Unifying Scientific Theories

Physical Concepts and Mathematical Structures

MARGARET MORRISON
University of Toronto

CAMBRIDGE
UNIVERSITY PRESS

multiplication of free parameters in order to account for distinct phenomena. Rigidity, on the other hand, not only minimizes the number of free parameters in the theory's domain but also rules out the addition of supplementary theoretical structure as a way of extending the theory's evidential base. These requirements are definitive of the unifying process, but as such they have very little to say about the nature of scientific explanation.

My discussion of unification in the subsequent chapters is motivated not only by what I see as errors and omissions in current philosophical analyses of the subject but also by historical investigation of what exactly was involved in paradigm cases of unification in both the physical and biological sciences. I want to stress at the outset that my emphasis is on the process of theory unification, something I want to distinguish from a metaphysical or even methodological thesis about the "unity of science" or a "unity of nature". What I want to show is that the methods involved in unifying theories need not commit one to a metaphysics of unity, of the kind that, say, Kepler advocated. As we saw earlier, Kepler's mathematical physics was rooted in the corresponding belief that nature was harmonious; hence there was a kind of one-to-one correspondence between the mathematical simplicity of physical laws and the mathematical simplicity of nature. Although some might claim that the motivation for theory unification embodies a belief in something like Keplerian metaphysics, I want to argue that there are good reasons, despite the presence of unified theories, for thinking such a belief to be mistaken. It is perfectly commonplace to have a high-level structural unity within a theoretical domain in the presence of a disunity at the level of explanatory models and phenomena. In addition to the electroweak case, population genetics, which is discussed in Chapter 7, is a case in point.

The purpose of this overview has not been to set out particular accounts of unification as models for the cases I intend to discuss. My intention has rather been to present a brief sampling of some ways in which unity and unification have been characterized throughout the history of science and philosophy and to give some sense of the diversity present in accounts of unity. I have also attempted to lay some groundwork for my argument that unity and explanatory power are different and frequently conflicting goals. Undoubtedly, strands of each of the views I have discussed can be found in the examples I shall present, something that serves to illustrate my point, namely, that although unified theories themselves may share structural similarities, no hard and fast conclusions can be drawn from that about nature itself. This is partly a consequence of the methods involved in theory unification, but it is also due to the fact that unity in science and nature can take on many disparate and contradictory interpretations and forms.

2

Unification, Realism and Inference

Unity – Explan – Truth

The question that occupies most of this chapter is whether or not the first word in the title – unification – bears any relation to the other two, and if so, how that relation ought to be construed. As mentioned in the introductory remarks, a common approach to fleshing out the notion of unification is to link it to explanation. A unified theory is thought to be one that can explain phenomena from different domains by showing either that the phenomena are essentially the same (e.g., light waves are simply electromagnetic waves) or that diverse phenomena obey the same laws, thereby suggesting some link between them. This explanatory power supposedly provides good evidence that the theory is true; hence, the best explanation, which typically will be the one that reveals some unity among the phenomena, should be seen as more likely to be true than its competitors. Of course, not all "best explanations" will perform a unifying function. There may be only one explanation of a particular phenomenon, and hence, by default, it will have to be considered the best. So embedded in the debate are two issues, one linking unity to explanatory power, and the other linking the concept of "best explanation" to increased likelihood of truth. This practice of drawing inferences to truth on the basis of explanatory power has been dubbed "inference to the best explanation" (IBE) and has been advocated by, among others, Harman (1965) and Thagard (1978).

More recently, however, there have been forceful criticisms by van Fraassen (1980), Cartwright (1983) and Friedman (1983) of the link between IBE and truth and its use as a methodological rule that forms the basis for inference. The complaints are varied. Some, particularly van Fraassen, emphasize the fact that explanation has to do with providing answers to "why" questions or organizing and systematizing our knowledge – pragmatic features that do not provide evidence for the literal truth of the background theory used in the explanation. Cartwright has argued that truth and explanation are, in fact, inversely related: Explanatory power requires broad general laws that do not accurately describe physical processes. But even for those who disagree about the pragmatic status of explanation or its relation to truth, the best available explanation may not be the one that we would want to accept, even provisionally. Friedman opposes IBE on the ground that it provides no guidance on the issue of whether we should construe theoretical structure literally or instrumentally. It simply fails to explain why theoretical structure should *ever*

