined value—only its integral has a value.

$\delta(x)$ is not a function of $x$ according to the usual mathematical definition of a function, which requires a function to have a definite value for each point in its domain, but is something more general, which we may call an ‘improper function’ to show up its difference from a function defined by the usual definition. (Dirac, 1958, p. 58)

Second, if at $x = 0$ the delta “function” lacks a sharply defined value, in what sense is its integral defined at all? Although Dirac doesn’t explicitly address this second question, clearly there is something delicate going on with the “function”. As a result, the use of the latter has to be restricted. Dirac is very clear about this point:

$\delta(x)$ is not a quantity which can be generally used in mathematical analysis like an ordinary function, but its use must be confined to certain simple types of expression for which it is obvious that no inconsistency can arise. (Dirac, 1958, p. 58)

In other words, a strategy to delimit the function’s appropriate scope needs to be found. And by determining the function’s scope, it’s then possible to dispense with it altogether.

2.2. Dispensing with the delta function

Given the worries mentioned above about the delta function, it comes as no surprise that Dirac provides a strategy to dispense with the latter. Although, as I will argue below, there’s more going on here than simply a worry about the peculiar behavior of the function, it’s worth being clear about Dirac’s dispensability strategy.

The first step in the strategy consists in identifying two significant properties of the delta function. These properties provide algebraic rules to manipulate the latter, and they motivate the function’s dispensability. The first property crucially depends on
the fact that the integral of the delta function is 1 in the neighborhood of 0. As Dirac
points out:

The most important property of \( \delta(x) \) is exemplified by the following equation
\[
\int_{-\infty}^{+\infty} f(x)\delta(x)\,dx = f(0),
\]
where \( f(x) \) is any continuous function of \( x \). We can easily see the validity of this
equation from the above picture of \( \delta(x) \). The left-hand side of (3) can depend only
on the values of \( f(x) \) very close to the origin, so that we may replace \( f(x) \) by its
value at the origin, \( f(0) \), without essential error. (Dirac, 1958, p. 59)

In other words, if \( f \) is a well-behaved, continuous function, with definite values for
each argument in \( f \)'s domain, the value of the integral in the left-hand side of (3)
ultimately only depends on the value of \( f \) at 0. After all, in this context the integral of
the delta function is 1.

The second property of the delta function also stresses an additional algebraic
feature. Dirac formulates it in the following way:

By making a change of origin in (3), we can deduce the formula
\[
\int_{-\infty}^{+\infty} f(x)\delta(x - a)\,dx = f(a),
\]
where \( a \) is any real number. Thus the process of multiplying a function of \( x \) by
\( \delta(x-a) \) and integrating over all \( x \) is equivalent to the process of substituting \( a \) for
\( x \). (Dirac, 1958, p. 59)

The same move discussed above in connection with the first property of the function
is also found in the second property. Again, the crucial feature of the function — the
fact that its integral is 1, when values close to 0 are considered—is used to obtain the
second property.

Given these two properties, the strategy to dispense with the delta function is not
hard to figure out. As long as the delta function occurs in an integral, it can be
eliminated. As Dirac notes:

Although an improper function does not itself have a well-defined value, when it
occurs as a factor in an integrand the integral has a well-defined value. In quantum
theory, whenever an improper function appears, it will be something which is to be
used ultimately in an integrand. Therefore it should be possible to rewrite the theory in
a form in which the improper functions appear all through only in integrands. One could
then eliminate the improper functions altogether. (Dirac, 1958, p. 59; italics added.)

In other words, as long as the use of the delta function is restricted to a factor in an
integrand, it becomes clear that the function is dispensable. After all, due to the two
properties discussed above, if the delta function occurs in an integral, it can be
completely eliminated. Dirac’s conclusion is then clear:

The use of improper functions thus does not involve any lack of rigor in the
theory, but is merely a convenient notation, enabling us to express in a concise form
certain relations which we could, if necessary, rewrite in a form not involving improper functions, but only in a cumbersome way which would tend to obscure the argument. (Dirac, 1958, p. 59; italics added.)

With these remarks, Dirac highlights the pragmatic character of the delta function. It is a “convenient notation”, useful to “express in a concise form” certain relations. And it’s only in a “cumbersome way”, which “would tend to obscure the argument”, that we could express—without invoking the function — the relations we want to express. Although ultimately dispensable, the function is no doubt useful. But the function’s usefulness is a pragmatic matter, which confers no ontological character to it.

2.3. The status of the delta function

It might be argued that the dispensability of the delta function paves the way for the function’s replacement by something else. In fact, the argument goes, with hindsight someone can indeed maintain that the delta function is not really a function. Not only in the sense acknowledged by Dirac (i.e., the delta “function” doesn’t have a definite value for each point in its domain), but also in a deeper sense. Given the “reformulation” of the delta “function” in Schwartz’s theory of distributions, the “function” is actually a distribution (see Colyvan, 2001, pp. 103–104, note 20). This also explains why the delta function is dispensable: it is not the relevant sort of object. What is needed to achieve consistently what Dirac aimed to achieve is not a function, but a distribution.

Is this response justified? Well, not quite. The suggestion that the delta function is ultimately a distribution misses a significant point about Dirac’s strategy; namely, the pragmatic role that the function plays in describing—in an elegant and simple way—the relevant relations. In fact, the introduction of the theory of distributions increases hugely the complexity of quantum theory as well as the size of the function space for QM. Not surprisingly, introducing distributions ends up “obscuring the argument”, in pretty much the way Dirac warned. There’s something to be said for Dirac’s strategy after all.

Moreover, any identification of the delta “function” with a distribution cannot be right. Even an improper function is not a distribution. To stipulate that they are the same objects based on the role they play in the respective theories (Dirac’s on the one hand, Schwartz’s on the other) is to assume that the objects in question behave in the...
same way. But they don’t—and they can’t. If Dirac’s formulation is simpler and inconsistent, and Schwartz’s is consistent but much more complex, there obviously is a significant difference between the two theories and their corresponding objects. No literal identification of these objects can be made.

