Rational Discovery of the Natural World: An Algebro-Geometric Response to Steiner

Robert H. C. Moir

Abstract. Steiner (1998) argues that the *mathematical methods* used to discover successful quantum theories are anthropocentric because they are "Pythagorean", *i.e.*, rely essentially on structural analogies, or "formalist", *i.e.*, rely entirely on syntactic analogies, and thus are inconsistent with naturalism. His argument, however, ignores the *empirical content* encoded in the algebraic form and geometric interpretation of physical theories. By arguing that quantum phenomena are *forms of behaviour*, not things, I argue that developing a theory capable of describing them requires an interpretive framework broad enough to include geometric structures capable of representing the forms, which set theory provides, and strategies of algebraic manipulation that can locate the required structures. The methods that Steiner finds so suspect or mysterious are entirely reasonable given two facts:

(1) discovering new theories requires algebraic and structural *variation* of old theories in order to access new forms of behaviour; and
(2) recovering the (algebraic and geometric) form of the prior theories is necessary *to retain their empirical support*.

Accordingly, I argue that the methods used to discover quantum theory are both rational and consistent with naturalism.

1. Introduction

In his (1998) book *The Applicability of Mathematics as a Philosophical Problem*, Steiner presents the following question concerning the applicability of mathematics: Why should it be that highly abstract mathematical methods, with no clear input from the world, allow us to discover the basic principles of the universe? This is a version of Wigner's famous problem concerning the "miracle of appropriateness of the language of mathematics for the formulation of the laws of physics" [**29**, p. 14]. The problem as posed by Wigner is rather vague and opinions expressed in the literature have differed with respect to whether it really is a problem (*cf.*, [**9, 15**]). I believe, however, that another more pointed question that Steiner poses calls strongly for a answer:

Is there a physical basis for the abstract
mathematical analogies used to discover quantum theories?

The question here focuses on whether a case can be made that there is a physical basis for the success of purely mathematical arguments in discovering deep, new

2010 *Mathematics Subject Classification.* Primary 00B13 Secondary 01-06.

physical principles. Steiner says that no such basis exists, and concludes that the success of these methods suggests quite the contrary, *viz.*, that, in fact, natural theology could provide the true explanation. Contrary to Steiner, I will suggest that the answer given to this question depends on what *kind of knowledge* of the world physical theories are taken to provide.

Views about the nature of physical laws can usually be divided into two broad categories: those that regard laws as summaries of sensory experience; and those that regard laws as articulating natural tendencies of action or interaction for physical entities or their essences. Views in the former category take laws to be *empirical-phenomenological*: laws correlate experiences, properties or behaviour but do not provide insight into ontology or the nature of an underlying reality. This class of views is compatible with most forms of instrumentalism and includes, of course, Hume's view of causation and positivistic views of laws, as well as those that accept Humean supervenience. Views in the latter of the two categories take laws to be *formal-ontological*: laws express the manner in which actual physical entities tend to behave or interact. The appellation 'formal' is added, first of all, because the entities described by or included in laws are in general unobservable and, hence, may only be characterized formally. Second of all, the inclusion of 'formal' is intended to include views where laws are not thought true of the world directly, but only of abstract models of the world, as well as to include views that take laws to express relations between essences of physical entities.. This class of views, which is compatible with most forms of realism, thus includes Aristotelian physics, where definitions express relations between essences of material substances [**16**] as well as views that take laws to apply only *ceteris paribus*, *e.g.*, [**6**].

Let us now consider the implications of the proven success of purely mathematical analogies in the discovery of the laws of quantum theory given the above two interpretations of the epistemological status of scientific laws.

On a *formal-ontological* view of the knowledge that laws provide, the success of abstract mathematical analogies, which have no physical interpretation, implies that reasoning using purely mathematical structures and constructions can allow us to learn what kinds of things occupy the universe and how they behave. This is highly suggestive of a deep connection between physical ontology and mathematical structures, which allows mathematical analogies to be successful in scientific discovery. Consequently, a formal-ontological view of laws supports an explanation of applicability in terms of Pythagoreanism, where the "being" or "essence" of things in the world is understood to be mathematical. If the things in the world just *are* mathematical, actually or essentially, then it is not a surprise that purely mathematical reasoning could be so successful.[1]

On the other hand, if one adopts an *empirical-phenomenological* view of the knowledge provided by scientific laws, the success of abstract, purely mathematical arguments implies that reasoning using purely mathematical structures and constructions can allow us to probe deeper and deeper into the nature of the world

[1] I do not mean to suggest that a Pythagorean explanation is the only possibility for those who adopt formal-ontological views of laws, I merely suggest that it is the most natural one.

of experience. This is highly suggestive of a deep connection between the structure of the world of experience and mathematical structures, implying then that the operations of our mind are somehow tuned to the exploration of the deepest aspects of the phenomenal world, far beyond immediate experience. Consequently, an empirical-phenomenological view of laws supports an explanation of applicability in terms of natural theology, in the following sense: the world appears to be structured in such a way that the tools and methods of mathematics, themselves products of the human mind, are all that is required to probe the deepest aspects of nature, which is suggestive of a divine hand and a special place for humans in the Great Chain of Being. Although not necessarily for these reasons, it is an explanation of the success of strategies of mathematical analogy in physics in terms of natural theology that Steiner favours.[2]

Although explanations in terms of Pythagoreanism and natural theology are not the only ones possible here, we may see that common views of the knowledge that scientific laws provide (formal-ontological and empirical-phenomenological) point strongly in the direction of anti-naturalistic explanations of the success of the abstract mathematical methods used to discover quantum theories, as Steiner argues.