Theoretician's dilemma

be taken literally. For example, consider two attitudes one might have toward the molecular model of a gas: Either one can be a realist and claim that gases really are just configurations of molecules, and the former can be reduced to or identified with the latter, or one can simply believe that the function of the kinetic theory is to supply a mathematical model for the observable behaviour of gases by associating gases or their properties with mathematical aspects of the model. In this case there is a mapping or correlation of the two domains, but not a literal identification; we have a representation, but not a reduction. We can think of the phenomenological and theoretical domains as being two structures $\mathcal{B}$ and $\mathcal{A}$. The realist sees the relation between these two as that of model to sub-model; $\mathcal{B}$ is a sub-model of $\mathcal{A}$, and hence the objects in $\mathcal{B}$ are identified with their counterparts in $\mathcal{A}$. The anti-realist, however, claims only that $\mathcal{B}$ is embeddable into $\mathcal{A}$; there is a mapping from one domain to the other, but no literal identification is made.

w/ don't answer

The important question, of course, is when to adopt one attitude rather than another. Part of Friedman's objection to IBE is that it provides no guidance on this issue. Regardless of whether we interpret theoretical structure as a mere representation of observable phenomena or as a literal reduction, we enjoy the same consequences vis-à-vis the observable realm. That is, we get the same explanations of the observable phenomena, the only difference being that the anti-realist says that the phenomena behave "as if" they were composed of molecules, rather than actually believing that to be so. In addition, we may have only one explanation of a particular phenomenon, one that might not be acceptable for a variety of reasons; nevertheless, if we apply the rule of IBE we are forced to accept it. Friedman's solution to this problem consists not in giving up this method of inference but rather in restricting its applicability. He argues that theoretical inference can be sanctioned when accompanied by unification, thereby linking unity, explanation and truth. Inference to the "unified explanation" is touted as superior because we get an accompanying increase in the confirmation value of the phenomena to be explained and greater confirmation than would accrue to the previously unconjoined (or non-unified) hypotheses. For instance, if we conjoin the atomic theory of molecular structure and the identification of chemical elements with different kinds of atoms, we can explain chemical bonding. This imparts more confirmation to the assumption that gases are simply molecular systems, a hypothesis that is also confirmed by the gas laws themselves.

or phenomena themselves?

Friedman provides persuasive arguments to suggest why one ought to be a realist about certain bits of theoretical structure that figure in the process of unification. Realism allows a literal interpretation of the relevant structure, which in turn affords our theories their unifying power and subsequently their confirmation. In other words, we simply cannot conjoin or unify hypotheses that we do not interpret literally, and, on his view, a literal interpretation requires realism. Without this ability to unify, there is no basis for increased confirmation and hence no basis for belief. Any theoretical structure not participating in unification can be

treated as purely representational, without any adverse consequences for the theory in question.

An important part of Friedman's programme is his characterization of the relationship between observational and theoretical structures as that of sub-model to model. For example, we can literally reduce (using an identity map in the model-theoretic characterization) the observational properties of gases to their molecular configurations, with the observable structure of kinetic theory construed as a sub-model of the larger theoretical structure (Friedman 1983, p. 240). Once the reduction is complete, we can then conjoin this theoretical structure with others to form a unified theory. Consequently, much of Friedman's discussion of unification focuses on the role of conjunctive inference and reduction in facilitating the process of unification.[1]

Before moving on to the details of Friedman's proposal, let me briefly mention what seem to be some obvious difficulties with the account. On the face of it, Friedman's argument seems problematic. He claims that we need realism for unification (1983, pp. 11, 244–5), but only if we have unification can we be realists (1983, pp. 7, 259–60). In looking at the details of his argument, particularly the association between unification and conjunction, one needs to distinguish between the logical and methodological aspects of conjunctive inference. Because any account of theory unification must respect, to some extent, the methodological practices of science, it is important to see whether or not the logical model of conjunctive inference as described by Friedman has any bearing on how theories are actually unified. I want to argue that neither a logical account nor a methodological account of *conjunction* is appropriate as a way of explicating unification.

A further criticism centres on Friedman's use of the model/sub-model approach to represent the relation between observational structure and theoretical structure. If the observable structure of our theories is to be construed as a sub-model of the larger theoretical structure (model) in the way that Friedman suggests, then it isn't immediately clear how we are to account for changes in the relationship between the observable and theoretical structures that result from theory evolution. That is, how do we relax the requirements on the identity relation to accommodate changing views on the nature of theoretical structure? Moreover, it is frequently the case that a particular empirical law (e.g., Boyle-Charles) and its approximations (the van der Waals law) require different theoretical models, depending on the context of application. My point here is simply that the ideal-gas law presupposes a different model of the molecular system than does the van der Waals law, and if we utilize the identity map in facilitating theoretical reductions, then we are committed to a unique interpretation of how the molecular system is constituted. That, in turn, makes it difficult to envision how different and sometimes incompatible aspects of theoretical structure are to be accounted for. Friedman claims that his approach accounts well for the way theories get confirmed over time, but the historical record shows that as the kinetic theory evolved, the molecular models underwent drastic changes, with the behaviour of gases often being explained by a variety of different and seemingly incompatible models of the phenomena. In light of these kinds of considerations it is difficult to see how the model/sub-model approach is an effective way to characterize scientific theories; the logical constraints on the model/sub-model relationship (the identity condition) are too tight to allow for the kind of looseness of fit that exists between the theoretical and observable structures of theories. In contrast, not only does the embedding account that Friedman attributes to the anti-realist provide a more accurate characterization of scientific practice, but also it need not fall prey to the criticisms he advances against it.