Of course, this is not an argument to adopt Dirac’s formulation. It is simply an argument that highlights the difficulty of absorbing Dirac’s theory within Schwartz’s. The relation between the two theories is much more complex than a simple identification of their ontologies may suggest. So, Dirac does have a point when he insists that, although ultimately dispensable, the delta function helps to express the relevant relations “in a concise form”. And even though we could rewrite QM without the delta function, this would be achieved “only in a cumbersome way which would tend to obscure the argument” (Dirac, 1958, p. 59).

Now, if failing to use the delta function tends “to obscure the argument”, does it mean that the function has an explanatory role? I don’t think so. The function does allow us to define the length of the basic vectors even when we can’t normalize them (because the parameters that label the vectors take on continuous values). But it’s not clear to me that the function explains why this definition is possible. As Dirac notes, the delta function only provides a useful notation that allows the definition to be carried out—a notation that is ultimately dispensable, and so it’s unclear whether it plays any explanatory role here.7

But two additional worries could be raised at this point. (a) Given that the delta “function” is admittedly not a function in the strict sense, how can Dirac know that using it within an integral yields something meaningful? We have good reason to believe that functions are well behaved in the context of integrals. But clearly, we don’t have any such reasons with regard to improper functions. Dirac can only hope that improper functions behave as functions in the context of an integral. But do they really behave in this way?

(b) Moreover, the real puzzling feature about the delta function is not (only) that it lacks a sharply defined value at some point in its domain and yet has an integral over the whole domain. The problem, rather, is that the non-zero integral seems to arise from the undefined point. How can the integral of any function be non-zero over a single point?8

These are fair concerns, and I think that to try to alleviate them, at least in part, Dirac indicated how the delta function could be dispensed with altogether. As a result, the function ultimately wouldn’t be needed. Of course, this doesn’t completely dispel the worries, given that, with regard to (a), only functions and numbers can be meaningfully used in an integral. If the delta function is neither of them, how can it be so used? In Dirac’s view, despite not being a function in its complete generality, the delta “function” behaves like a function under most circumstances—that’s the

---

7Similarly, I don’t think that the delta function plays an explanatory role when it’s used to establish the equivalence between matrix and wave mechanics. Ultimately, the function only provides a notational device to relate differential and integral operators. And as von Neumann showed, the use of the function in this context is clearly dispensable (see von Neumann, 1932).

8I owe this second worry to Mark Colyvan (personal correspondence).
reason why he highlighted the two properties of the function discussed above (see Eqs. (3) and (4)). These properties yield the intended results “without essential error” (Dirac, 1958, p. 59)—again, a significant pragmatic move made by Dirac. There certainly is the possibility of error—and Dirac openly admits that. But given Eqs. (3) and (4), that possibility is substantially restricted, allowing the results to be obtained without significant trouble.

Similarly, with regard to (b), Dirac’s idea is that the delta function “vanishes everywhere except inside a small domain, of length \( \varepsilon \) say, surrounding the origin \( x = 0 \)” (Dirac, 1958, p. 58). However, the value of the function “is so large inside this domain that its integral over this domain is unity” (ibid.). So, strictly speaking, the non-zero integral is not arising from the undefined point, but from the small domain in the immediate neighborhood of that point. Of course, Dirac is perfectly aware of the issue as to whether there is a function satisfying this condition. That’s why he ultimately insists on the function’s dispensability.

But is Dirac justified in claiming that the delta function is dispensable? To answer this question, recall, first, that Dirac’s elimination strategy involves identifying two properties of the delta function (expressed in Eqs. (3) and (4), above). These properties not only allow one to operate algebraically with the delta function, but also to eliminate it eventually; after all, on the right-hand side of Eqs. (3) and (4), the delta function doesn’t figure anymore. So, as long as the delta function is only used as a factor in an integrand—which is the only context in which Dirac needs to use the function — and as long as equations (3) and (4) are invoked to operate with the function — which is something Dirac has to do in any case—his strategy works. In other words, Dirac’s dispensability strategy is crucially based on two mathematical properties of the delta function — those expressed by equations (3) and (4). Moreover, by using the delta function as a factor in an integrand, Dirac has no difficulty in using the function to define the length of the basic vectors of observables with continuous spectra. And by invoking later the two mathematical properties of the delta function just considered, Dirac is then able to dispense with the function altogether.

3. The significance of Dirac’s strategy

3.1. The role of mathematics in physics

The discussion above indicates that Dirac is clear about his attitude with regard to the role of mathematics in physics and the nature of applied mathematics. Mathematics plays a pragmatic role: it’s useful in expressing relations among physical quantities, and in establishing (and shortening) derivations. These roles, I insist, are pragmatic at best. The expressive and inferential usefulness of

---

9So, Dirac’s dispensability strategy is not based on the idea that mathematical theories are ultimately tools of derivation in science. Rather, the strategy explores mathematical properties of the delta function to do the dispensability job.
mathematics don’t justify a commitment to the truth of the relevant mathematical theories. As Dirac emphasizes, mathematics is a tool—it’s only a tool.

Mathematics is the tool specially suited for dealing with abstract concepts of any kind and there is no limit to its power in this field. For this reason a book on the new physics [i.e. quantum mechanics], if not purely descriptive of experimental work, must be essentially mathematical. All the same the mathematics is only a tool and one should learn to hold the physical ideas in one’s mind without reference to the mathematical form. (Dirac, 1958, p. viii; italics added)

Dirac’s attitude toward mathematics is, of course, not uncommon among physicists. After all, what matters is not the mathematical content of the empirical theories under consideration, but the physical content. To highlight this instrumental role of mathematics, Dirac also emphasizes the fact that he tries (or at least tried in Dirac, 1958) to “keep the physics to the forefront”, and that he also examines “the physical meaning underlying the formalism wherever possible” (Dirac, 1958, p. viii). This suggests that there is a significant demarcation between the mathematics and the physics, in that the physics leads the way while the “physical meaning” of the mathematical formalism is explored.

None of this is particularly new, of course. But it’s worth highlighting the point, given that it paves the way for Dirac’s own attitude with regard to the delta function. The function clearly has a crucial role in Dirac’s approach, given that it guarantees that every self-adjoint operator can be put in diagonal form.10 Moreover, as we saw, the delta function is also used in the case in which the eigenvalues of an observable are not discrete, and so the basic vectors cannot be normalized (see Dirac, 1958, p. 62). In this case, “to fix the numerical factors by which the basic vectors may be multiplied” (Dirac, 1958, p. 58), the delta function was introduced as a “convenient method of handling this question” (ibid.).