In this paper I will advocate for an alternative view of the knowledge provided by scientific theories, *viz.*, that *laws specify abstract patterns or forms of behaviour that are exhibited by certain phenomena under certain conditions or circumstances.* The view I intend here cuts across the two categories of views concerning the nature of laws presented above. On the one hand it views laws as *formal* in both the sense that laws, in general, are not considered true of the world directly, but only of abstract models of the world, and in the sense that unobservable phenomena can be characterized formally. On the other hand, the formal characterization of unobservable phenomena is *phenomenological* rather than being ontological in the usual sense. This is because what are normally taken to be "entities" are understood merely as forms taken on by phenomena under certain circumstances, rather than as externally existing objects. Thus, the view intended here is not only phenomenological in the usual sense of "phenomenological theories" but also in a manner that includes "metaphysics" in the sense of existence of unobservable forms of behaviour that can be characterized formally. I will argue that on this *formal-phenomenological* construal of the knowledge provided by the basic equations of a scientific theory, abstract mathematical analogies are actually *necessary* to adapt or generalize theories to describe new phenomena. As a consequence, I will argue that, on this reading of the kind of knowledge provided by scientific theories, the evidence from physics that Steiner presents actually *supports* naturalism.

2. Steiner's Argument Against Naturalism

Steiner [**25**] argues that two strategies of mathematical analogy, which he calls "Pythagorean" and "formalist", have played a central role in the discovery of theories in physics since 1850. In particular, he argues convincingly that these analogical

[2] I do not mean to suggest that an explanation in terms of natural theology is the only possibility for those who adopt empirical-phenomenological views of laws, I merely suggest that it seems to be the most natural one.

strategies were crucial in the discovery of theories in (non-relativistic) quantum mechanics, quantum field theory and particle physics, and that similar methods played a less pronounced role in the discovery of earlier theories by Newton and Maxwell. On the basis of his examination of historical evidence, he draws the conclusion that the success of these strategies of mathematical analogy presents a significant challenge to naturalism. Steiner [**25**, p. 60] summarizes his argument against naturalism as the conclusion of the following two premises:

I. Both the Pythagorean and formalist strategies are anthropocentric; nevertheless,
II. Both Pythagorean and formalist analogies played a crucial role in the fundamental physical discoveries in the twentieth century.

In this section we will examine some of the key features of Steiner's arguments for these two premises.

These terms "Pythagorean" and "formalist" pick out, respectively, two different types of mathematical analogies:

(i) structural analogies, *i.e.*, those based on similar mathematical structure;
(ii) syntactic analogies, *i.e.*, those based on equations with similar form.

By "Pythagorean" analogies, Steiner means structural analogies that "were *then* inexpressible in any other language but that or pure mathematics" [**25**, p. 3].[3] By "formalist" strategies, then, he intends syntactic analogies without regard to the mathematical meaning of the equations.[4] The problem presented by the success of "Pythagorean" analogies is that with no evident connection to a physical motivation for the strategy, it is simply a relation between mathematical structures that is allowing the discovery of deep new theories of the world. Successful "formalist" strategies are all the more problematic since they are determined solely by the form of the equations, which ultimately we choose, and is obviously not determined by the physical world. The evidence that the analogies used by physicists to discover quantum theories were indeed "Pythagorean" or "formalist" in the above senses comes from the consideration of detailed outlines of the actual arguments physicists used to motivate new equations. We now consider an example of each of the two strategies.

One simple, though important, example Steiner provides is that of an argument used by Schrödinger [**24**] to derive the famous equation bearing his name. We consider this argument in detail because we will revisit it later. The following is a version of Steiner's account [**25**, pp. 79-80] of Schrödinger's reasoning. First, in analogy to optical waves, Schrödinger assumes that a particle of constant energy E can be described in terms of a wave with energy $E = \hbar\nu$, where ν is the frequency,

---

[3]Steiner makes a big point of the significance of the physicists' understanding of what they were doing. This is less significant for my purposes because I am arguing that their methods are simply extensions of strategies of discovery that had been successful in the past. My concern is making a convincing case that we can understand their methods as continuous with prior strategies that do have a sound physical basis, which makes it relatively unimportant whether or not they understood what that basis was.

[4]I introduce the two kinds of analogy as "structural" and "syntactic" to separate the compelling evidence Steiner actually presents from the loaded evaluation implied by calling them "Pythagorean" and "formalist".

which yields the wavefunction

$$\psi(x,y,z,t) = f(x,y,z)e^{-i\frac{E}{\hbar}t}. \tag{2.1}$$

The governing wave equation for this function is then

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V(x,y,z)\right)\psi = E\psi, \tag{2.2}$$

which we recognize as the time-independent Schrödinger equation for a free particle. Then, Schrödinger differentiates the expression (2.1) for the wavefunction which yields, after rearranging constants,

$$i\hbar\frac{\partial\psi}{\partial t} = E\psi. \tag{2.3}$$

Note that the validity of this equation can be construed as justifying an interpretation of the energy E in terms of the "time-evolution" operator $i\hbar\frac{\partial}{\partial t}$. Finally, using an analogy to the classical expression $E = \frac{\mathbf{p}^2}{2m} + V$ for the total energy in terms of kinetic and potential energy, resulting in the momentum $\mathbf{p}$ being interpreted in terms of the "space-evolution" operator $i\hbar\nabla$, the governing equation (2.2) motivates the replacement of E in (2.3) to obtain

$$i\hbar\frac{\partial\psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(x,y,z)\right)\psi, \tag{2.4}$$

which we recognize as the time-dependent Schrödinger equation for a free particle.