Finally, Friedman emphasizes the connection between his views on unification and those of Whewell on the consilience of inductions. However, a traditional Whewellian view incorporates a great deal of conceptual change in the evolution from one theoretical picture to its successor, something that cannot be accommodated on the simple "logical" picture of theory conjunction that Friedman puts forth.[2]

?

In the end, however, regardless of whether or not Friedman's model of theory unification is viable, there are independent reasons for rejecting unification as a justification for scientific realism or a basis for inference. Of course, this still leaves open the question of whether or not unification should be equated with explanatory power. That issue will be taken up in some of the remaining chapters.

2.1. The Friedman Model

A typical scientific explanation is characterized in the following manner: We postulate a theoretical structure $\mathcal{A} = \langle A, R_1, \ldots, R_n \rangle$ (where A is the domain of individuals, and $R_1, \ldots, R_n$ are physical relations defined on A) possessing certain mathematical properties. We also have an observational substructure $\mathcal{B} = \langle B, R'_1, \ldots, R'_m \rangle$ $(m \leq n)$, and $\mathcal{A}$ functions as an explanation or reduction of the properties of $\mathcal{B}$. Using the kinetic theory, we can explain the observable properties of gases characterized by $\mathcal{B}$ by embedding them in $\mathcal{A}$, where $\mathcal{A}$ is literally construed as the world of molecular theory. This enables us to account for the behaviour of gases by identifying them with large configurations of molecules that interact according to the laws of Newtonian mechanics. Because of the properties and relations provided by the theoretical structure, we can derive laws that govern the behaviour of observable objects. By contrast, if we remained strictly on the phenomenological level we would not be able to accurately formulate a law like the van der Waals gas law because we would be unable to appeal to the account of intermolecular forces provided by the higher-level theoretical structure.

Friedman sees the correct relationship between $\mathcal{A}$ and $\mathcal{B}$ as that of model to sub-model, where $B \subseteq A$ and $R'_i = R_i/B$ $(i \leq m)$. This characterization affords us a literal identification of the elements in $\mathcal{B}$ and $\mathcal{A}$, which in turn results in the larger structure $\mathcal{A}$ "inducing" theoretical properties and relations on objects in $\mathcal{B}$, properties necessary for stating accurate laws about observable objects (Friedman 1983, p. 240). Contrast this with what Friedman terms the representational account. On that view we do not interpret $\mathcal{A}$ literally (as the molecular world); rather, it is construed as a mathematical representation. Instead of asserting that $\mathcal{B}$ is a sub-model of $\mathcal{A}$, we claim only that $\mathcal{A}$ is embeddable into $\mathcal{B}$; there exists a one-one map $\phi : B \to A$ such that $\phi(R'_i) = R_{i/\phi(B)}$ $(i \leq m)$. $\mathcal{A}$ does not "induce" the necessary theoretical properties on objects in $\mathcal{B}$ unless, of course, those properties are definable from the observational properties $R_1, \ldots, R_m$. Consequently, we could have two different embeddings ϕ and ψ of $\mathcal{B}$ into $\mathcal{A}$ such that for some property R_j $(j > m)$ and some $b \in B$, $R_j[\phi(b)]$ & $-R_j[\psi(b)]$. This difficulty is avoided on the sub-model interpretation because of the uniqueness of the mapping (the identity map).