Throughout this discussion, as we saw, Dirac stresses the pragmatic character of the delta function. He also considers the function simply as a “new mathematical notation” (Dirac, 1958, p. 58). The crucial question then arises: Is Dirac’s attitude toward the delta function — in particular his attempt to establish the function’s dispensability—only the result of the peculiar properties of the function (its inconsistency and unusual character), or is there something more general going on? That is, should Dirac’s pragmatic attitude toward the delta function, and his strategy of dispensing with it, be extended to other mathematical entities used in physics? Or is the pragmatic attitude only warranted in the case of ill-behaved mathematical entities (such as, inconsistent objects or objects defined in an unusual way)?

I think there is something more general going on, and Dirac’s attitude should be extended to any mathematical entity referred to in physics. In fact, Dirac’s emphasis (a) on the physical content of his mathematical formulation of quantum mechanics, and (b) on the pragmatic use of mathematics already indicates that he would resist the reifying moves of those who believe in the existence—and indispensability—of

---

10This is the property that Dirac needs in order to establish the equivalence between wave and matrix mechanics. For an insightful discussion of this point, see von Neumann (1932, pp. 3–33).
mathematical objects. The fact that the delta function is inconsistent and unusually defined is certainly a source of worry. But, interestingly enough, Dirac doesn’t raise the issue of the function’s inconsistency (or the way in which it has been defined) as a motivation for its dispensability. The inconsistency only motivates a restricted use of the delta function. As Dirac points out: “$\delta(x)$ is not a quantity which can be generally used in mathematical analysis like an ordinary function, but its use must be confined to certain simple types of expression for which it is obvious that no inconsistency can arise” (Dirac, 1958, p. 58; italics added). In other words, even though the delta function’s inconsistency might be a reason for its dispensability, it actually isn’t. Not, at any rate, according to Dirac.

Rather Dirac’s attitude toward the delta function could easily be the same as that advanced by von Neumann. If our physical theories require inconsistent concepts, we should incorporate the latter concepts into our current mathematical theories in order to preserve the physics. In fact, it wouldn’t be the first time in which the scientific community changed standards and theories in mathematics due to a demand from physics: the introduction of the calculus in the context of Newtonian mechanics did exactly that. As von Neumann argues:

There would be no objection here [to the introduction of inconsistent mathematical concepts, such as the delta function] if these concepts, which cannot be incorporated into the present day framework of analysis, were intrinsically necessary for the physical theory. Thus, as Newtonian mechanics first brought about the development of the infinitesimal calculus, which, in its original form, was undoubtedly not self-consistent, so quantum mechanics might suggest a new structure for our “analysis of infinitely many variables”—i.e., the mathematical technique would have to be changed, and not the physical theory. (von Neumann, 1932, p. ix)

Of course, I am not suggesting that von Neumann or Dirac would willingly introduce inconsistent mathematical concepts. The point is that they would introduce these concepts if the latter turn out to be indispensable to our physical theories. So, the issue has less to do with the status of the concepts in question (e.g. whether they are inconsistent or not), and it has more to do with the role that these concepts play in physics (i.e. whether they are indispensable or not).

As it turns out, the delta function is doubly dispensable: Dirac devised one strategy (namely, to use the function only in restricted conditions) and von Neumann devised another strategy (namely, to bypass the use of the delta function altogether, by formulating quantum mechanics in terms of Hilbert spaces). So, the delta function is not indispensable after all, and so there is no reason to postulate its existence.

3.2. Two ways of dispensing with mathematical entities

This raises the broader issue of the conditions under which we would be justified in believing in the existence of mathematical entities (referred to in our physical theories). And the answer I want to defend here, which I think both Dirac and von Neumann would agree with, is that we are never justified in believing in these
entities. Nothing in the practice of physics actually requires the commitment to these entities. And as we saw, the fact that Dirac highlights that we should “keep the physics to the forefront”, articulating “the physical meaning underlying the formalism wherever possible” (Dirac, 1958, p. viii), clearly indicates that in his view it is possible to distinguish the mathematical and the physical content of a physical theory. Our commitment is to the latter, not to the former. Note that, for Dirac, this is the case even though physical theories can only be expressed in terms of mathematical notions. So, mathematics is used in theory construction, but no claim is ever made about the existence of the corresponding mathematical entities.

Is the physicist then committed to a double standard, denying the existence of those very entities that he or she uses to formulate the best physical theories? W.V. Quine would say so (see Quine, 1953, 1960). In his view, we ought to be ontologically committed to all (and only) the entities that are indispensable to our best theories of the world.11

But I don’t think Quine is warranted here. First, we often quantify over entities whose existence we have no reason to believe. Things like “The average star has 2.4 planets” and fictional characters (such as Sherlock Holmes or Macbeth) provide obvious examples. Note that although we may have good evidence that the average star has 2.4 planets, we may not know the precise number of stars and planets in the universe, and so we cannot dispense with the notion of “average star”. But this doesn’t mean that we are committed to the existence of average stars!12 Moreover, there is no way of understanding the Holmes stories without asserting thinks like “There is a detective who lives in Baker Street”. Our best theories of fiction require us to assert such existential claims. But this doesn’t mean that we are committed to the existence of Sherlock Holmes! Similarly, there may not be any way of formulating quantum mechanics without quantifying over abstract entities (whether they are vectors in a Hilbert space or complex numbers). But this doesn’t mean—and physicists typically don’t mean—that we are committed to the existence of the corresponding entities. But if mathematics is being used, what are the moves that allow us to justify our lack of commitment to mathematical entities?

There are two moves here. The first is to indicate that the relevant mathematical notions are not indispensable after all. They are, at best, convenient notational devices that help the derivation of the intended results. In principle, the latter results could be established with no recourse to mathematical notions. As we saw, this is precisely Dirac’s strategy in the case of the delta function — “a new mathematical notation”, in his own words (Dirac, 1958, p. 58). Of course, it’s not enough simply to say that the offending notions are only useful notations, one needs to establish that

11This is, of course, the first premise in Quine’s famous indispensability argument, which is meant to provide the reason why we should believe in the existence of mathematical entities. The second premise of the argument states that mathematical entities are indispensable to our best theories of the world. From which it then follows that we ought to be ontologically committed to mathematical entities. (For a careful discussion and defense of the indispensability argument, see Colyvan, 2001; for critical responses, see, e.g., Maddy, 1997; Azzouni, 2004, and references quoted therein. As will become clear shortly, I don’t think the argument goes through.)

