Ruetsche, Laura (2011) *Interpreting Quantum Theories*, New York: Oxford University Press, Chapter 1.

— 1 —

Exegesis Saves: Interpreting Physical Theories

1.1 Why be so particular?

If, like Plato's Eleatic Stranger, we had a mania for dichotomous division, we might sort philosophers into two types: the generalists, distinguished by their sweeping and lofty concerns with (for instance) the natures of truth, meaning, virtue, and knowledge; and the particularists, distinguished by their obsessive attention to earthly details, such as the allocation of resources in an industrial society, or the physics of quantum non-locality. In his 1974 presidential address to the Eastern Division of the American Philosophical Association,[1] John Rawls considers his particular philosophy, moral theory, which he understands as

> the study of substantive moral conceptions, that is, how the basic notions of the right, the good, and moral worth may be arranged to form different moral structures. Moral theory tries to identify the chief similarities and differences between these structures and to characterize the way in which they are related to our moral sensibilities and natural attitudes, and to determine the conditions they must satisfy if they are to play their expected role in human life. (1974, 5)

Philosophers of physics may recognize their own labors as analogously particularized. Studying substantive physical theories, philosophers of physics seek to articulate the structures by which those theories represent, predict, and explain. They try to identify the chief similarities and differences between prominent theoretical structures, and to characterize the way those structures are related to empirical data as well as to metaphysical home

[1] Brought to my attention by Jon Mandle, to whom I am grateful.

truths. They seek to determine the grounds for, and the meaning of, acceptance of theoretical structures as successful science.

In his presidential address, Rawls takes on the prejudice that moral theorists can't get on with their jobs until generalists—epistemologists, philosophers of language, philosophers of mind—have finished theirs. 'It is thought,' he observes,

> first that other philosophical questions cannot be satisfactorily resolved until the problems of epistemology, or nowadays the theory of meaning, are already settled; and second that these prior questions can be investigated independently: their answers neither rest on nor require any conclusions from other parts of philosophy. (1974, 6)

The prejudice is that generalist philosophy can be conducted in isolation from particularist philosophy, and indeed that the former's deliverances serve the latter in some sort of legislative role. Brought home to the philosophy of physics, one expression of the prejudice is the thought that one must consult a generalist's account of explanation to tell whether quantum mechanics (QM) explains the distant correlations it predicts. Another expression is the expectation that the scientific realism debate can be conducted (and maybe even settled) in abstraction from the details of particular scientific theories.

Against the expression of prejudice in the sphere of moral philosophy, Rawls contends not only that 'preoccupation with the problems that define [generalists'] subjects may get in the way and block the path to advance' (1974, 6) but also that ground level engagement with the structures of substantive moral conceptions may clear that path. Contending that 'the analysis of moral concepts, the existence of objective moral truths, and the nature of persons and personal identity, depend on an understanding of these structures,' Rawls concludes that '[t]he problems of moral philosophy that tie in with the theory of meaning and epistemology, metaphysics, and the philosophy of mind, must call upon moral theory' (1974, 6). His conclusion liberates the particularist, and gives moral theory a voice in moral philosophy more generally cast.

I mean this book as an exercise in liberatory particularism. Its primary focus is the interpretation of Quantum Field Theory (QFT) and the thermodynamic limit of Quantum Statistical Mechanics (QSM)— theories I gather under the heading "QM_∞." (The subscript will be explained presently.)

The book's primary aim is to suggest that Rawls was right about more than moral theory. I will contend that questions generalists in the philosophy of science typically address in lofty abstraction—questions such as, "What are laws of nature?" and "What epistemic attitude should we strike toward successful physical theories?"—are transformed when brought into contact with QM_∞. For example: it is generally supposed by those who consider the matter in lofty abstraction that to interpret a physical theory is to identify the set of worlds possible according to that theory. The idea is that a theory says what it says by allowing some states of affairs (the worlds possible according to the theory) and forbidding others (the worlds by its lights physically impossible). It is also usually supposed that it's the job of a theory's laws to circumscribe its collection of possible worlds: its possible worlds are exactly those consistent with its laws. While which (if any!) of a theory's possible worlds is actual may be a contingent matter (the historical accident of what the initial conditions were, say), what the theory's possible worlds are depends only on its laws.[2] Thus loftily considered, physical possibility is what one might call 'unimodal.' Everything that's physically possible is physically possible in the same way, and interpretation is the *pristine* business of identifying a theory's physical possibilities by bringing one's own metaphysical scruples and technical sophistication to bear on its laws.[3]

Underlying this methodological *ideal of pristine interpretation* is a distinction between what holds at each world of which a theory T is true and what varies—as John Earman put it in his 2002 presidential address to the Philosophy of Science Association, 'a distinction between what holds as a consequence of the laws of physics and what is compatible with but does not follow directly from those laws' (2004a, 1229).[4] The class of what applies in all settings where T applies includes T's laws, as well as metaphysical, methodological, and mathematical truths; the class of what changes from setting to setting includes initial/boundary conditions, as well as practical

[2]See, for example, Suppe (1977, 226-8) and van Fraassen (1989, 222-223) on 'laws of coexistence' and 'laws of succession': the former determine which instantaneous configurations are 'physically possible' (Suppe 1977, 226); the latter 'select the physically possible trajectories in the phase space' (*ibid*).

[3]I refer those who are wondering whether an uninterpreted theory even *has* laws to §1.2's contention that the theories we care about are already partially interpreted.