As Friedman points out, the representationalist account does not prevent us from generating accurate laws; we simply do so by adding new primitive properties and relations to $\mathcal{B}$ instead of deriving them directly from higher-level structure. However, on this account we provide explanations only in response to particular observable events. There exists no background structure that can be appealed to in attempting to furnish a unified account of various observable phenomena. As a result, the representationalist account provides explanations that are less powerful, and hence it proves unhelpful when confirmation of laws is at issue. The literal construal is preferred because it yields greater unifying power and increased confirmation; for example, we can *conjoin* molecular theory with atomic theory to explain chemical bonding, atomic energy and many other phenomena. Consequently, the molecular hypothesis will pick up confirmation in all the areas in which it is applied. The theoretical description then receives confirmation from indirect evidence (chemical, thermal and electrical phenomena) that it "transfers" to the phenomenological description. Without this transfer of confirmation the phenomenological description receives confirmation only from the behaviour of gases. So in cases where the confirmation of the theoretical description exceeds the prior probability of the phenomenological description, the latter receives the appropriate boost in confirmation as well. Hence the phenomenological description is better confirmed in the context of a total theory that includes theoretical description than in the context of a theory that excludes such description. The literal interpretation can thereby be seen as better confirmed, more plausible and less *ad hoc* (Friedman 1983, p. 241).

? ? only if you flip flop? increase or level? ?

2.2. The Importance of Conjunction

Friedman claims two virtues for his reductivist programme. First, there is a type of theoretical inference (specifically, conjunction) that is valid on the hypothesis of a genuine reduction, but not in the case of a representation. This is because the structural relations that hold between the model and the sub-model guarantee the identity of the observational and theoretical structures. Because Friedman sees $\mathcal{A}$ as a "real particular world in its own right" (1983, p. 246), it follows that the identity between $\mathcal{B}$ and $\mathcal{A}$ is also one that should be understood as literally true. For example, where Δ_1 and Δ_2 are classes of models, a reduction facilitates the inference

$$\frac{\begin{array}{lcl}\langle B, R_1\rangle \subseteq \mathcal{A} & \text{and} & \mathcal{A} \in \Delta_1 \\ \langle B, R_2\rangle \subseteq \mathcal{A} & \text{and} & \mathcal{A} \in \Delta_2\end{array}}{\langle B, R_1, R_2\rangle \subseteq \mathcal{A} \quad \text{and} \quad \mathcal{A} \in \Delta_1 \cap \Delta_2}$$

whereas the non-literal interpretation restricts the inference to the following form:

$$\frac{\begin{array}{lcl}\exists \mathcal{A} \exists \phi\colon \langle B, R_1\rangle \to \mathcal{A} & \text{and} & \mathcal{A} \in \Delta_1 \\ \exists \mathcal{A}' \exists \psi\colon \langle B, R_2\rangle \to \mathcal{A}' & \text{and} & \mathcal{A}' \in \Delta_2\end{array}}{\exists \mathcal{A}'' \exists \chi\colon \langle B, R_1, R_2\rangle \to \mathcal{A}'' \quad \text{and} \quad \mathcal{A}'' \in \Delta_1 \cap \Delta_2}$$

This latter inference is invalid, because $\mathcal{A}$ and $\mathcal{A}'$ are different models, but even if they were the same, we would require some guarantee that the mappings ϕ and ψ had the mapping χ in common. This, of course, is not needed on the reductionist account, because in all cases the mapping (the identity map) is the same. Consequently, we are able to obtain a single joint reduction that is already entailed by our original hypothesis, while on the representationalist account we require the addition of some new piece of theoretical structure.

The second virtue also concerns the utility of the conjunctive inference rule. According to Friedman, our theories evolve by conjunction. Certain assumptions about molecular structure play a role in the explanation of the gas laws, and these, together with further assumptions, figure in the explanation of chemical combination. As a result, the theoretical assumptions receive confirmation at two different times. These advantages also extend to the case of observational predictions. Suppose we have two reductions $\mathcal{A}$ and $\mathcal{B}$, each of which receives an individual boost in confirmation at time t_1 and t_2, respectively. If their conjunction implies a prediction P at t_3 that does not follow from either conjunct individually, then both conjuncts receive repeated boosts in confirmation at t_3 if the prediction is borne out. On the representational schema $[\exists\, \phi(A), \exists\, \psi(B)]$ we cannot derive the same observational prediction at t_3 by a simple conjunction; we need a new joint representation $\exists \chi(A \,\&\, B)$. The disadvantage of this approach is that the joint representation is formulated as a response to a new observational situation, rather than as a result of the theory's evolution over time by conjunction of hypotheses. Consequently, there is no common unified structure whose parts could participate in the increased confirmation.