12See Melia (1995) for a fascinating discussion of this example and others of this sort.
they can actually be dispensed with. And, as we saw, that’s exactly what Dirac did as well, providing a mechanism in which the “improper functions” would only appear in an integrand. As a result, Dirac concluded, “one could then eliminate the improper functions altogether” (ibid., p. 59).

Of course, this provides at best a piecemeal dispensability strategy, which ultimately would need to be applied to each individual mathematical notion used in physics—obviously, a virtually endless task.13 So, a more systematic approach needs to be developed, an approach that doesn’t depend so much on the particular features of the mathematical notions involved, but could be applied across the board. To articulate this approach, the second move is made.

The second move is to distinguish between quantifier commitment and ontological commitment (see Azzouni, 1998, 2004). A quantifier commitment is simply a commitment we incur when we quantify over certain entities. An ontological commitment is a commitment to the existence of certain entities. Note that the former doesn’t entail the latter, since we may quantify over entities whose existence we are not committed to. (The cases mentioned above of fictional entities and average stars clearly illustrate that.) Quine identified these two types of commitment in the case of entities that are indispensable to our best theories of the world: ontological commitment is determined by the quantification over those entities that are indispensable to our best theories. After all, if an entity is indispensable to our best theories, we can’t rewrite such theories without quantifying over the entities in question. And if the theories are the best we have, we can’t do without them either. And so, if we quantify over $X$, and $X$ is indispensable to our best theories, we are ontologically committed to $X$, and $X$ exists. So, in Quine’s picture, in the case of indispensable entities, quantifier commitment and ontological commitment go hand in hand.

Despite Quine’s identification of the two types of commitment, there’s no reason for doing that—even if we grant the indispensability of the entities in question. As noted above, we often quantify over (and thus have a quantifier commitment to) fictional characters. But we don’t take this quantification to provide any reason to believe in the existence of these fictions (even if we grant that quantification over these entities is indispensable). After all, to claim that something exists—to be committed to the existence of the objects in question—we require an additional condition: a criterion of existence needs to be satisfied.

It is, of course, no trivial task to specify such a criterion. And for our present purposes, there’s no need to. It’s enough to provide sufficient conditions for someone to believe in the existence of certain objects. Here are some such conditions: we have observed a given object; we have interacted with it; we have tracked it; or we have developed mechanisms of instrumental access to this object (mechanisms that can be refined). Although any of these conditions can be defeated, taken together they provide significant (independent and sufficient) grounds to be committed to the

13Moreover, as Mark Colyvan pointed out (in correspondence), Dirac’s strategy is not conclusive either. After all, the delta function is being replaced with additional mathematics. And for the nominalist at least, the additional mathematics would still need to be dispensed with.
existence of the objects in question. If conditions of this sort are not met, it doesn’t entail that the objects in question don’t exist (the conditions are only sufficient, after all). But, in this case, the scientific community will typically be suspicious about the existence of the corresponding entities, and will often raise doubts about their existence. ¹⁴

Now, let us contrast the conditions mentioned above with something much weaker. Suppose that the criteria of existence (or the conditions just mentioned) are not satisfied, but we quantify over objects that play a theoretical role in the description of the phenomena. Suppose that quantification over these objects helps to simplify the description of the phenomena, yields a theory that is familiar, fecund and has a large scope. Do these theoretical virtues provide reasons to believe in the existence of the corresponding objects—even though the criteria of existence are not met? I don’t think so. The point is, of course, contentious, and typically, realists and anti-realists assess the issue very differently. But the crucial idea is that simplicity, familiarity, fecundity and scope are pragmatic reasons (see van Fraassen, 1980).

Simple and familiar theories are simple and familiar to us, given our interests, expectations and abilities. A theory’s simplicity or familiarity, on its own, gives no reason to believe that it is true, unless we assume that the world itself is simple or familiar. But it’s not clear how to articulate this claim without postulating the truth of simple or familiar theories—which begs the question.

Similarly, fecundity is also a pragmatic feature: whether a theory is fecund depends on the kind of problems we use the theory to solve; whether a theory yields new consequences is also a contextual matter, which depends on the particular domain to which we apply the theory. So, again, fecundity ultimately depends on us, the users of the theory. It also depends, of course, on the relation between the theory and the world. But, ultimately, we select the theory because it helps us to solve problems we care about. Why should this be a reason to believe that the theory is true? False theories (such as Newtonian mechanics) also are extremely fecund in solving problems.

Finally, the more scope a theory has, the more content it has, and so the less likely it is that the theory is true (van Fraassen, 1985). In other words, having a large scope, on its own, provides no reason to believe that the theory is true. And, again, the case of Newtonian mechanics clearly illustrates that.

¹⁴Azzouni provides (tentatively) one set of criteria of existence in terms of what he calls “thick epistemic access” (see Azzouni, 1997). In his view, we have a thick form of epistemic access to an object if the epistemic access to this object is (i) robust, (ii) can be refined, (iii) enables us to track the object, and is such that (iv) certain properties of the object itself play a role in how we come to know other properties of the object. Observation provides the paramount example of thick epistemic access. Azzouni then distinguishes this thick form of epistemic access from a thin form of access. According to the latter, our access to an object is through a theory that has five virtues: (i) simplicity, (ii) familiarity, (iii) scope, (iv) fecundity, and (v) success under testing. (These are the familiar Quinean theoretical virtues; see Quine, 1976, p. 247.) The point can be presented in the following way: as opposed to thick epistemic access, thin epistemic access provides no reason to believe in the existence of the objects in question—at best, it provides a reason to accept the theory in which these objects are studied. (Azzouni doesn’t use the distinction between acceptance and belief to make his case, but as will become clear, the distinction is useful in the discussion that follows.)
As a result, if a theory satisfies the above theoretical virtues, we have at best reasons to **accept** the objects in question, but we **don’t** have reasons for **belief**.\(^{15}\) It may be **useful** to postulate such objects. But, if only the theoretical virtues are satisfied, and the criterion of existence is not, then we don’t have reason to believe in the existence of the objects in question.\(^{16}\)

Of course, these considerations are not conclusive—and they are not meant to be. The point here is only to **motivate** a distinction between quantifier commitment and ontological commitment. As noted, quantification over certain objects—even in the context of theories that satisfy usual theoretical virtues (simplicity, scope, familiarity etc.)—**on its own** is not enough to guarantee the existence of these objects. Theories that are simple, have large scope, are familiar etc. (such as, e.g., Ptolemy’s theory) may turn out to be false. To be justified in claiming that certain objects exist, an **additional** requirement needs to be met: a criterion of existence.