[4]Chapter **??** addresses Earman and Wigner's contention that this 'sharp distinction' is *presupposed* by notions of symmetry and invariance.

considerations parochial to the settings which give rise to them. To adhere to the ideal of pristine interpretation is to invoke only considerations from the former class when circumscribing the collection of worlds that are possible according to T. Its to aim at a commodity aptly described as (say) *the* interpretation of QM, all told, rather than a set of commentaries on the theory's subspecies and/or piecemeal applications. A virtue Everett famously claims for his own relative state formulation of QM aptly expresses the limiting case of ideal interpretation: all extra-theoretical considerations drop away, and 'the theory itself sets the framework for its interpretation' (Everett 1957, 462).

The 'Samurai TV repairman' portrayed by John Belushi on *Saturday Night Live* in the late 70s is an apt caricature of the pristine interpreter and his commitment to unimodality. No matter what ailing appliance was brought forth for his attention, Belushi's TV repairman would respond in the same way. Without pausing to examine the workings of the ailing appliance, he'd brandish an enormous sword conveniently hung from his belt, and gleefully cut the TV set in two. Just so, T's pristine interpreter, confronted with the space of logically possible worlds, splits off the worlds possible according to T in one fell swoop. Just as Belushi's TV repairman defiantly ignores the interior details of the appliance before him, the pristine interpreter consigns differences *between* circumstances that fall under T to 'the astronomer, geographer, geologist, etc.' (as Wigner, Houtappel, and van Dam (1965, 596) once memorably put it). For the pristine interpreter, variations between worlds to which T applies make no difference to what's possible according to T.

The ideal of pristine interpretation is not often directly voiced.[5] But it pervades the practice of philosophy of physics over the last half century. Witness competing interpretations of spacetime theories, which by muscular application of differential geometry and the like, tie their banners to competing accounts of the metaphysics of substance and identity.[6] And observe how Redhead's classic *Quantum Mechanics, non-locality, and realism* characterizes the task of interpreting QM:

> [An interpretation of QM] is simply some account of the nature of the external worlds and/or our epistemological relation

[5]An exception may be Fine (cf. 1986, 148), whose NOA parts ways with the ideal.

[6]The vast literature on the dreaded hole argument is a good place to see this in action.

> to it that serves to *explain* how it is that the statistical regularities predicted by the formalism with the minimal statistical interpretation come out the way they do ... (1989, 44).

By suggesting that an interpretation of QM should emanate from general principles of metaphysics and epistemology, Redhead evokes the ideal of pristine interpretation. He moreover implicates the *generality* of the principles informing an interpretation in its capacity (constitutive, he reckons, of its status *as* an interpretation) to explain. Contending that 'theories that lack an interpretation ... simply do not contribute to our *understanding* of the natural world' (45), Redhead attributes the explanatory oomph of an interpretation X of a theory Y to 'the "unifying" effect of X. A few general principles about the nature of reality expressed in X comprehend a wide variety of seemingly unconnected observational regularities, including Y' (45). The mathematician Irving Segal expresses a similar commitment to pristine interpretation when he describes his favored strategy for interpreting QFT: 'The proper sophistication, *based on a mixture of operational and mathematical considerations*, gives however a unique and transparent formulation' (Segal 1959, 343; italics mine).

Against the ideal of pristine interpretation, I will argue that the best way to support QM_∞'s explanatory aspirations is give up on the idea that a theory's laws on their own, or even in concert with duly general mathematical, metaphysical, methodological, etc., considerations, delimit what worlds are possible according to that theory. Instead, those worlds should be characterized in different ways for different extranomic (factual or material or explanatory or maybe even practical) circumstances. Far from unimodal, physical possibility fractures on these circumstances into a kaleidoscope of varieties, varieties indexed the extranomic contingencies that condition them. The details of particular settings matter in something like the way laws are generally supposed to matter: they matter to what generalizations hold (and perhaps also to what properties are appropriate for involvement in generalizations) and thus to what explanations those generalizations support. On this view, contingency adulterates the delimitation of physical possibility. The philosopher's task is accordingly adulterated. It's to notice which contingencies matter, to understand how they matter, and to try to decide whether their mattering makes a difference to characteristically philosophical concerns.

My argument supposes that a theory ought to be interpreted in a way that enables its explanatory capacity. The nature and significance of a theory's explanatory capacity is a central bone of contention in the scientific realism debate, and an interpretation of a theory is what a realist about that theory believes. So it shouldn't be surprising that my discussion of QM_∞ has a collateral consequence for the realism debate. It is that there often isn't a single interpretation under which a theory enjoys the full range of virtues realists are wont to cite as reasons for believing that theory. The theory manifests different virtues under different ways to characterize its set of possible worlds. One might be tempted to count this as a strike against realism — evidence of an "underdetermination of interpretation by theory" exacerbating whatever underdetermination of theory by evidence may obtain. But I think it points to a mistaken attitude toward interpretation, an attitude with which most parties to the realism debate operate.

Reflecting the prejudice Rawls bemoans, the attitude is that interpretation is an afterthought. Once the lofty generalist decides whether or not we should believe successful scientific theory T, the narrow particularist tells us what we believe (or not) when T is set equal to (say) QFT. One problem with this attitude is that arguments which look compelling when conducted about T in lofty abstraction can disintegrate when T is instantiated as a genuine, interpreted theory. Another problem is that accounts of theoretical virtues (such as empirical adequacy, explanatory reach, internal and external coherence, continuity with future science, and so on) that adhere cleanly to T considered in abstraction may disintegrate when T is instantiated and interpreted. The lofty debate is conducted in terms that obscure how real theories possess the virtues they possess. It is often a theory *under an interpretation* that predicts, explains and promotes understanding. To the disappointment of the realist, there may not be a single interpretation under which a given theory accomplishes all those things. But it is to misunderstands this circumstance to treat it as grounds *against* epistemic commitment to a theory. A theory that underdetermines its own interpretation can be capable of a sort of semantic indecision that's a scientific resource, a ground for commitment to that theory. Or so I will contend.