Friedman is certainly right about the shortcomings of the representationalist/instrumentalist programmes in accounting for the role of conjunctive inference. The original version of the conjunction objection advanced by Putnam (1975) was

roughly the following: When a scientist accepts a theory, he believes it to be true; it is only by having such a belief that he is able to perform the appropriate conjunctions. Thus, because theory conjunction is a desirable virtue, our epistemic attitudes must be such as to allow for this practice. Friedman changes the tone somewhat by taking the argument one step further, claiming that the product of these conjunctions, the part of the theoretical structure that unifies the other parts, is what is to be believed or interpreted literally. However, because both he and Putnam claim that our theories evolve by conjunction, it appears that we must have some *prior* belief in the truth of our hypotheses in order to achieve the desired outcome. Although Friedman cites unification as a justification for the literal interpretation of theoretical structure, it is interesting to note that on his account we cannot achieve a unification unless we *first* adopt a reductivist approach that construes the theoretical structure as literally true. In other words, in order to have a unified theoretical structure, we must be able to conjoin our theories, which in turn requires the belief that they are true. But this was the *same* belief for which unification was thought to provide a justifying condition. Hence it appears as though we cannot simply limit belief to the unifying part of the theory; we need a stronger form of realism to motivate this model of unification.

Circle

One of the difficulties in Friedman's discussion of conjunction is his failure to distinguish between semantic realism and epistemological realism. This is especially important, because the semantic realist can quite consistently provide a literal interpretation of theoretical structure while withholding belief that the interpretation is in fact *true* (e.g., that what the kinetic theory says about molecules is a true description of their constitution).[3] This version of realism denies the tenets of classical instrumentalism by interpreting theoretical structure literally, thereby enabling us to appeal to theoretical structure for derivations and entailments of phenomenological laws without relying solely on phenomenological properties as the basis for theoretical explanation. But this need not entail belief that the structure is real. It requires only that it be interpreted literally – without the "as if" clause. It may be real, but there is no commitment either way. But this latter approach does not conform to the demands of the original conjunction objection, namely, that belief in the truth of our theories is the only way to make sense of conjunction. Because we do not claim truth for our hypotheses, we cannot motivate conjunction. In other words, conjunction seems to require epistemological realism as well. But in light of some recent literature is seems fairly obvious that this strong form of realism is no longer tenable; it is simply too strong a requirement to claim truth for theoretical hypotheses. Hence, we should perhaps interpret Friedman's conjunction model as distinct from Putnam's model, advocating the semantic version rather than the epistemological version of realism. But what are the consequences of this interpretation? Perhaps most important, it would render Friedman's position no different from that of van Fraassen (1980), one that he rejects (Friedman 1983, p. 220) because it fails to sanction the kind of realism about theoretical structure that Friedman needs for conjunction and unification. As he

yes

Van F

points out (1983, p. 246), it is this kind of epistemological anti-realism that has the undesirable consequence of allowing for different embeddings of observational structure into theoretical structure.

Independent of the logical issues about truth and inference is the question whether or not the actual practice of science and the evolution of theories can be modelled on the approach Friedman describes. It is important here to distinguish between the "in principle" use of conjunctive inference (the logical issue) and its legitimation in specific instances (whether or not science actually proceeds in this way). No one objects to the use of conjunction as a logical rule that guarantees truth, *if* we begin with true conjuncts. But making this assumption in the case of scientific theories is simply to ignore the very problem of realism. Moreover, when we conjoin theories, we rarely, if ever, do so strictly on the basis of logical principles. A complicated process of testing and manipulation is involved in order to ensure a relatively successful outcome. The criticism levied against the anti-realist by Putnam and Boyd emphasized that by denying knowledge of the truth of our theories, we have no guarantee that the conjunction will be successful and no way to make sense of the practice of theory conjunction. But isn't that exactly the point? We do not and should not have any guarantee that the practice will be successful. At the level of theory construction, unification is nothing like the kind of straightforward process that this account suggests. Moreover, if we consider what is involved in theory conjunction, realism presents no methodological advantage. Scientists initially bring theories together with an eye to further testing and successful prediction; only if the conjunction survives empirical tests and yields accurate predictions will the theory be believed or accepted. Hence, our expectations for theory conjunction are validated on the basis of predictive success; but to equate that predictive success with truth would yield the kind of instrumentalism that realists typically object to because it cannot account for the theory conjunction.

A realist might want to maintain that successful conjunction provides evidence for or is an indicator of truth; and indeed Friedman himself seems to suggest that in claiming that through conjunction, theories and pieces of theoretical structure pick up boosts in confirmation (1983, p. 254). Although no "confirmation theory" is provided by Friedman, it is reasonable to assume that, like most realists, he equates increased confirmation with increased likelihood that the theory is true. If one interprets this structure as a literally true description of reality, then difficulties arise in cases where there is more than one way of identifying observational structure with theoretical structure. Approximate truth fares no better, because logical laws like conjunction and transitivity cannot always be successfully applied to terms and hypotheses that are true only in a limited domain. Consequently there is no reason to assume that theoretical conjunctions involving these terms will be truth-preserving. Similarly, any theory that defines truth epistemically or instrumentally will also prevent application of the conjunction rule.

or Confirmation itself?