The "Pythagorean" reasoning, Steiner observes, enters in two main places in this argument. First of all, the argument begins with an analogy to classical optics, yielding a wavefunction (2.1) and corresponding governing wave equation (2.2). The argument then abstracts away from this intuitive optical analogy to obtain a wave equation (2.4) that has no analogue in classical optics and where purely notational complex numbers in the classical context, *i.e.*, in equation (2.1), become an essential component of the resulting wave equation (2.4). This is "Pythagorean" because it is the mathematical expression (2.1) that becomes the basis for the derivation of the wave equation and not its physical meaning under the optical analogy. Second of all, the wave equation (2.4), which was derived using solutions of the special form (2.1), is then considered to be generally valid, which allows superpositions of waves as physical solutions, not valid in the optical analogy. Moreover, if (2.4) is taken to continue to be valid with a time varying V, then the equation permits solutions that do not exhibit ordinary "wavelike" characteristics at all, instead being "smeared out" mass distributions. Steiner also makes the point that there is a "formalist" element in the argument as a result of taking an equation that was derived assuming constant energy to then be valid for cases where energy is not constant.

Steiner provides many examples of "Pythagorean" reasoning such as this, which base analogies on mathematical structures and not physical arguments. Many of these are from particle physics which rely heavily on structural analogies involving symmetry groups. Steiner also provides another class of examples that are even more problematic from a naturalistic point of view because they are analogies based on the *form of the notation* alone. A striking example of such formalist analogies discussed by Steiner [**25**, pp. 157-160*ff.*] is Dirac's "factoring" argument to derive the famous equation bearing his name. After attempts failed to derive a valid equation for the electron by extending the classical-quantum analogies ($E \to i\hbar\frac{\partial}{\partial t}, \mathbf{p} \to i\hbar\nabla$) from the classical energy formula $E = \frac{\mathbf{p}^2}{2m} + V$ to the relativistic version

$$E^2 = \mathbf{p}^2 - m^2, \tag{2.5}$$

a substitution which results in the Klein-Gordon equation, Dirac tried an approach based on the *form* of (2.5). Dirac reasoned that for there to be a valid equation for the electron, first order in space and time variables, it must be the quantum correlate of a factorization of $E^2 - \mathbf{p}^2 + m^2$, *i.e.*, where E and $\mathbf{p}$ are interpreted as operators. Adopting the ansatz that such a factorization existed, Dirac derived the relationships between the coefficients of the factors that had to obtain for the factorization to be valid. He then looked for mathematical structures that could solve the equations, finding that four dimensional spinor fields, with the coefficients interpreted as 4×4 matrices, worked. This allowed the derivation of the Dirac equation, and led Dirac to predict the existence of the positron.

Thus, in this case it was not an analogy between two mathematical structures that allowed the derivation of new laws, it was an analogy based purely on the form of an equation that allowed the discovery of new mathematical structures that could be the subject of new laws. In both these two cases, it is clear that mathematical analogies are driving the arguments. These two cases provide an illustration of the kind of compelling evidence Steiner presents to justify the second (II.) of his two main premises above. Addressing the matter of what to make of this evidence and whether there is any naturalistic justification for these methods will have to wait until the next section.

Since empirical or physical considerations appear to play no role in these arguments, and because Steiner argues that "there is no naturalistic definition of 'mathematics,'" [**25**, p. 108] Steiner concludes that the evidence from physicists' methods on its own, *i.e.*, without regard to the nature of mathematics, is most suggestive of a Pythagorean explanation. The explanation of the success of these methods, then, being that pure mathematical analogy is so successful in theory discovery because the world *is* mathematical; since the world is actually or essentially a mathematical structure, careful mathematical analogies based on known structures and equations can help us to find the deeper mathematical structures that nature instantiates. Steiner advises against a Pythagorean explanation, however, based centrally on his contention that, though it is purely mathematical reasoning that allowed the discoveries, mathematics itself is *anthropocentric.*

This leads us to the reasons underlying the first (I.) of Steiner's two main premises stated above. Steiner's argument for the anthropocentrism of mathematics, and hence against both Pythagoreanism and naturalism, is based centrally on the following claim:

- What counts as mathematics is determined by considerations of *beauty* and *convenience*; that is to say, there is no *objective* criterion for what counts as a mathematical structure.

Steiner does not devote much of the book to supporting this claim, since most of the effort is spent in making an airtight case for premise II. Let us then consider the sort of evidence he provides.

Steiner provides evidence that it is the human aesthetic sense that determines what counts as mathematics. The evidence consists primarily of the recurring example of how the game of chess and its "theorems" do not count as mathematics, together with a selection of quotes from von Neumann and G. H. Hardy, who emphasize the character of mathematics as a creative art and how the selection of mathematical structures for study is primarily aesthetic [**25**, pp. 63-66]. The authority of these twentieth century mathematicians, working in a time where abstract algebra began to dominate the study of mathematics, is supposed to make reasonable the claim that *all* of mathematics is driven by aesthetic concerns, not objectively determined criteria. The evidence for convenience being the other significant determinant of what mathematics consists in is somewhat more varied, being based on how certain mathematical concepts, *i.e.*, complex numbers, potentials and Taylor series, that were historically, or *are* fairly generally in the case of Taylor series, introduced to aid calculation but end up having physical significance.[5] Although this is a paucity of evidence for a claim as grand as "mathematics is anthropocentric," we may grant that *if* all mathematics derives primarily from the human aesthetic sense, and not by any objective criterion, and secondarily only from computational convenience, there is a strong case for anthropocentrism of mathematics.