The balance of this prefatory chapter reviews a lofty account—the best lofty account I can muster—of the interpretation of physical theories. It also sketches QM_∞ in sufficient depth to suggest how its exploration could refine and redirect this account, and with it, our thought about the nature of theories and theoretical virtue.

1.2 Interpreting physical theories

This section reviews an account of what it is to interpret a physical theory that I take to be widespread, although often implicit. Explicit variations on the account are most often found in the course of discussions of the model-theoretic "semantic view" of theories. I greet this circumstance with trepidation. There is a vast literature evaluating the semantic view, and another vast literature examining models in the sciences. I fear that my invocation here of the notion of model will be taken for a stand on issues debated in those literatures. I don't mean to take a stand; I mean only to help myself to the two uncomplicated and appealing (if vague) ideas that mathematical physics is formulated in terms of mathematical structures, and that mathematical structures can have physical instantiations.

The account of interpretation starts with the idea, shared in some form by philosophers as various as Ludwig Wittgenstein and David Lewis, that the content of a proposition is given by the set of possible worlds of which that proposition is true. The idea is widely embraced because there is much to recommend it. After all, someone who understands a proposition—who grasps its content—is able, at least in principle, to recognize when it obtains, as well as to distinguish the states of affairs that make it true from those that falsify it. Extended in a straightforward way to theories entire, the idea becomes

> [The Standard Account] The content of a theory is given by the set of worlds of which that theory is true.

This idea is sufficiently deep-seated and widespread that I've called it the *standard account*.

A testament to the deepseatedness of the standard account is how squarely it frames projects central to metaphysics and philosophy of physics. A basic metaphysical concern is how to zone the space of *logically possible* worlds. Take a logically possible world to be a maximal set of consistent propositions.[7] A significant zoning restriction is to consider only those possible worlds consistent with propositions stating the laws of nature. The zone so defined is the space of *nomologically possible* worlds.

[7] I'll soon admit that the propositional calculus is an inappropriate vehicle for the content of physical theories.

There are sweeping and homely ways for philosophers of science to concern themselves with nomologically possible worlds. The sweeping way addresses the principles behind the zoning restriction, by articulating and evaluating general accounts of laws of nature. Van Fraassen describes the homely way:

> When we come to a specific theory, there is an immediate philosophical question, which concerns the content alone: *how can the world possibly be the way this theory says it is?* This is for me the foundational question *par excellence*. (van Fraassen 1989, 193)

Take some physical theory T, and someone who can't say what laws of nature (in general) are. This someone might nevertheless have enough of a handle on T to embark on the project of characterizing and individuating the worlds of which T would be true—the task (van Fraassen remarks) of equipping T with content.

To characterize the worlds of which a theory T would be true is to supply as good an account of possibility-according-to-T as anyone could demand. It is to explicate the genus of nomological possibility that is physical possibility, with T acting as the species. It is also to answer on behalf of T what van Fraassen calls '*the question of interpretation*: Under what conditions is the theory true? What does it say the world is like?' (1991, 242). To interpret a physical theory is not only to equip it with content but also to explicate the notion of physical possibility allied with it.

Interpreting a theory articulates its truth conditions, but it doesn't follow that one has to be committed to the truth of theories to care about interpreting them. Van Fraassen is agnostic, and he cares. So might you, if you think interpretation promotes understanding of the theory (presuming that understanding is afforded by the grasp of truth conditions), or articulates its explanatory potential (presuming that a theory explains a phenomenon by situating it in the space of possibilities its laws allow), or serves what Sellars identifies as the philosopher's characteristic task of 'understand[ing] how things in the broadest possible sense of the term hang together in the broadest possible sense of the term' (1963, 37) (presuming, as Sellars does, that among the hanging things are our best scientific theories). Describing what the world would be like if the theory were true hardly commits one to believing that it *is* true; characterizing a theory's possible worlds or models hardly commits one to claiming that the actual world is among them. The old fashioned realist takes on these extra commitments: believing that the theory is true, she numbers the actual world among its models.

On the standard account, then, to interpret a theory is to characterize the worlds possible according to it. These possible worlds are (i) models (in something like the logician's sense) of the theory, and (ii) characterized as physical. I'm not going to try to say exactly what (ii) amounts to, although I hope it's clear that models characterized in terms of configurations of matter in spacetime count, and that models characterized in terms of sequences of natural numbers don't (unless, of course, they are accompanied by an informative account of why we should regard sequences of natural numbers as physical).

It is essential to understand that, in the present sense of interpretation, *The vast majority of the theories philosophers of physics dwell upon are already partially interpreted.* Otherwise they wouldn't be theories of *physics*. These theories typically come under philosophical scutiny already having been equipped, by tradition and by lore, with an interpretive core almost universally acknowledged as uncontroversial. In the case of General Relativity, this received core includes Einstein's field equations, along with the idea that their solutions describe space and time, energy and matter; in the case of QM, this received core incorporates canonical commutation relations, along with the quantization and statistical algorithms. There may also be, by tradition and by lore, theoretical foci of longstanding disagreement: the interpretation of the Born rule, say. And sustained philosophical attention to a theory is likely to trigger additional interpretive disputes that arise along with the attempt to extend and refine its received core: disputes about how exactly a substantivalist should individuate models of the Einstein field equations, say.