But is this all that can be said about the issue of conjunction? Demopoulos (1982) claims that no realist would deny that correction often occurs prior to conjunction – a point that leaves the anti-realist with the problem of accounting for the conjunction of corrected theories.[4] It seems, however, that because the issue of correction has been introduced, the logical and methodological aspects of Putnam's original conjunction objection have changed. The point of the conjunction objection, in its early formulation at least, was that the truth predicate gave realist epistemology a distinct advantage over its rivals who denied the truth of scientific theories. If one could not expect true predictions, then what could possibly be the purpose of conjoining theories?[5] But if the motive for correction is to facilitate theory conjunction, which presumably it is, then the truth predicate that was initially applied to our theories has little, if anything, to do with the methodological process. The conjunction of corrected theories then becomes an empirical process of bringing together two theories that have been *designed specifically* for that purpose. If they have previously been corrected to ensure, as it were, successful prediction, then there is no reason for the realist to claim any kind of epistemic or methodological superiority. The anti-realist simply explains the practice of conjunction as one that is crucial in the search for theories that are equipped to explain and predict a variety of phenomena. The issue (and practice) becomes a methodological one that involves trial and error, rather than simply a logical operation encompassing semantical and epistemological considerations.

good

It seems, then, that if Friedman's account of theory conjunction shares the presuppositions of the original Putnam-Boyd formulation, it is fraught with many of the difficulties that led many realists to soften their line on whether or not certain theories could or should be believed to be true in the sense required by conjunction. If, on the other hand, he intends his account to involve no more than semantic realism, an argument for a literal interpretation of theoretical structure (in the manner of van Fraassen), then he cannot motivate the kind of uniqueness that the model/sub-model approach and the identity map guarantee.

Much of Friedman's discussion of conjunction depends on the viability of his model/sub-model account of the relationship between observations and theoretical structure. Ironically, it is possible to show that the kinds of reductions achieved by means of the identity map can actually *prevent* a literal interpretation of theoretical structure – a result that seems to leave the embedding approach as the more accurate representation of scientific practice.

2.3. Reduction versus Representation

2.3.1. *Is Reduction a Viable Approach?*

The traditional philosophical problems associated with reduction focused on the relationship between thermodynamics and statistical mechanics and dealt with the identification of concepts such as temperature, mean kinetic energy and entropy. However, the idealized nature of the assumptions of statistical mechanics exposes a more serious problem than those linguistic debates would indicate. Consider the following case, in which we have a geometric representation of a mechanical system. From a physical point of view, we describe the mechanical system G with s degrees of freedom by values of the Hamiltonian variables: $q_1, q_2, \ldots, q_s; p_1, p_2, \ldots, p_s$. The equations of motion assume the following form:

$$dq_i/dt = \partial H/\partial p_i, \qquad dp_i/dt = -\partial H/\partial q_i \qquad (1 \leq i \leq s)$$

where H is the so-called Hamiltonian function of the $2s$ variables $q_1, \ldots, p_s$. The Hamiltonian function expresses the energy of a system in terms of momenta p and positional coordinates q. Now consider a Euclidean space Γ of $2s$ dimensions whose points are determined by the Cartesian coordinates $q_1, \ldots, p_s$. To each possible state of the mechanical system G there corresponds a uniquely determined point of the space Γ that can be called the image point of the given system. The whole space Γ is the phase space of the system. The dynamic coordinates of a point in Γ are simply the Hamiltonian variables of the given system G. Any function of these variables is called ~~the~~ phase function of the system. The most important phase function is the Hamiltonian function $H(q_1, \ldots, p_s)$, which determines the mechanical nature of the system (in virtue of the fact that it determines the equations of motion). The total energy E of the system can be represented as $E(q_1, \ldots, q_s; p_1, \ldots, p_s)$. For an isolated part of the system, this function has a constant value; hence for any constant a, the region of the phase space for the point where $E = a$ is an invariant part of phase space. Such regions can be referred to as surfaces of constant energy. Hence, $\sum_{(x)}$ is the surface of constant energy where $E = x$; for $x_1 < x_2$, the surface $\sum_{(x_1)}$ is situated entirely inside $\sum_{(x_2)}$. The family of surfaces of constant energy can be represented as a family of concentric hyperspheres. The structure function of the given system can be defined as the measure (volume) $\Omega(x)$ of the surface of constant energy $\sum_{(x)}$. This structure function determines certain features of the mechanical structure of the corresponding physical system, as well as geometrical aspects of phase space.