Based on his evidence for premises I. and II. above, Steiner concludes that "the truly great discoveries in contemporary physics were made possible only by abandoning—often covertly and even unconsciously—the naturalistic point of view" [**25**, pp. 59-60]. And since Steiner argues that because what determines what counts as mathematics is determined by us, the success of "Pythagorean" and "formalist" strategies is at odds even with Pythagoreanism, making the evidence compatible, *inter alia*, with natural theology. This follows, as was intimated above, because the success of mathematical analogies in physics, *given that mathematics is anthropocentric*, is quite suggestive of it being a divine contrivance that the world has been constructed in a way that the conceptual frameworks we develop because of their apparent beauty to us are tuned to the discovery of nature's deepest laws.

One final point that bears mention is Steiner's point that it is a "global strategy" that was successful in discovering the laws of modern physics using purely mathematical analogies. There are far more examples of failure using mathematical analogies than successes. But this does not affect Steiner's argument given that it was a mathematical *taxonomy* that supported the analogies that were successful. It was the overall *scheme* that was successful, not any particular instance, and led successfully to valid new laws in a relatively short period of time.

[5] He also criticizes the practice of introducing perturbation expansions of equations in parameters previously taken to be constant as a "fiction" where the "formalism [is] leading the scientist" [**25**, p. 68]

3. Problems for Steiner's Argument from the History of Analysis

We now turn to consider some of the problems with Steiner's argument. I will argue that there are two general problems with his argument that derive from ignoring two important lines of evidence from the history of mathematics and physics, the first of which is taken up in this section and the second in the section following:

(1) In the period from the 17th to the (early) 19th century, the development of mathematics and physics was closely interdependent; and
(2) Prior physical theories are, and indeed must be, continually used as a basis for the discovery of new theories, which places strong constraints on the kinds of mathematical structures and methods available to formulate new theories.

I will argue that the matter raised by (1) shows that there are objective origins for mathematical structures and that the matter raised by (2) shows, *inter alia*, that the structures needed to formulate new theories were neither arbitrary nor determined by criteria of aesthetic appeal or convenience. These considerations, therefore, serve to undermine the claim that mathematics is anthropocentric, thereby undermining the challenge to naturalism posed by Steiner's argument. Indeed, I will argue in the next section that the structural and syntactic analogies essential in the discovery of modern physical theories, without their "Pythagorean" or "formalist" sheen, can be provided with a naturalistic motivation, which I will argue can lead to an explanation of the physical basis for mathematical analogy in theory discovery.

Because the calculus was not given a rigorous foundation until the late nineteenth century, much of mathematical analysis had to look to intuitive, analogical and physical reasoning to justify its results and methods. This led to a situation from the sixteenth through the middle of the nineteenth century where the development of mathematics was driven in large measure by abstract treatment of scientific problems and phenomena [**1, 18, 19, 26**]. At a foundational level, both Newton's and Leibniz's methods are based on a intuitive notion of continuum, with Newton's concept rooted in our intuition of flow or continuous motion of points in space [**5**, 82],[**12**, p. 97]. Importantly, physical considerations often played a central role in the development of methods of solution of ordinary differential equations (ODE) and partial differential equations (PDE); indeed, the first problems involving partial differential equations were based in the study of continuum mechanics [**1**, pp. 336-7]. The influence of physics on mathematics is also reflected in the fact that the greatest contributors to mathematics in the seventeenth and eighteenth centuries were just as much, or even more, influential as scientists than as mathematicians, examples including Descartes, Newton, the Bernoullis, Euler, and Lagrange, to name a few. Perhaps the strongest indication of the interdependence of mathematics and physics in this period is reflected in the fact that the lack of a rigorous foundation forced the use of physical intuition and accordance with phenomena as substitutes for mathematical proof [**11**, p. 110],[**19**, p. 395,pp. 617-8].

The importance of the reliance on physical intuition, physical meaning, and physical phenomena for the justification of mathematical arguments in the development of mathematical analysis, as well as how important physics was as a source of problems and methods in analysis, provides very strong evidence against Steiner's claims of the non-objective origins of mathematics. The development of analysis is seen to be inextricably entangled with physics and experience of natural phenomena. The historical evidence also shows that the very development of the kinds of equations central to the formulation of physical laws, *viz.*, differential equations, as well as many of their methods of solution, arose from the desire to treat physical problems mathematically. Newton's approach to the calculus in terms of fluxions (time derivatives of quantities in motion) led naturally to quadrature problems relating fluxions of two quantities that were to be solved by determining how the two quantities were related, which expresses the elementary problem of solving a(n ordinary) differential equation [**1**, p. 326]. More generally, physical considerations were central in the development of techniques of solution of ODE [**1**, p. 336]. In the case of PDE, they originated in the study of continuum mechanics, as mentioned above, and later saw significant developments come from potential theory, which, far from being determined primarily by computational convenience as Steiner claims, arise from consideration of the attractions and repulsions of discrete mass points [**1**, p. 337].

That ODE and PDE as classes of equations, and to a large degree their methods of solution, have origins in physical intuition, physical interpretation, or experience of physical phenomena, shows quite clearly that the theory of differential equations is neither determined primarily by aesthetic predilections of human beings nor primarily by computational convenience. This also indicates that differential equations that could be interpreted physically were more likely to be developed and studied throughout the development of analysis. And since differential equations specify or constrain the spatial and temporal variation of the quantities they relate, differential equations encoded[6] the form of behaviour of the natural phenomena to which they could be applied. In turn, the solutions to these equations provided an abstract representation of forms of behaviour exhibited by natural phenomena under the conditions that a differential equation and its solutions could be applied. In this way, the equations of continuum mechanics, including elasticity theory, plastic deformation and fluid mechanics, arguably originate in abstraction of forms or patterns of behaviour in experience. They are clearly not to be regarded as "true" in the sense of an ontologically correct specification of the dynamics of matter, but they do characterize the *form of behaviour* of certain kinds of matter under certain ranges of physical conditions. This shows that the original phenomenological theories of physics naturally lend themselves to a formal-phenomenological interpretation of the sort discussed in the introduction.