Where does this leave us with regard to the existence of mathematical entities? In a nutshell, we have, at best, **pragmatic** reasons to **accept** such entities. Mathematical entities are simply not the kind of thing we are able to track, interact with, or devise mechanisms of instrumental access to. Our only access to these entities is **through our mathematical theories**. The theories may be simple, fecund, familiar to us etc. But this gives us at best reasons to **accept** and work with these theories. It doesn’t give us reason to **believe** in the existence of the corresponding objects.\(^{17}\)

In this way, by clearly distinguishing quantifier commitment and ontological commitment, it’s possible to resist the indispensability argument. Its first premise is denied: the fact that we quantify over certain entities—even indispensably so—doesn’t entail that we ought to be ontologically committed to these entities. As a result, since we have no reason to accept the first premise of the argument, we are no longer obliged to accept its conclusion. In the end, the issue is left open as to whether mathematical objects do exist or not. The point is that simply by quantifying over these objects one need not be committed to their existence—nothing in scientific practice requires that.

For these reasons, it’s not surprising that Dirac had the attitude he had toward the delta function. As we saw, he always emphasized the pragmatic character of the function, and never thought that the function’s usefulness provides any reason to believe in its existence. Moreover, the second move—highlighting the distinction between quantifier commitment and ontological commitment—also explains why

---

\(^{15}\)The distinction between acceptance and belief is articulated in detail by van Fraassen (see van Fraassen, 1980, 1985).

\(^{16}\)Maddy provides a very interesting case study about belief in the existence of atoms that beautifully illustrates this claim (see Maddy, 1997). Theories invoking the existence of atoms have been around for a long time, and these theories even satisfied the theoretical virtues discussed above. But the scientific community only stopped being skeptical about the existence of atoms when there was independent evidence for these objects. In other words, it was only when the criteria of existence were satisfied that the community incorporated atoms in the ontology of science. (Although Maddy doesn’t present the point in these terms, this is a plausible reading of her case study.)

\(^{17}\)Using Azzouni’s terminology, we have at best **thin** epistemic access to mathematical entities; we never have **thick** epistemic access to them.
Dirac only considered mathematics as a *tool*, as something that provides convenient notational devices to express relations among objects, but nothing beyond that. As long as mathematical objects fail to satisfy the criterion of existence, we will always be justified in avoiding being committed to their existence—as Dirac certainly was.

4. Objections and responses

Three objections could be raised against the considerations so far. I’ll consider each of them in turn.

4.1. Begging the question?

First, it might be objected that even if we grant the importance of distinguishing between quantifier commitment and ontological commitment, the particular criteria of existence suggested above (or the conditions that motivate such criteria) are inadequate. Recall that the criteria of existence involve being able to observe a given object, interact with it, track it etc. But the trouble here, the objection goes, is that these criteria seem to favor an empiricist epistemology (or something like an empiricist epistemology). After all, the criteria emphasize one form or another of *empirical* access to the objects we claim to exist—and obviously mathematical objects fail this. As a result, the criteria beg the question against the existence of mathematical entities.

In response, recall that the criteria suggested here provide only sufficient conditions for belief in the existence of certain objects; they do not provide necessary conditions as well. So, the fact that the criteria are not met by mathematical objects doesn’t entail that such objects don’t exist. Moreover, note also that the criteria suggested here seem to inform much of the theoretical practice of physicists. And as I tried to argue, Dirac’s own attitude toward the role of the delta function in theory construction, and more generally his attitude toward the role of mathematics in physics, can be better understood as the result of adopting these criteria. And so, at least to this extent, the criteria seem to be reflected in the physicist’s practice. Furthermore, if there is something about the way in which mathematics is used in science that doesn’t allow (or, at least, doesn’t require) the users to believe in the existence of mathematical entities, this is not a prejudice against mathematical entities. It’s simply a reflection on the way in which the practice goes. Of course, my point here is not to claim that scientific practice *demands* that we should all be anti-platonists. My point is only that the practice doesn’t require us to be platonists. Thus, to the extent that the criteria above seem to be informed by the practice, and so are motivated independently of the locus where philosophical disputes are typically conducted, no questions should be begged here. Again, I insist, I’m not suggesting that scientific practice alone settles the issue surrounding such metaphysical controversies. My point is that, even though the practice may not uniquely determine the relevant criteria, it seems to suggest that considerations as those discussed above *do* play a role in the practice of physics.
Given that the practice is typically independent of this sort of controversy, the proposal advanced here doesn’t beg the question.

4.2. Reifying mathematics?

But a second objection emerges at this point. At the beginning of the Principles, Dirac highlights two methods of presenting the “mathematical form” of quantum mechanics (Dirac, 1958, p. viii). The first is the “symbolic method”, which “deals directly in an abstract way with the quantities of fundamental importance (the invariants, etc., of the transformations)” (ibid.). The second is the “method of coordinates or representations”, which deals with “sets of numbers corresponding to these quantities” (ibid.). As Dirac points out, although the method of coordinates is more familiar (at least it was in the 1930s!), the symbolic method “seems to go more deeply into the nature of things” (ibid.). The latter method also has advantages in expressive power, given that it “enables one to express the physical laws in a neat and concise way” (ibid.), and it is the method that Dirac eventually uses in the Principles. Now, and here comes the objection, it’s not possible to make sense of the idea that the symbolic method goes “more deeply into the nature of things” except by reifying mathematics. That is, except by maintaining that mathematical theories are true, and mathematical entities exist. Thus, in the end, Dirac needs to be committed to the existence of mathematical entities to make sense of his own practice!