Assume that the total energy $E(x_1, \ldots, x_n)$ of a system can be represented as the sum of two terms E_1 and E_2, and $(x_1, \ldots, x_n)$ denotes the dynamical coordinates of a point of the space Γ (the product of the phase spaces of all the components). Each phase function and the total energy E of the given system are functions of these n variables. $E_1 = E_1(x_1, \ldots, x_k)$ depends on *some* of the dynamical coordinates, whereas $E_2 = E_2(x_{k+1}, \ldots, x_n)$ depends on the remaining coordinates. Given this characterization, we say that the set of dynamical coordinates $(x_1, \ldots, x_n)$ of a particular system is decomposed into the components $(x_1, \ldots, x_k)$ and $(x_{k+1}, \ldots, x_n)$. However, a peculiarity results when we try to interpret each component as a separate physical system contained in the given system.[6] Although each materially isolated part of the system usually determines a certain component of the system, some components or sets of coordinates do *not* correspond to any *materially isolated* part of the system. The isolated character of these components defines (in the sense given earlier by the definition of a component) pure

Set of coords

Components are not ...

energetical aspects of the system. For example, consider a system of one material particle, with the components of velocity and mass being u, v, w, m; if its energy E reduces to kinetic energy (because there are no intermolecular forces), we have $E = m/2(u^2 + v^2 + w^2)$. Although u is a component of the system whose energy is $mu^2/2$, it doesn't correspond to any material aspect of the system. Although it *mathematically* represents the u component of velocity, its relation to an isolated material aspect of the mechanical system is more problematic. First of all, it is not even possible to say that any particular molecule has exactly some stated velocity; instead, because of the way in which probabilities enter in the calculation of the velocity distribution function, it cannot be applied to a singular situation. In order to make a substantial claim about velocities, we must consider a number of molecules in some range of velocity (Feynman 1963, vol. 1, pp. 40–5).

Even if we consider a gas with a wide distribution of molecular velocities, it isn't clear that separation of velocity into its three components would result in anything that could be isolated in any *physical* way within the mechanical system. Because velocity is a vector quantity, we assume that it is possible to *physically* isolate each component because we can do so *mathematically*.[7]

What are the consequences of an example like this for Friedman's view? That is, if we subscribe to reduction as a theoretical goal, how do we interpret a theoretical situation of this sort? If we literally identify the energy of a particular molecule with the mathematical representation $E = m/2(u^2 + v^2 + w^2)$, then there are aspects of the latter that seem to lack any materially isolatable counterparts having physical significance. In other words, the notion of a literally true identification in this case seems too strong.[8] Nevertheless, as a mathematical representation it contains certain parameters that are crucial for modelling a statistical system. Each component is a group of dynamic coordinates, and it has a definite energy, with its own phase space (the phase space of the system Γ being the product of the phase spaces Γ_1 and Γ_2 of its two components). Moreover, each component also has its own structure function, and taken together, they determine the structure function of the given system. Indeed, the law governing the composition of the structure function is one of the most important formulas in statistical mechanics.

An additional problem is the methodological paradox that arises from decomposition of the system into components, something that results in exclusion of the possibility of any energetical interaction between particles defined as components. The irony is that statistical mechanics invariably assumes that particles of matter are in a state of intensive energy interaction, where the energy of one particle is transferred to another through the process of collisions. In fact, its methods are based precisely on the possibility of such an energy exchange. Quite simply, if the total energy of a gas is expressed as the *independent* energies of the two components (the energies of the molecules), then the assumptions of conservation and velocity distribution are violated, because each assumption requires that the particles interact. If the Hamiltonian expressing the energy of the system is a sum of functions, each of which depends only on the dynamic coordinates of a single particle (and

Tedious!

Cross terms from ...

Point – Friedman conjunction lacks cross-terms.

represents the Hamiltonian of this particle), then the entire system of equations governing the motion of the system splits into component systems. Each component system describes the motion of some separate particle and is not connected to any other particle (Khinchin 1949). As a result, the energy of each particle expressed by its Hamiltonian function appears as an integral of equations of motion and remains constant.[9] From the fact that the particles are independent and the fact that the sum of the energies is constant, it follows that the individual energies must be constant as well. But because this conclusion violates conservation of energy, we must deny the claim that the total energy is the sum of n independent individual energies. In other words, the way the mathematical model describes the system violates some of the structural constraints of statistical mechanics; hence there is good reason not to interpret the mathematical representation as a literally true account of the mechanical system.