That the theories philosophers care about arrive already partially interpreted explains how the pristine interpreter can take a theory's laws to constrain its interpretation: laws can be part of what's left in the received core of a theory when you strip away what different interpreters disagree about. I think that partially interpreted theories can explain, predict, and so on; but I also think (and will try to argue in what follows) that sometimes a theoretical explanation's *bona fides* can be secured *only* by saying more about its interpretation than the received core does.

We can distinguish two phases in the interpretation of a physical theory (see (van Fraassen 1989, Ch. 9), or (Suppe 1977, 221-230)). One phase is *structure-specifying.* In this phase, the interpreter characterizes the structures by which the theory would represent physical reality. For example, the interpreter of general relativity might nominate structures of the form $(\mathcal{M}, G_{ab}, T_{ab})$, where the first entry is a manifold and the second and third entries are tensor fields defined on that manifold and satisfying Einstein's Field Equations. (Notice that the presentation proceeds with less formality than is generally demanded of axiom systems admitting models in the logician's sense. This is entirely appropriate. The philosopher of physics interprets physics, not logic. Notice also the structure-specification is an interpretive enterprise in which physicists, qua physicists, engage. This underscores the observation of the last paragraph, that physical theories come to philosophers' attention already partially interpreted.) In the other, *semantic*, phase of interpretation, the interpreter identifies the physical worlds which model the theory so structured. So our sample interpreter might declare that a collection of spacetime points conceived as substances and bearing properties encoded (in a way the interpreter is beholden to elaborate) by the stress energy and curvature tensors instantiates the structure $(\mathcal{M}, G_{ab}, T_{ab})$. Moving from the structure-specifying through the semantic phase, the interpreter characterizes, and characterizes *as physical*, the worlds possible according to the theory. We will see that, like thermodynamic phases, the semantic and structure-specifying phases of interpretation can intermingle and coexist.

For the myriad physical theories admitting a Hamiltonian formulation, the structure-specifying phase of interpretation decomposes neatly into three elements (see, again, van Fraassen or Suppe, *op. cit.*). The first element is a specification of the theory's *state space*. In the mercifully simple case of the classical theory of a particle of mass m moving in one dimension, the state space is comprised of ordered pairs (q, p) of position and momentum values. The second element is a specification of the set of physical magnitudes or *observables* the theory recognizes. In the same simple case, observables are functions from elements of state space to the real numbers $\mathbb{R}$; e.g., the physical magnitude *kinetic energy* E is given by the function $f_E(q, p) = p^2/2m$. Together, these two elements constitute a *kinematics* for the theory.

The third element of structure specification is the theory's *dynamics*, that is, its account of the time development of states and observables. In classical mechanics, dynamics takes the form of trajectories through state space determined by Hamilton's equations, once a Hamiltonian function

$H(q, p)$ is provided. According to this structure specification for classical mechanics, the structures by which the theory would represent physical reality are state spaces with dynamical trajectories imposed.[8] Insofar as these trajectories are all instances of a dynamical scheme furnished by Hamilton's equations, the structures form 'a family of state space types' (van Fraassen 1989, 223), with the Hamiltonian scheme supplying the kinship relation. 'Thus theories are structures, these structures being phase spaces with configurations imposed on them in accordance with the laws of the theory' (Suppe 1977, 227).

In the *semantic* phase of interpretation, the project is to characterize models of the theory—instantiations of the structure specified—*in physical terms*. Straightforwardly construed, a semantics can take the form of an account of which propositions attributing determinate values to magnitudes recognized by the theory are true of a system represented by a state of the theory. The characterization qualifies as physical if we have grounds for commitment to the status of (some of) the magnitudes in question as physical. A straightforward semantics for the simple classical theory is: *"Observable O has value x in state (q, p)" is true just in case* $f_o(q, p) = x$. But semantics needn't be so straightforward. It can, for instance, be conducted in terms exogenous to those used to present the theory structure. Terms supplied by antecedent ontology are an example, as when an interpreter of QFT takes Fock space structures to describe possible *particle* configurations.

Let us summarize the standard account: to interpret a physical theory is to characterize the worlds possible according to that theory. Two phases of this characterization can be distinguished. One phase identifies the theory's structures: its states, observables, and dynamics. The other characterizes the physical situations that count as models of the theory so structured. Interpretation is an exercise in nomic articulation: a theory's laws guide the characterization of its possible worlds; the interpretation of a theory is at the same time an explication of the notion of nomological possibility allied with the theory.

Understanding interpretation as an exercise in nomic articulation enables us to notice one way the ideal of pristine interpretation is reflected in

[8]Each trajectory corresponds to what I've been calling a "world" possible according to the theory.

the laws of nature literature.[9] An interpretation identifies the worlds possible according to a theory. So do that theory's laws. According to most accounts, laws of nature are (or imply) generalizations that aren't accidental. The intuition is that "All unsupported bodies in the vicinity of the earth fall toward it" is, or has something to do with the, nomic; "all the coins in my pocket are dimes" does not. John Carroll offers a diagnosis of what's un-nomic about propositions like the latter: 'What blocks it from being a law is that something in nature, or really a certain sort of initial condition of the universe, an absence of something in nature, explains the regularity' (Carroll 2007, 74). Refracted through this concern with distinguishing real laws from pretender generalizations, the idea that a theory's laws determine what's possible according to that theory, assumes a different form. It becomes the idea that what's possible, according to a theory shouldn't depend on accidental particulars, like actual initial or boundary conditions. But such non-dependence is exactly what the commitment to unimodality requires, and the ideal of pristine interpretation pursues. By allowing only considerations of metaphysics and theoretical structure to guide the identification of the worlds possible according to a theory, the ideal insulates that identification from the influence of contingencies and accidents.