I don’t think this worry is justified. First, note that, once again, Dirac highlights the pragmatic character of mathematics—at least of the mathematics used in physics. As he points out, the symbolic method “enables one to express the physical laws in a neat and concise way” (Dirac, 1958, p. viii; italics added). However, the ability to express physical laws neatly and concisely is a pragmatic feature. It indicates that mathematics is useful to us. But, as pointed out above, this usefulness provides no reason to believe that the mathematical theories in question are true. Thus, when Dirac asserts that the symbolic method goes “more deeply into the nature of things” (ibid.), this is a reflection on the symbolic method’s expressive power, not a reification of mathematics. After all, using this method, physical laws can indeed be formulated in an elegant and concise form. But, to achieve this result, as we saw, there’s no need to reify mathematics—and Dirac didn’t.

Of course, one can always try to reify mathematics—platonism is not incoherent after all! For example, the platonist may understand Dirac’s claim that the symbolic method “goes more deeply into the nature of things” (ibid.) as an expression that the method yields valuable explanations that couldn’t be obtained otherwise. But, as suggested above with the case of the delta function, nothing in the practice of mathematics or physics requires this reification. And, as we will see in the next section, this point can be extended to other uses of mathematics—even those that led to novel discoveries.

4.3. Accommodating the discovery of antimatter?

At this point, the third and final objection can be formulated. It might be granted that perhaps the uses of mathematics discussed so far (the delta function and the
symbolic method more generally) are pragmatic in character. However, the objection goes, other parts of Dirac’s work clearly illustrate that at least some uses of mathematics cannot be characterized in pragmatic terms. In particular, Dirac’s famous discovery of antimatter (Dirac, 1928a, b, 1930, 1931) can only be understood by reifying mathematics. After all, it was by following the guidance of the mathematical formalism that Dirac predicted the existence of (what later would be called) the positron. In the end, mathematics does play an indispensable role, and no such discovery would be possible without reference to mathematical terms.

Colyvan has recently presented this point as follows:

In classical physics one occasionally comes across solutions to equations that are discarded because they are taken to be “non-physical”. Examples include negative energy solutions to dynamical systems. This situation arose for Paul Dirac in 1928 when he was studying the solutions of the equation of relativistic quantum mechanics that now bears his name. This equation describes the behaviour of electrons and hydrogen atoms, but was found to also describe particles with negative energies. It must have been tempting for Dirac to simply dismiss such solutions as “non-physical”; however, strange things are known to occur in quantum mechanics, and intuitions about what is “non-physical” are not so clear. So Dirac investigated the possibility of negative energy solutions and, in particular, to give an account of why a particle cannot make a transition from a positive energy state to a negative one. (Colyvan, 2001, p. 84)

This investigation would turn out to be extremely fruitful. After all,

Dirac realised that the Pauli exclusion principle would prevent electrons from dropping back to negative energy states if such states were already occupied by negative energy electrons so widespread as to be undetectable. Furthermore, if a negative energy electron was raised to a positive energy state, it would leave behind an unoccupied negative energy state. This unoccupied negative energy state would act like a positively charged electron or a “positron”. Thus, Dirac, by his faith in the mathematical part of relativistic quantum mechanics and his reluctance to discard what looked like non-physical solutions, predicted the positron. (Colyvan, 2001, p. 84; italics added.)

Colyvan then concludes:

Dirac’s equation play[s] a significant role in predicting a novel entity despite the relevant solutions seeming non-physical […]. It is hard to see how a nominalised version of Dirac’s theory would have had the same predictive success. (Colyvan, 2001, p. 84)\(^\text{18}\)

It turns out that the situation was more complicated, and even in the case of the positron, the use of mathematics was at best pragmatic. There is no doubt that the

mathematics played a *heuristic* role, but this is not the same as an *indispensable* role. Ultimately, the work is done by *interpreting* the mathematical formalism—by assigning a physical meaning to the formalism—rather than by the formalism alone. But, as we will see, the *mathematical formalism doesn’t uniquely determine its physical interpretations*, given that it’s compatible with radically different physical scenarios. And so, ultimately, the mathematical formalism *doesn’t* establish what is physically significant. As a result, the formalism cannot have a “significant role in predicting a novel entity” (Colyvan, 2001, p. 84), since it’s compatible with fundamentally different physical entities. It is no surprise that the acceptance of the positron only happened after *independent evidence* was found for it—that is, only after the *criteria of existence* have been satisfied. And this only happened after much controversy about what the experiments established. Given that the mathematical entities quantified over throughout this episode *don’t* meet the criteria of existence, nothing requires their existence—and as we will see, Dirac didn’t claim they did.

It’s important to realize that Dirac himself was initially unsure about how to interpret the existence of negative energy solutions to his equation. In his first 1928 paper, Dirac explicitly acknowledged the problem. The familiar relation between energy, momentum and mass, \( E^2 = c^2 p^2 + m^2 c^4 \), has two roots:

\[
E = \pm c \sqrt{(p^2 + m^2 c^2)}.
\]

But it’s not clear what should be done with the negative solutions. As Dirac points out:

One gets over the difficulty on the classical theory by arbitrarily excluding those solutions that have a negative \( [E] \). *One cannot do this in the quantum theory, since in general a perturbation will cause transitions from states with \([E]\) positive to states with \([E]\) negative.* Such a transition would appear experimentally as the electron suddenly changes its charge from \(-e\) to \(e\), a phenomenon which has not been observed. The true relativity wave equation should thus be such that its solutions split up into two non-combining sets referring respectively to the charge \(-e\) and to the charge \(e\). (Dirac, 1928a, italics added; see also Schweber, 1994, p. 61, and Pais, 1986, pp. 347–348.)