Now, the notation developed to express and manipulate differential equations is not entirely arbitrary either. The Leibnizian notation for differentials, derivatives and integrals allows one to reason intuitively in terms of tiny changes, rates of change as ratios of such changes and integrals as sums of tiny parts. Hence, the notation makes reasoning using physical intuition quite natural. This is why the physical arguments in the derivation of differential equations used and continue

[6] In the interest of brevity I will not consider in detail the sense of "encode" I intend here. It will suffice for our purposes to specify that it is the *flows* that dynamical (hyperbolic, parabolic, ordinary) differential equations pick out which are the *forms* or *patterns* of behaviour that the equations encode.

to make use of differentials (see, *e.g.*, [**4**],[**10**]). The dominance of the algebraic Leibnizian notation in differential equations is no accident.

For differential equations, the notation reflects both the mathematical and physical meaning of the equations. For example,

$$\frac{d^2x}{dt^2} = \frac{F}{m}$$

reflects the mathematical meaning that the curvature of $x(t)$ with respect to time is equal to the ratio of F to m; it also reflects the physical meaning of Newton's second law as "a change in motion is proportional to the motive force impressed and takes place along the straight line in which that force is impressed" [**23**, 416]. This also means that a (large or small) change of $F(x)$ from one function to another, or m from one value to another, has a definite effect on the mathematical structure defining the dynamics, by changing the direction or the magnitude of the vector field. With an understanding of the physical meaning of the equation, then, it is possible to reason about the physical effects of variations or changes to the differential equation. If we regard the flow of the differential equation, *i.e.*, the continuous time transition map on phase space, as characterizing the form of behaviour the equation picks out in general, then understanding the physical meaning allows reasoning about what the effects are of variations of the equation on the form of behaviour. In this way, the algebraic form of differential equations *covaries* with their mathematical meaning in terms of functions on geometric spaces, allowing translation of algebraic arguments into physical ones.

Thus, we can regard the Leibnizian notation for the calculus as being the solution to an important practical problem, *viz.*, how can we develop a symbolic notation that humans can easily manipulate in such a way that understanding simply the meaning of the symbols allows insight into abstract representations of physical behaviour. This might appear like a matter of convenience, but when we realize that without this property of the notation it would not be feasible to successfully use differential equation methods widely in the manner required by mathematical physics, we see there is an *objective* constraint that is driving the choice of notation. This is to say, that being able to find such a sort of notation or not being able to do so marks objectively the difference between methods that permit the kind of sophisticated intuitive reasoning that has been essential in the historical development of physics and methods that would not allow physics as we know it to develop. Thus, the choice of notation is determined in large measure by what kinds of inferences it renders *feasible*, *i.e.*, possible in the actual world, and not merely by concerns of practical convenience. Indeed, this is supported by the limited developments in Britain from those who pursued Newton's geometric approach to the calculus compared to the fruitful developments on the Continent resulting from the use of the algebraic Leibnizian notation. As a simple example to drive the point, try doing complex arithmetical calculations in Roman numerals rather than using positional notation—the notation has a strong impact on what inferences are feasible, *i.e.*, *actually* possible rather than simply possible in principle.

Viewed in terms of feasible access to knowledge of forms of behaviour of natural phenomena, then, neither the structures that abstract patterns of physical behaviour nor the notation we choose to present those abstract patterns are arbitrary nor determined solely by beauty and convenience. The significance of this for our purposes is that the mathematical structures, *i.e.* solutions of differential equations, are dictated by the form of behaviour of phenomena; and that the notation and methods are dictated, centrally, by the requirement of feasible means of physical insight and the study of the mathematical structures that represent the phenomena. The point here is that it is the combination of the type of equations, the physical representation of their mathematical meaning *and* the specialized notation, that allowed physicists to feasibly access mathematical representations of such a wide variety of forms of behaviour. In this way, differential equations and their theory as developed in the seventeenth to mid-nineteenth centuries were a solution to a central practical problem for science: How can we develop mathematical structures and methods that allow us to extend and deepen our knowledge of the behaviour of natural phenomena. Beauty and convenience were certainly relevant constraints, but the primary constraint was feasible insight into forms of physical behaviour.

4. A Naturalistic Understanding of the Content and Function of Theories and Methods of Their Discovery

The evidence provided in the previous section shows the distinct sense in which the origin and the content of classical analysis, particularly differential equations, was abstract representation of forms behaviour of phenomena, which could then, *inter alia*, be studied to gain insight into behaviour.

The theories so formed initially involved phenomena that could be directly observed, like the rigid motion of bodies, static fluids, fluid flow, elastic behaviour of solids, motion of celestial bodies, *etc.* These are the sorts of phenomena covered by differential equations in the manner considered in the previous section. As phenomena became more removed from experience, however, mathematical analogies to prior theories played an increasingly important role.

One example of this was the development of a theory of light as a transverse wave in a medium, the luminiferous aether, that was consistent with experiment. After empirical evidence from Arago led to the hypothesis of light as a transverse wave, a model of the aether as an elastic solid was sought, since it was known that an elastic solid allows transverse waves [**28**]. Thus, an analogy was drawn to a phenomenon, *i.e.*, elastic solids, for which equations of motion were known, in the interest of developing an equation of motion for light. By deriving physical boundary conditions for real elastic solids, Green was able to show that modeling the aether as an ordinary elastic solid was not consistent with experiment [**28**]. This then initiated the search for *variations* of the equations of an elastic solid that were empirically adequate, which was accomplished first by MacCullagh [**28**].