1.3 In praise, and in defense, of the standard account

The practice of philosophers of physics speaks in favor of the standard account of interpretation. Most things philosophers of physics count as interpretive questions are interpretive questions in its sense. A question about a physical theory's state space: does the state of a quantum system include, *a la* Bohm, a determinate configuration? A question about the observables appropriate to a physical theory: must the observables of general relativity be generally covariant, and if so, in what sense? A question about the dynamics of a physical theory: can a full Newtonian dynamics be formulated on a reduced phase space encoding information only about *relative* positions and velocities? A question about a physical theory's semantics: do all propositions attributing determinate values to observables pertaining to quantum systems have truth values, or just some of them? If the latter, which ones? The list could be extended. The standard account

[9] A useful recent anthology is Carroll 2004.

of interpretation enables us to regard its entries as different expressions of a single enterprise.

The words of some philosophers of physics speak against the standard account. Matthias Frisch (2005), among others (see, for example, the essays in (Morgan and Morrison 1999)), has recently challenged the idea that theories represent by picking out sets of possible worlds as models of their fundamental mathematical structures. But I'm not ready to give up on that idea yet. I take Frisch's consideration of classical electrodynamics to establish that the theory is not simple but manifold, that conjoining its different variations can induce inconsistency, and that physicists blithely ignore this as they opportunistically leap from one formulation to another as the problems they confront demand. I am not hostile to a single one of those points. Indeed, I will try to establish their QM_∞ analogs in the balance of this book. But I do not conclude from them that classical electrodynamics is inconsistent, and so unsusceptible to interpretation in the standard sense. Rather, I conclude that electrodynamics admits multiple interpretations in the standard sense, and celebrate the opportunity this affords philosophers to articulate and adjudicate those interpretations, as well as to reflect on what their presence might be telling us about physics.

1.4 Criteria of adequacy for interpretations

This section announces some criteria of adequacy for interpretation as it is standardly understood. These criteria are provisional; they're meant to help initiate, not terminate, the interrogation of interpretive options. The exigencies of interpreting QM_∞ may induce us to revise our sense of what it takes to comprehend the physical world by means of an empirical theory. To my mind, the capacity of interpretive projects to inspire such revision is part of what makes them philosophical.

In "The Theoretician's Dilemma," Hempel (1965) considers the case for taking the content of a scientific theory to be its "Craigified" part—that is, the connections it draws between observable phenomena, connections presented without the intermediary of (nonlogical) theoretical apparatus. Hempel resists the Craigification of theoretical content because he attributes theories a function which he considers their Craigified parts ill-suited to fulfill. That function is inductive systematization. What promotes it, Hempel suggests, is the systematic unification between observable phenomena effected by the theoretical apparatus.

Without going deeper into the details of Hempel's position, I claim it as an inspiration for one criterion of interpretive adequacy. Supposing that explanation is a central function of physical theories, the criterion is: *An interpretation should enable the theory it interprets to discharge its explanatory duties.* I don't need—I'm not sure there exists—a categorical imperative for identifying these duties. In applying the criterion, we can start from the explanatory aspirations of physicists using a theory, and aim for a reflective equilibrium between the assessment of interpretations for sustaining those aspirations, and the assessment of those aspirations as appropriate.

"Explanatory duties" is not only vague but also synechdochical. There are a host of things a theory ought to do or be, and an interpretation of a theory is adequate insofar as it enables the theory to do or be these things. An interpreted theory ought to be logically consistent, empirically adequate (if we think the theory is), and maybe even simple and mathematically elegant. What's more, an interpreted theory ought to make sense of experimental practice. Huge tracts of the literature on the interpretation of QM concern whether it is possible to give an interpretation of quantum mechanics consistent with the empirical predictions of the minimally interpreted quantum statistical algorithm. Still more of this literature tackles the measurement problem, which constrains interpretation to secure the fundamental presupposition of laboratory practice, that experiments have outcomes.

I've been proceeding as though each theory were an island, complete unto itself. In this mode of address, adequacy of the sort under discussion is a purely internal affair: an interpretation ought to equip a theory with content sufficient to that theory's own aims. But externally-oriented questions of sufficiency arise as well. We may, for instance, want to interpret a theory in a way that makes sense of its fit with environing theories and their interpretations. Thus in the case of the thermodynamic limit of QSM, we might hope for an interpretation that brings the theory into a substantial explanatory relationship with thermodynamics. Again, existing theories are conceptual resources from which new theories are forged. Thus we might want an an interpreted theory to serve as such a resource. Semiclassical quantum gravity gives a good example of how this desire plays out in the interpretation of QFT.

Semiclassical quantum gravity considers a quantum field on spacetime manifold subject to Einstein's field equations, with a quantum commodity (the expectation value $\langle T_{ab} \rangle$ of the stress energy tensor T_{ab}) substituted for a classical one (plain old T_{ab})[10]:

$$G_{ab} = 8\pi \langle T_{ab} \rangle \tag{1.1}$$

One point of this hybridization is to glean from it what a purebred quantum theory of gravity might look like. An interpretation of QFT that withheld physical significance from a stress energy observable (or something like it) would blunt that point. It would fail to recognize enough observables to sustain QFT in this developmental task.