Two points should be noted here. First, Dirac emphasizes the existence of a *physical reason* why in quantum theory negative energy solutions to the energy-momentum-mass equation *cannot* be excluded. After all, a perturbation could cause transitions from positive energy states to negative energy states. If such transitions are not observed, it’s crucial to identify the *physical* reason for that. So, the issue isn’t whether our intuitions about what is “non-physical” in QM are clear or not (Colyvan, 2001, p. 84). The issue is to identify the *physical process* that explains why

---

19 I’ll elaborate on this distinction at the end of this section.

20 Interestingly enough, the empirical detection of the positron was largely done independently of Dirac’s theory. In fact, the first person to report the positron’s detection was Carl Anderson (see Anderson, 1932a, b), who insisted that he was too busy operating his equipment to be able to read Dirac’s papers (see Anderson, 1966; Pais, 1986, p. 352). (I’ll return to this point below.)
we don't observe transitions in which an electron changes its charge from \(-e\) to \(e\). In other words, as Dirac conceptualizes the problem, the issue is physical, not mathematical.

Second, at this point in 1928, because Dirac wasn’t clear about the appropriate physical interpretation of the equation, he argued that eventually the negative energy solutions should be excluded. As he insisted, in the end, “half of the solutions must be rejected as referring to the charge \(+e\) of the electron” (Dirac, 1928a; see also Pais, 1986, p. 348).

It was only after Dirac developed a physical interpretation of the equation — an account that identified the physical process described by the negative energy solutions—that he took the latter solutions seriously. So, Dirac was not blindly following the mathematics wherever it took him. It was not a matter of having “faith in the mathematical part of relativistic quantum mechanics” and being reluctant “to discard what looked like non-physical solutions” (Colyvan, 2001, p. 84). Dirac first identified what he considered physically significant features of the situation, and only then took seriously the negative solutions. In the absence of these physically significant features, he would simply rule out the negative energy solutions to his equations. In the end, these physically significant features were only identified in 1930 (and so two years after the original formulation of the problem). And the features were developed in terms of Dirac’s “hole” theory (see Dirac, 1930). As Dirac claims, in a now famous passage:

Let us assume that there are so many electrons in the world, that all the states of negative energy are occupied except perhaps a few of small velocity. […] Only the small departure from exact uniformity, brought about by some of the negative-energy states being unoccupied, can we hope to observe. […] We are therefore led to the assumption that the holes in the distribution of negative electrons are the protons. (Dirac, 1930; see Schweber, 1994, p. 62)

In other words, the issue here is of a physical, not mathematical nature. And, I insist, even though the mathematics has a heuristic role here, Dirac didn’t consider the negative energy solutions as significant until a physical process was clearly put forward. This is, of course, an additional example of Dirac’s strategy of keeping “the physics to the forefront” (Dirac, 1958, p. viii); a strategy that informed so much of his approach.

The important point—and Dirac himself was explicit about this—is that what was guiding the physical construction here was not the mathematics, but actually, it was chemistry, in particular, the chemical theory of valency. Commenting on Dirac’s strategy, Schweber puts the point very clearly:

The idea of the hole theory was suggested to Dirac by the chemical theory of valency in which one is used to the idea of electrons in an atom forming closed shells which do not contribute at all to the valency. One gets a contribution from an electron outside closed shells and also a possible contribution coming from an incomplete shell or hole in a closed shell. (Schweber, 1994, p. 62)
In fact, Dirac insists:

One could apply the same idea to the negative energy states and assume that normally all the negative energy states are filled up with electrons, in the same way in which closed shells in the chemical atom are filled up. (Dirac, 1978, p. 50)

But, alas, Dirac’s interpretation, in terms of “holes”, was far from adequate! Heisenberg, Pauli, Oppenheimer, and Weyl, among others, raised several objections. For example, in Dirac’s theory, the electron and the proton get the same mass, and the theory entails a much too high rate of annihilation of protons and electrons into $\gamma$-rays to be empirically adequate.21 Because of these criticisms, in 1931 Dirac eventually gave up his “hole” theory:

It thus appears that we must abandon the identification of the holes with protons and must find some other interpretation for them.22 Following Oppenheimer, we can assume that in the world as we know it, all, and not nearly all, of the negative energy states are filled. A hole, if there were one, would be a new kind of particle, unknown to experimental physics, having the same mass and opposite charge to an electron. We should not expect to find any of them in nature, on account of their rapid rate of recombination with electrons, but if they could be produced in high vacuum they would be quite stable and amenable to observation. (Dirac, 1931, italics added; see Schweber, 1994, p. 67.)

This prediction, as is well known, was ultimately vindicated (even though the data didn’t support the existence of “holes” as Dirac initially conceived of them). In fact, in 1932, Anderson reported the detection of the positron (Anderson, 1932a,b). And in the following year, other experimentalists confirmed Anderson’s result (see Blackett & and Occhialini, 1933, and Blackett, 1933). But the positron was certainly not the “hole” initially postulated by Dirac.23

Given all the vicissitudes of the process of theory construction that led to the positron’s detection, it is no wonder that Anderson thought that the discovery of the positron was completely accidental. When asked about the influence of Dirac’s work on his own ideas, Anderson was very clear:

Yes, I knew about the Dirac theory. […] But I was not familiar in detail with Dirac’s work. I was too busy operating this piece of equipment to have much time to read his papers. (Anderson (1966); see Pais, 1986, p. 352.)

---

21See Schweber (1994, pp. 65–67) for a discussion of and reference to the arguments put forward by the critics of Dirac’s hole theory.

22Note that the crucial issue, as Dirac emphasizes, is to identify a physical interpretation of the formalism—that’s where the work is done.

23As Pauli wrote in a letter to Dirac in 1933: “I do not believe on your perception of ‘holes’, even if the existence of the ‘anti-electron’ is proved” (quoted in Schweber, 1994, p. 68). Talking about Anderson’s discovery and the confirmation by Blackett and Occhialini, Bohr added: “Even if all this turns out to be true, of one thing I am certain: that it has nothing to do with Dirac’s theory of holes!” (also quoted in Schweber, 1994, p. 68).
Anderson then added:

[The] highly esoteric character [of Dirac’s papers] was apparently not in tune with most of the scientific thinking of the day. [...] The discovery of the positron was wholly accidental. (Anderson and Anderson, 1983; see Pais, 1986, p. 352.)