This episode from optics shows how a mathematical analogy was used to determine equations of motion that governed the form of behaviour of light waves, which are not directly observable. The approach started with known equations that admit transverse wave solutions and then required variation of these equations to *extend* physical theory to describe the sort of transverse waves exhibited by light. Just as in the case of ordinary phenomenological theories, these equations are obviously not ontologically correct, since we know now that even the description of light as an electromagnetic wave is a macroscopic approximation. The analogical strategy did, however, succeed in producing equations that capture dominant aspects of the way that microscopic light waves *behave*. A similar, but more involved, strategy was employed by William Thomson and Maxwell to develop equations of motion governing electrical and magnetic phenomena [**7**]. Independently of the mechanical and aether models Maxwell used in the process of finding a consistent set of equations that agreed with experiment, the result were equations that captured more and more aspects of the *behaviour* of electric and magnetic phenomena, culminating in equations that governed the behaviour of electromagnetic fields and the waves they permit.

The key point here is the following. If we continue to understand the function of differential equations in science as specifying abstract patterns or forms of behaviour that certain physical phenomena exhibit under certain circumstances, then mathematical analogies play a clear and rational epistemic role. Starting with differential equations we know are physically valid for phenomena under certain conditions and admit solutions with behaviour similar in relevant respects to that of some new phenomenon, it stands to reason that a modification of the equations may lead to new equations that govern the form of behaviour exhibited by the new phenomenon. The idea is that the variation of old equations can allow access to new forms of behaviour not previously captured by physical theory. And particularly when the phenomena whose form we are modeling are not accessible in experience, the only epistemic access we have to these forms is through the use of mathematics.

With (distinctly) quantum phenomena being even further removed from experience than those of optics and electromagnetism, the reliance on abstract mathematical analogy became even more important for discovering quantum theories. Here once again, however, strategies of mathematical analogy are rational and naturalistically justifiable provided that we regard the aim of these strategies as accessing new forms of behaviour not accessible to current physical theory.

Consider Schrödinger's derivation of (2.4) discussed above in this connection. He begins with the assumption that under certain circumstances a quantum particle can be described as a wave of the form (2.1), an assumption Steiner admits had physical justification [**25**, p. 79]. Differentiation of this wavefunction together with an analogy between its governing PDE and the classical expression for the energy then led to Schrödinger's equation (2.4). Contrary to Steiner's implications, however, the argument does not require the abandonment of fixed energy in a way that is contradictory or conceptually problematic. The move from the assumption of a wavefunction of form (2.1) with constant energy E permits differentiation to obtain (2.3). Replacement of E by the operator from the governing wave equation (2.2) for the initially assumed wavefunction (2.1) then yields the Schrödinger equation (2.4). All of this is rigorous and at fixed energy, as Steiner admits, since it simply requires a time-varying wavefunction with fixed energy, as in (2.1).

The importance of this strategy for Schrödinger is that it allowed the extension of the governing equation (2.2) for the syncronic situation to the equation of motion (2.4) for the corresponding diachronic situation. Once this equation has been reached, it is natural for a strategy of discovery of laws governing new phenomena that we suppose that this equation is valid in cases beyond the special one of the solution (2.1) that was used to obtain it. Thus, it is only by *interpreting* (2.4) *as generally valid*, or adopting it as an ansatz that (2.4) is generally valid, that permits *additional* solutions with non-constant energy. Furthermore, the superposed solutions this general validity thereby allows are actually characteristic of quantum phenomena, which indicates the nature of the discovery that had been made. This leads us to a key point: analogies based on *variation* can, and must if a discovery is to be made, lead to new behaviour. Thus, interpreting strategies of mathematical analogy in theory discovery as part of a naturalistically justified tactic for finding new forms of behaviour and the laws that govern them, the generalization step that Steiner finds naturalistically problematic is actually just the sort of variational strategy needed here to extend our understanding of nature.

Thus, we see that it is the quantum-classical analogies ($E \to i\hbar\frac{\partial}{\partial t}, \mathbf{p} \to i\hbar\nabla$) to *access* a new equation of motion for dynamical quantum phenomena *together* with the assumption that this equation applies generally that really drive the reasoning in Schrödinger's argument. So provided that these analogies can be understood as supporting the adaptation of classical equations of motion to quantum phenomena, involving a significant change in scale, then Schrödinger's analogical strategy is quite naturally comprehended as an extension of the analogical methods used in optics and electrodynamics, only in the interest of a much more extreme extension of physical theory. Indeed, there is a justification for these analogies in terms of generators of canonical transformations in Hamiltonian mechanics, which is what actually provided the central background framework to support the extension of classical theory to quantum phenomena (see [**10**], chapter 9).[7]

We also saw above that these same quantum-classical analogies drove the argument Dirac provides to derive his equation for the electron. This argument falls in Steiner's particularly problematic class of "formalist" strategies, since it is based on the form of the relativistic energy equation (2.5). The symbols used are clearly irrelevant to the analogy; what matters is the relationship between the physical quantities of energy, momentum and mass that the equation expresses. Given that this is a relativistic expression for energy, we would expect a relativistic theory to satisfy it. But what kinds of mathematical structures E, $\mathbf{p}$ and m refer to need not necessarily matter, provided that the structures can be interpreted in terms of energy, momentum and mass in such a way that they are related according to (2.5). Indeed, it is not unreasonable to expect that a successful extension of physical theory to relativistic electrons would require new mathematical structures to characterize their behaviour.[8] Understood in this way, Dirac's anstaz is not purely

[7]Note that it is more rigorous analogies based on converting Poisson brackets to commutation relations, or recovering Poisson bracket relations from commutation relations in a classical limit, that ultimately justified the "correspondence principle" implied by the analogies ($E \to i\hbar\frac{\partial}{\partial t}, \mathbf{p} \to i\hbar\nabla$) [**10**, p. 390].