If the foregoing criteria of adequacy had a bumper sticker on their car, that bumper sticker would say:

Physically significant is significant for physics.

To interpret a physical theory is to delimit what that theory recognizes as physically significant. The criteria demand the delimitation to respect how the theory matters to the present and future of physics.

The criteria discussed so far all fall roughly under the heading of 'sufficiency.' The demand that an interpretation of a theory say *enough* about the world of which that theory is true to make sense of the theory as successful science. Should we also demand interpretations not say *too much*? Although intuitively appealing, the demand is difficult to motivate ecumenically. It is preaching to converted Ockhamists to declare parsimony desirable in itself. Perhaps it helps to observe that parsimony can serve recognized epistemological and metaphysical goals. For instance, there is a failure of parsimony that frustrates the venerable goal of *determinism*. For a theory to have any shot at determinism, its dynamics must specify how the values of some set of magnitudes at one time depend on the values of some set of magnitudes at another. Unparsimoniously recognizing more magnitudes than are dynamically tractable, or more states than can be put into one-to-one correspondence with valuations of dynamically tractable magnitudes, the structure-specifying phase of an interpretation

[10] Good sources to consult for a review of the basics of differential geometry and its use in spacetime theories include the appendix to Friedman (1983) and Wald (1984). Instead of a comprehensive formal review, I will issue informal sketches of key ideas as they come up, and trust to context and prose explanations to orient the uninitiated reader through the rest.

can threaten determinism. Of course, its semantic phase could work to defuse this threat, as some defenses of substantivalism against the hole argument do (see, for example, Brighouse 1994).

1.5 Realism and Pristine Interpretation

There is a loftier way to motivate the 'sufficiency' criteria just announced. It is to adopt the desideratum

> An interpretation of T should, as far as possible, attribute T the virtues realists are wont to cite as reasons for believing that T is true or approximately true.

The desideratum motivates the 'sufficiency' criteria because these virtues include inductive systematization and other varieties of empirical success, invoked by the realist's Miracles Argument, as well as continuity with successor theories, invoked by standard realist replies to the pessimistic meta-induction (see Psillos (1999) for a recent anatomy of scientific realism debates). Adopting the desideratum above preserves a bridge between generalists and specialists in the form of the idea that an interpretation of T is what a realist about T believes. But it preserves that bridge in a way that neither prejudges the outcome of the generalists' debate nor implies the insulation of that debate from particularist inquiries. Even supposing T admits an interpretation outfitting it with a full panoply of virtues, antirealists can refrain from regarding those virtues as reasons for belief. And particularists' close examination of an instantiation of T may reveal that 'as far as possible' isn't very far at all: T may exhibit different virtues only under different interpretations.

We are in a position to exhibit a connection between realism and pristine interpretation. Suppose that in different settings, T exhibited different virtues salient to the realist. Applied to theoretical high energy physics, it afforded the inductive systematization implicit in the notion of 'particle' and exhibited consistency with another regnant theory; applied to quantum optics, it sacrificed some of that consistency for the sake of joining *other* regnant physical theories in the task of modelling collider phenomena (later chapters try to put some meat on the bones of these suppositions). And suppose further that the interpretation explicating the virtues T exhibits in one setting is different from the interpretation explicating the virtues

T exhibits in another. Then there isn't a single interpretation of T under which it exhibits all the virtues realists regard as reasons for believing T. The impediment this presents too realism is: to be wholly virtuous, T requires different interpretations in different settings. Pristine interpretation is a strategy for negotiating this impediment. The ideal of pristine interpretation is to ignore details of T's particular applications—details that are merely consistent with but not compelled by T's laws— when identifying the worlds possible according to T. Pristinely interpreted, T won't admit different interpretations under different settings.

There is no guarantee that the strategy will succeed. It may be that under no pristine interpretation does T emerge as wholly virtuous. But despairing of pristine interpretation is frightfully close to despairing of warranted realism. If T only exhibits *some* of its virtue in each setting to which it applies, unpristinely keying T's contents to its circumstances of application guarantees that there is no single interpretation under which T possesses all its virtues.

1.6 Interpreting QM_∞

Here's one reason to think QM_∞ could use some interpretation. What I'll call theories of *ordinary QM* concern systems of finitely many particles in Euclidean space. There is a standard notion of what a quantum theory of such a system requires. A quantum theory requires a *Hilbert space representation*, that is, a Hilbert space, on which act symmetric operators obeying relations characteristic of the system quantized—in general, canonical commutation relations, aka CCRs, (for mechanical systems) or canonical anticommutation relations, aka CARs, (for spin ones). Possible states for the quantum system are density matrices on the representing Hilbert space; observables pertaining to the system are self-adjoint operators on the representing Hilbert space. Most interpreters of ordinary QM take quantum kinematics, in the form determined by a Hilbert space representation, as their point of departure.

After that, they typically diverge. They disagree about whether to supplement these kinematics' bare quantum states with hidden variables. They disagree about whether quantum *dynamics* (that is, the time evolution of quantum states and observables) is collapse-ridden. They disagree about whether a quantum system can exhibit a determinate observable value its state cannot predict with certainty. They disagree about whether quantum reality comprises one world or many, and about the role in its constitution of minds, the environment, and the biorthogonal decomposition theorem. But amid all these disagreements, it doesn't occur to interpreters of ordinary QM to worry that they're not even talking about the same Hilbert space structure to begin with.