Of course, the considerations above don’t diminish at all the importance and significance of Dirac’s discovery. They only highlight that the work was not done by the mathematics. The significant work, as one would expect, was done by the physics—by developing a fruitful physical interpretation of Dirac’s equation. This interpretation identifies the type of entity and the physical properties that this entity should have for the theory to be empirically adequate. As Dirac points out, what needs to be found is a “new kind of particle”, one that has “the same mass and opposite charge to an electron” (Dirac, 1931). But these features are not determined by the mathematical formalism. The latter formalism is compatible with both (i) Dirac’s initial identification of his “holes” with protons, and (ii) Dirac’s later identification of the “holes” with a new kind of particle. And so, clearly, the interpretations are underdetermined by the formalism. Given the underdetermination, the mathematics doesn’t determine what is physically relevant, and given that the mathematical objects referred to in the formalism fail to satisfy the criteria of existence, no reification of the mathematics is justified either. I conclude that, just as with the delta function, the use of mathematics in the case of the positron was heuristic and pragmatic at best. There’s no need to reify mathematics to accommodate this episode.

In other words, (a) in the case above, the crucial work was done by interpreting the mathematical formalism (without a physical interpretation, no empirical predictions could ever be obtained from the latter). But (b) the mathematical formalism underdetermines its physical interpretations. So, it’s ultimately silent about what is physically significant—as one would expect from a mathematical theory. Thus, I think it is misleading to say that the mathematics played an indispensable role (particularly if this is used to justify a commitment to the existence of mathematical entities). Mathematics certainly played a heuristic role. But this is a very different issue.

So how do we distinguish between the indispensable and the heuristic roles of mathematics? A mathematical theory plays an indispensable role if to achieve this role (whatever it turns out to be) the mathematical theory cannot be eliminated. A mathematical theory plays a heuristic role if the mathematical theory is fruitful in generating new ideas (particularly, but not exclusively, if those turn out to be well supported). While the indispensability argument is taken to provide a powerful reason for belief in the existence of mathematical entities (based on the indispensable role of mathematics in science), it’s not plausible to suppose that the heuristic role of mathematics provides such a reason. After all, the ability to generate new ideas gives no reason to believe in the existence of the entities that motivate these ideas. Reading the Holmes stories may generate new ideas about, say, psychological motives of certain actions, but clearly, this gives us no reason to believe in the existence of Sherlock Holmes. Similarly, in the case of the positron, the mathematics at best suggested that there might be some object yet to be discovered. But it didn’t
determine which object was that, nor did it specify its physical properties. As we saw, the mathematics was compatible with radically different physical interpretations. So, the role of mathematics clearly was heuristic.

But was this role also indispensable? That is, could Dirac have predicted the existence of antimatter without referring to mathematical entities? I think he could, because in the end, I think that’s what he actually did. But I grant it’s a controversial issue how to assess a counterfactual condition such as the above. The platonist will insist, of course, that reference to mathematical entities cannot be eliminated, and so we ought to believe in the existence of the corresponding entities. But is this really the case? The considerations above indicate that, when the details of the process of theory construction are examined, there is no need to assign any ontological significance to the indispensable role of mathematics. After all, the predictive work was ultimately done by the physical interpretation of the formalism. In other words, even if mathematics played an indispensable role, it doesn’t follow that we are justified in believing in the existence of mathematical objects. To believe in the latter objects, the criteria of existence need to be satisfied. Given that the criteria are not met, we are not required to believe in the existence of these objects. So, even if the use of mathematics turned out to be indispensable, this is not sufficient to draw ontological conclusions about mathematics. In the end, the mathematics plays at best a heuristic role.

In fact, I think Dirac would agree with this assessment. Discussing renormalization at the end of his life, Dirac insisted:

> These rules of renormalization give surprisingly, excessively good agreement with experiments. Most physicists say that these working rules are, therefore, correct. I feel that is not an adequate reason. Just because the results happen to be in agreement with observation does not prove that one’s theory is correct. (Dirac, 1987; see Pais, 1998, p. 28.)

The same point applies to those who wish to reify the mathematics in the case of the positron.

5. Conclusion

As I argued above, Dirac was clear about the usefulness of the delta function, but he was also clear about its ultimate dispensability. This attitude, however, is not (and should not) be restricted only to the delta function, but should be extended to other mathematical objects, as the use of mathematics in the positron case illustrates. After all, despite the theoretical usefulness of postulating mathematical objects, our access to them is not independently motivated to justify a commitment to their existence.

The picture of the application of mathematics that emerges from Dirac’s view is one in which the crucial work is carried out by suitable physical interpretations of the mathematical formalism, without (i) the commitment to the truth of the relevant mathematical theories or (ii) the existence of the corresponding mathematical objects. With regard to (i), as argued above, in order to play a heuristic role,
mathematical theories need not to true;\textsuperscript{24} they only need to be consistent with the empirical data, given suitable physical interpretations of the mathematical formalism. And even if mathematical theories were to play an indispensable role, that wouldn’t entail that we ought to be ontologically committed to the corresponding objects either, since quantification over mathematical entities alone doesn’t require commitment to their existence. After all, moving now to (ii), the commitment to the existence of certain objects—including mathematical entities—requires the satisfaction of certain criteria of existence. If mathematical objects fail to meet such criteria, nothing requires us to be ontologically committed to them.

By interpreting Dirac’s work via this framework, it becomes clear how, throughout his career, Dirac managed to apply mathematics in insightful and successful ways without having to believe in the existence of mathematical objects. This is, no doubt, an interesting example to follow.

\textbf{Acknowledgements}

My thanks go to Jody Azzouni, Mark Colyvan, Steven French, Alirio Rosales, Harvey Siegel, and Bas van Fraassen for helpful discussions. In particular, Mark Colyvan and Jody Azzouni provided extensive and insightful comments on earlier versions of this paper, for which I am extremely grateful. Their comments led to substantial improvements. Finally, I would like to acknowledge the comments of two anonymous referees, which also led to several changes in the final version of this paper. The material is based upon work supported by a grant from the National Science Foundation, NSF 01-157, NIRT. All opinions expressed here are mine and do not necessarily reflect those of the National Science Foundation.

\textbf{References}


\textsuperscript{24}Hartry Field provided a very interesting program to justify the claim that good mathematical theories need not be true (see Field, 1980, 1989). I actually have a lot of sympathy for the program (see, e.g., Bueno, 1999). But as opposed to the approach suggested in the present paper, Field’s program is not based on mathematical or scientific practice (it’s not meant to be), and it yields a much more complex strategy to dispense with mathematical entities.