[8]I mean here new to physics and not necessarily new to mathematics.

"formalistic" and is reasonable in the interests of a search for equations of motion that characterize the behaviour of the electron. Furthermore, if we understand this episode as a successful search for such an equation that encodes behaviour in a hitherto inaccessible level of phenomena, then it is considerably less mysterious, though certainly still surprising, that finding the equation that encodes the behaviour of such phenomena also reveals new phenomena that occur at that level, such as the existence of antimatter.

The validity of the argument that the kinds of mathematical analogies used to discover quantum theories are rational and naturalistic when we adopt a formal-phenomenological interpretation of physical laws depends essentially on it being the case that quantum phenomena are forms of behaviour, just as classical phenomena characterized by differential equations are. As has been mentioned, the phenomenological theories of the seventeenth to nineteenth centuries clearly do not provide laws of interaction for entities, since matter is now known (post Perrin) to not be continuous. Thus, a formal-phenomenological interpretation of phenomenological theories, including theories in wave optics and electromagnetism, is quite natural given the above considerations. Yet, quantum theories, which I am arguing were developed by varying abstract forms of behaviour, are sometimes taken to magically(!) reveal the fundamental ontology of the world—a radical shift from phenomenology to ontology using the *same variety of discovery strategy*—variation of equations and their solutions. The continuity of analogical strategies in theory discovery in physics therefore suggests that *quantum phenomena are forms of behaviour* (of matter-energy stuff), so that the *abstract form* of quantum phenomena can be characterized mathematically, just as classical phenomena are.

There are, however, deeper physical reasons to suppose that quantum phenomena, including all the particles of the Standard Model of particle physics, are actually forms of behaviour (of matter-energy) and not things. Halverson and Clifton [**14**], building on results from Malament [**21**], have provided rigorous "no-go theorems" that rule out *any* particle ontology in a relativistic quantum theory. They even rule out a localizable particle ontology supervenient on fundamental fields in quantum field theory. Thus, it turns out to not be possible to interpret the particles of the Standard Model as actual entities. Halverson and Clifton show, however, that it is possible to have a relativistic quantum field theory that, for all practical purposes, has field excitations that *behave* like localizable particles, therby explaining the appearance of particles empirically. Thus, we have rigorous results that show that what may seem to us distinctly as particles, *viz.*, electrons, protons, positrons, *etc.*, are merely forms of behaviour of matter-energy-stuff, whatever that may be, under certain measurement conditions.

The no-go results cited above were originally taken to support a field ontology, but it is now known ([**2**],[**13**]) that even natural field ontologies are not possible in quantum field theory.[9] Thus, despite a strong desire and considerable effort to establish some fundamental ontology, there is strong evidence that even our

[9] Lupher [**20**] shows that in the context of algebraic quantum field theory an ontology of "smeared" fields *is* possible. But even if an ontology is consistent with a given theory, this does not necessarily give us reason to believe that it is veridical (see, *e.g.*, [**8**]).

most foundational theories establish only new levels of phenomenology, forms of behaviour at different scales of space, time and energy. This conclusion is in line with the view, held by some, that there will never be a fundamental theory and that all physics provides access to are effective forms of behaviour at a given scale (see [**17**] for a discussion of the issues underlying this debate).

We are now in a position to consider my argument concerning the physical basis of mathematical analogy in theory discovery. In the case of the use of strategies of structural analogy in theory discovery, we can understand the mathematical structures determined by differential equations (individual solutions, invariants, flows, *etc.*) as abstract representations of forms of behaviour that phenomena can exhibit under certain conditions. So it is rational strategy to vary the abstract forms of behaviour (by varying the structures picked out by the equations) of empirically successful theories in order to extend theories to new phenomena or new scales. The physical basis for this idea is that, on a formal-phenomenological view of theories, physical theories can be understood as all describing the same stuff but simply different modes of it at different scales and under different conditions. Therefore, varying the structures picked out by equations, in a manner consistent with a change of mode, scale, or conditions, can allow the description of new forms of behaviour.

In cases of the use of strategies of syntactic analogy in theory discovery, such strategies can be rational on account of the fact that we can understand the algebraic form of differential equations as covarying with its mathematical interpretation in terms of geometric or topological spaces, the structures of which are the abstract representations of forms of physical behaviour. As a result, the algebraic form can encode important relationships between physical quantities that we expect, on physical grounds, to be preserved or varied in a shift to new phenomena. This can be construed as an *algebro-geometric* reading of the equations, so that reasoning based on algebraic form is not actually purely syntactic as a result of the covariation of the algebraic form and its geometric and topological interpretation.

Moreover, by varying the old equations we ensure that the old equations are recovered by reversing or undoing the variation, along with their physical meaning and the empirically demonstrated forms of behaviour that they encode. In this way, the equations of the old theory act as a *surrogate for agreement with experiment.* This is because by showing that one recovers the equations of an old theory, either directly or in an appropriate limit, one also shows that the new theory exhibits (exactly or effectively) *the same forms of behaviour* under the conditions covered by the old theory.