Interpreters of ordinary QM don't worry about this because the Stone-von Neumann theorem suggests that they needn't. The theorem states that all Hilbert space representations of the CCRs for a particular classical Hamiltonian theory of finitely many particles in Euclidean space stand to one another in a mathematical relation called *unitary equivalence*. Unitary equivalence is widely accepted as a standard of physical equivalence for Hilbert space representations. It follows that all these representations are simply and unalarmingly different ways of expressing the same quantum kinematics. For finitely many spin systems subject to CARs, the Jordan-Wigner theorem likewise guarantees uniqueness. For systems of finitely many particles, the directive "quantize!" has a unique outcome. That is why many philosophical discussions of QM have proceeded in the scope of the assumption that the basic kinematics of a quantum theory takes the form of a Hilbert space representation.

One of my missions here is to chronicle the disintegration of that assumption in the context of QM_∞, and the consequent disruption to an interpretive landscape familiar from philosophers' discussions of ordinary QM. The uniqueness theorems do not extend to QFTs, which one obtains not by quantizing a system of finitely many particles, but by quantizing a field, an entity defined at every point of space(time). Neither do they apply to the thermodynamic limit of quantum statistical mechanics, where the number of systems one considers and the volume they occupy are taken to be infinite. This brings us to the provocative feature of QM_∞, not shared by ordinary QM. *According to very same criterion of physical equivalence by whose lights Hilbert space representations of the CCRs/CARs for ordinary quantum theories are reassuringly unique, the CCRs/CARs of a QM_∞ theory can admit infinitely many physically inequivalent Hilbert space representations.*

Something has gone terribly wrong—but what? Are we mistaken about what a quantum theory is, or about when Hilbert space representations are physically equivalent? Has some crucial aspect of what it is to be *physical*—a consideration that would alleviate the apparent non-uniqueness—escaped our notice? Was it fundamentally misguided to expect "the theory of the mass m free boson field" (and other terms presuming to denoting particular theories of QM_∞) to have a single referent?

These are among the questions I aim to explore here. My means of exploration will be the articulation and evaluation of a variety of accounts of the content of theories of QM_∞ in a variety of circumstances where unitarily inequivalent representations arise. I will try to argue that no overarching account—no single interpretation of QM_∞—is adequate to all of those circumstances. I think that this should attenuate our commitment to the ideal of pristine interpretation, as well as complicate our notion of physical possibility and our conduct of the scientific realism debate, and I will try to suggest how.

The plot of the book is as follows. Chapter 2, "Quantizing," reviews the procedure of Hamiltonian quantization, explicates the Stone-von Neumann theorem as a theorem about the uniqueness, up to unitary equivalence, of quantizations thus obtained, and identifies the presuppositions underlying the reception of this result as a proof of the *physical equivalence* of those quantizations. These are assumptions about the shape of quantum theories and the nature of physical equivalence. Chapter 3 shows how theories of QM_∞ and other sorts of extraordinary QM escape the clutches of the Stone-von Neumann and Jordan-Wigner theorems to admit unitarily inequivalent Hilbert space representations—sometimes continuously many of them.

Suggestively, there is a level of abstraction at which even unitarily inequivalent Hilbert space representations of a theory of QM_∞ share a common structure, the structure of an abstract algebra. The aim of Chapter 4, "Representation without Taxation," is to introduce the technicalia needed to ascend to this level of abstraction. The introduction is informal, and meant to be accessible to readers who encountered QM through its philosophy.[11] Thus I will assume readers to have a working acquaintance with vector spaces, operators, and the so-called axioms of ordinary non-relativistic QM, but I won't presuppose systematic mathematical knowledge beyond that. My governing principle is that if you can be familiar with standard philosophy of QM texts without knowing about X, I should at least gistify X for you. (To minimize tedium and annoyance for readers with more extensive mathematical backgrounds, some gistifications will be exiled to

[11] For instance through expositions such as (Redhead 1980), (Hughes 1989), or (Albert 1994).

Scholia or footnotes.)

Once the basic notions of algebraic quantum theory have been introduced, they can be used to characterize interpretations of QM_∞ and the challenges they face. Chapter 5 presents a preliminary overview of candidate accounts. These range from "Hilbert Space Conservatism" which continues to insist, in the face of QM_∞'s superabundance of unitarily inequivalent representations, that theories of QM_∞ take the same form as theories of ordinary QM, and are physically equivalent only when unitarily equivalent; to "Algebraic Imperialism" (so labelled by Arageorgis 1995), which identifies the content of a theory of QM_∞ in terms of the abstract algebraic structure common to all its concrete Hilbert space representations. Both the Conservative and the Imperialist promise to adhere to the ideal of pristine interpretation: their identification of the worlds possible according to a theory of QM_∞ rests on considerations so general that the adulterating option, of contouring the theory's content in response to non-"nomic" details of its applications, is suppressed. Chapter 5 unkindly labels such pristine interpretive positions "extremist."

The algebraic apparatus developed in Chapter 4 makes visible a startling mathematical possibility. The backdrop against which it emerges as startling is provided by a family of pristine approaches to ordinary quantum semantics, a family which includes most familiar interpretations of QM. Advocates of these *maximal beable strategies* take the task of quantum semantics to be the specification of the largest set of propositions concerning a quantum system that can be attributed simultaneously determinate truth values *obedient to classical truth tables*. Gripped by different metaphysical scruples and technical insights, different versions of the maximal beable strategy identify this "maximal beable algebra" in different ways. But for each, the identification hinges on *a privileged set of pure states* of the system under scrutiny (e.g. possible endpoints of collapse, states of determine position, "value states"). The startling mathematical possibility which emerges from Chapter 4 is the possibility of algebras of quantum observables which admit no pure normal (that is, countably additive) states. Chapters 6-7 continue Chapter 4's unpunishing exposition of algebraic notions by explaining how such algebras of observables come to be instantiated in QM_∞, and how they stymie maximal beable strategies that apply so well to ordinary QM. The upshot is that QM_∞ rewrites what serve ordinary QM as the usual rules for quantum semantics.

Chapters 8-13 confront candidate interpretations with QM_∞ phenomena they might be expected to save. Chapter 8 asks, "Is Particle Physics Particle Physics"[12] That is, should we understand QFT as a theory about particles? A notorious case against rests on the Unruh effect, according to which an observer accelerating through Minkowski spacetime in its QFT vacuum state—a state distinguished by the absence of particles—finds herself bombarded by . . . particles. Chapter 8 dissects this case against, to reveal the working therein of unitarily inequivalent representations of the quantum field theoretic commutation relations. Chapter 9 characterizes the circumstances conducive to a particle interpretation, and argues that many QFTs meriting interpretation fall outside thsoe circumstances.

This is already a refutation of a species of extremism that would subject *every* theory of QM_∞ to a particle interpretation. Chapter 9 also targets the extremism of Hilbert Space Conservatism by producing examples of theories of QM_∞ with the following feature: different applications of the theory are explicated by different (and ergo competing) Conservative interpretations of it. Pristinely pursued, Conservatism undermines the empirical reach of such a theory.

Chapter 10 is one of the few places in the book that makes contact with the QFTs high energy particle physics test with scattering experiments. (Chapter 13's treatment of broken symmetry in QFT is another place.) It investigates a route from the physicists' techniques of calculation and detection—techniques which employ eerily particulate Feynman diagrams and cloud chamber tracks — to a particle interpretation of QFT. It travels that route only long enough to conclude that however such a "phenomenological particle interpretation" might be executed, it will differ in fundamental respects from particle interpretations of the sort at issue in Chapters 7 and 8.

Chapter 9 suggested that there's a particular extremist, the Hilbert Space Conservative, who can't interpret a particular theory of QM_∞ in a way that explicates all its empirical virtues. Chapters 11-13, which treat phase structure and spontaneous symmetry breaking in QM_∞, argue for a more radical conclusion: there occur in QM_∞ *individual explanations* that *no* extremist can make sense of. Chapter 11, "Putting Unitary Inequivalence to Work," presents phase structure—understood as the existence, at a single critical temperature, of distinct equilibrium states, corresponding to distinct pure thermodynamic phases available to a system in equilibrium at

[12]Some of this chapter is based on Arageorgis, et al, 2003.

that temperature—as an *explanandum* that might require an *explanans* no extremist interpretation can sustain. To construct an interpretation that makes sense of phase structure, I suggest, we should suspend an assumption, shared by all extremisms, about how to specify the content of physical theories. This is the unimodality assumption that physical theories sort states of affairs into the unqualifiedly possible and the unqualifiedly impossible, the assumption motivating the ideal of pristine interpretation. Using examples from QSM, I develop principles for *adulterated* interpretation, ones that take some of the mystery out of the suggestion that what's possible according to a theory could vary along with its circumstances of application.

Chapter 11's case that no extremism can sustain the explanations considered there is dubbed "the Coalesced Structures Argument." This argument is liable to be met with an immediate, and apparently devastating, objection. The objection is that the thermodynamic limit, which inspires the Coalesced Structures Argument, is a rank idealization from which no serious foundational conclusions can be drawn. Bluntly put, the objection is that the Coalesced Structures Argument is irrelevant because steaming cups of coffee are finite. Chapters 12 and 13, which treat broken symmetry in QSM and QFT respectively, start out as an end-run around this objection. Broken symmetry in QM_∞ shares with phase structure many of the features on which the Coalesced Structures Argument turns. Moreover, broken symmetry is reputed to characterize theories of QM_∞ that come by their ∞ honestly, for instance, the QFTs allied into the Standard Model. The end run Chapters 12 and 13 attempt is to devise a QFT analog of the Coalesced Structures Argument. The attempt does not fail so much as evaporate into speculation, for the simple reason that the QFTs that are reputed to break symmetry aren't formulated explicitly enough to tell whether or how the concepts at play in the Coalesced Structures Argument apply. But this observation prompts another about the apparently devastating objection to the phase argument: it rests on a *guess* about the future of physics. So too does my willingness to rest foundational conclusions on the Coalesced Structures Argument. Chapter 13 closes by explaining why I like my guess better.

I claim in this book that 'extremist' interpretations of QM_∞ fail to make sense of the full range of empirical successes those theories enjoy, but that claim is not the point of this book. The point of this book is how different the debate over scientific realism looks in the wake of the claim. Chapter 14, "Re: Interpreting Physical Theories," develops the point. It brings the particularist labors of the previous dozen chapters home to roost in the generalists' debating halls. It tries to persuade the generalist that he should care about the presence and role in QM_∞ of unitarily inequivalent representations, by describing the lessons the present particularist sees in QM_∞, lessons about the natures of interpretation, physical possibility, theoretical virtue, and grounds for warranted belief in our best scientific theories.

Chapter 14 also describes a subset of the things I wish I could figure out but haven't managed to. My hope (which is founded on the possibly unreasonable supposition that some readers will make it as far as Chapter 14) is that some kind readers will help.

.