In this way, then, the strategies of both structural and syntactic analogy can be understood to have a sound physical basis. For strategies of variation to be successful, a definite interpretational framework within which the structural variation is understood to occur is required. This has been ensured since the establishment of axiomatic set theory as a foundational framework for mathematics. The universality of ideal forms of description in classical sets ensures the *existence* of structures that can extend known physical forms of behaviour to new modes, scales or conditions. Thus, the key requirement of a strategy of discovery is that it provide a means of manipulating the algebraic form of equations known to be valid, implying covariation of the geometric and topological interpretation of the equations, in order to *locate* forms in the universe of sets that capture the form of behaviour of phenomena in some new mode, under different conditions, or at different scales. But this is precisely what the "Pythagorean" and "formalist" strategies can be understood to accomplish when we adopt a formal-phenomenological view of the knowledge that physical theories provide. Seen in this new light, a strong case can be made that the large body of examples of structural and syntactic analogy in theory discovery in modern physics are actually examples of a naturalistic strategy of discovery of new theories. Moreover, such strategies are all we have to work with when dealing with phenomena that we can only access or understand through the use of equations and mathematical structures. Therefore, though these achievements and discoveries in twentieth century modern physics are marvellous, striking and surprising, they are certainly not magical or inexplicable.

References

1. Tom Archibald, *Differential equations: A historical overview to circa 1900*, A history of analysis, vol. 24, AMS Bookstore, 2003, pp. 325–354.
2. David J. Baker, *Against field interpretations of quantum field theory*, The British Journal for the Philosophy of Science **60** (2009), no. 3, 585–609.
3. Robert W. Batterman, *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence*, Oxford University Press, USA, 2002.
4. John L. Bell, *A primer of inifinitesimal analysis*, Cambridge University Press, 1998.
5. ______, *The continuous and the infinitesimal in mathematics and philosophy*, Polimetrica, 2005.
6. Nancy Cartwright, *The dappled world: A study of the boundaries of science*, Cambridge University Press, 1999.
7. Olivier Darrigol, *Electrodynamics from ampère to einstein*, Oxford University Press, 2003.
8. Chris Fields, *A physics-based metaphysics is a metaphysics-based metaphysics*, Acta Analytica (2013), 1–18.
9. Steven French, *The reasonable effectiveness of mathematics: Partial structures and the application of group theory to physics*, Synthese **125** (2000), no. 1-2, 103–120.
10. Herbert Goldstein, Charles Poole, and John Safko, *Classical mechanics*, Addison Wesley, 2002.
11. Ivor Grattan-Guinness, *Mathematical physics in france 1800-1840; knowledge, activity, and historiography*, Mathematical Perspectives: Essays on the History of Mathematics in Honor of Kurt.-R. Biermann (1981), 95–138.
12. Niccolò Guicciardini, *Newton's method and leibniz's calculus*, A history of analysis (2003), 73–103.
13. H Halvorson and M Mueger, *Algebraic quantum field theory*, Philosophy of Physics (2007), 731–922.
14. Hans Halvorson and Rob Clifton, *No place for particles in relativistic quantum theories?*, Philosophy of Science **69** (2002), no. 1, 1–28.
15. Richard W. Hamming, *The unreasonable effectiveness of mathematics*, The American Mathematical Monthly **87** (1980), no. 2, 81–90.
16. Rom Harré, *Laws of nature*, A Companion to the Philosophy of Science (W. H. Newton-Smith, ed.), Blackwell Companions to Philosophy, Blackwell, 2001, pp. 213–223.
17. Stephan Hartmann, *Effective field theories, reductionism and scientific explanation*, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics **32** (2001), no. 2, 267–304.
18. Hans N. Jahnke, *Algebraic analysis in the 18th century*, A history of analysis **24** (2003), 105–136.

19. Morris Kline, *Mathematical thought from ancient to modern times*, vol. 2, Oxford University Press (1990), 1972.
20. Tracy Lupher, *Not particles, not quite fields: An ontology for quantum field theory*, Physics and Metaphysics (2010), 155–173.
21. David B. Malament, *In defense of dogma: Why there cannot be a relativistic quantum mechanics of (localizable) particles*, Perpectives on Quantum Reality (R. Clifton, ed.), Kluwer, 1995, pp. 1–10.
22. Robert H. C. Moir, *Structures in real theory application: A study in feasible epistemology*, PhD. Thesis. University of Western Ontario - Electronic Thesis and Dissertation Repository. Paper 1578. url: http://ir.lib.uwo.ca/etd/1578, 2013.
23. Isaac Newton, *Mathematical principles of natural philosophy*, The Principia: Mathematical Principles of Natural Philosophy (I. Bernard Cohen and Anne Whitman, eds.), University of California Press (1999), 1729.
24. Erwin Schrödinger, *Quantization as a problem of proper values*, Collected papers on wave mechanics (1926), 1–12.
25. Mark Steiner, *The applicability of mathematics as a philosophical problem*, Harvard University Press, 2002.
26. John Stillwell, *Mathematics and its history*, The Australian Mathematical Society (2002), 168.
27. Paul Teller, *An interpretive introduction to quantum field theory*, Princeton University Press, 1997.
28. Edmund T. Whittaker, *A history of the theories of aether and electricity: Vol 1 the modern theories 1900-1926*, Nelson, 1961.
29. Eugene P. Wigner, *The unreasonable effectiveness of mathematics in the natural sciences. richard courant lecture in mathematical sciences delivered at new york university, may 11, 1959*, Communications on pure and applied mathematics **13** (1960), no. 1, 1–14.

Western University
Department of Applied Mathematics
Current address: Middlesex College 275a, 1151 Richmond St., London, ON CANADA N6A 5B7
E-mail address: robert@moir.net