For practical purposes, the difficulty is resolved by idealizing assumptions that consider particles of matter as *approximately* isolated energetical components. Although the precise characterization of energy contains terms that depend simultaneously on the energy of several particles, and also allows for energy interaction between them, these forces of interaction manifest themselves only at very small distances. Consequently, the "mixed terms" in the energy equation (those that represent mutual potential energy of particles) will be negligible compared with the kinetic energy of the particles and therefore will be of little importance in the evaluation of averages. In a majority of cases, such as calculation of the Boyle-Charles law, we can neglect these terms and still arrive at a good *quantitative* approximation; we simply assume that the energy of the system equals the sum of component energies. However, on a *qualitative* analysis the mixed terms are extremely important, because they provide the basis for an understanding of energy exchange between particles, the very core of statistical mechanics. Hence, we sacrifice explanatory power for predictive success.

These examples raise some fairly serious difficulties for the kind of literal reductivist approach outlined by Friedman. Even if we disregard the problem of identifying temperature and mean kinetic energy across theoretical boundaries, a more significant difficulty arises in the case of *identifying*, in the way suggested by the model/sub-model approach, the constituents of the system postulated by classical statistical mechanics with its individual particles. The structural presuppositions involved are radically different in each case. Although we can ignore these assumptions in some cases of quantitative prediction, that is not the important issue. As Friedman himself suggests, if we are interested in purely phenomenological laws, then there is no reason to prefer a reduction to a representation (1983, p. 241). But the motivation for Friedman's account is to achieve a literal interpretation of theoretical structure, which in turn will yield greater confirmation of hypotheses, something he sees as guaranteed by the model/sub-model approach.

yes! it conflicts

If we recall the constraints involved in the relationship between a model and its sub-model, we see that they are structurally similar insofar as the interpretation of

each relation, function and constant symbol in the sub-model $\mathcal{B}$ is the restriction of the corresponding interpretation in the model $\mathcal{A}$. Equivalently, for every atomic formula ϕ and assignment (s) in $\mathcal{B}$, $\mathcal{B} \models \phi(s)$ iff $\mathcal{A} \models \phi(s)$. Applied to our physical example, we see that a *literal* identification of the properties of individuals of the mechanical system $\mathcal{B}$ cannot be accomplished given the structural constraints on $\mathcal{A}$, the mathematical representation of the statistical system. A literal identification of $\mathcal{B}$ with $\mathcal{A}$ would preclude the formal mathematical model of the statistical theory from accounting for specific parameters (the possibility of energy exchange between particles) that must be interpreted literally if we are to have a proper understanding of the theory's physical foundations. This difficulty can be countered on a representationalist account, where we have an embedding of the properties in $\mathcal{B}$ into $\mathcal{A}$. We do not claim a literal *identification* of one with the other, but instead correlate, by way of an embedding map, certain features of $\mathcal{B}$ with features of $\mathcal{A}$. Every aspect of $\mathcal{B}$ need not have a counterpart in any one model of the statistical theory. Instead, the theory may have several models, each suited to a particular application. In this case the relationship between corresponding elements of $\mathcal{A}$ and $\mathcal{B}$ is not uniquely specified by the identity map, and hence there can be a variety of ways that the "reduced" entities/theory can be correlated with the reducing theory or model. Given the logical properties of the model/sub-model relationship, we demand that the relations and functions specified by the identity map be preserved over time, in the way we think of inference rules as truth-preserving. But the representational approach allows the relationship between $\mathcal{A}$ and $\mathcal{B}$ to change over time, something that is *prima facie* ruled out by a literal identification of their corresponding elements. Although these problems do not necessarily deal directly with the straightforward reduction of observational structure to theoretical structure, they do deal with the idea of reducing and identifying physical concepts/properties with theoretical or mathematical representations. What the examples expose are the difficulties associated with reduction as a methodological strategy. But as we shall see later, the demand for different models to account for the same phenomena also arises in the more narrowly defined context in which we have a simple reduction of observable entities to their theoretical counterparts. This situation poses obvious problems for the model/sub-model approach and the accompanying idea that we can correlate the elements in each model by means of an identity map.

2.3.2. *The Problem of Many Models*

In Friedman's kinetic-theory example he claims that given an appropriate theory of molecular structure and intermolecular forces, we can explicitly define the a and b terms (those representing molecular size and intermolecular forces) and go on to derive the van der Waals law from the kinetic theory – something we cannot do if we remain at the phenomenological level. Although I am in agreement with his claims about the disadvantages of a purely phenomenological approach, even if we acknowledge the need for a literal interpretation of theoretical structure it does not follow that the model/sub-model approach can be vindicated. We shall see that in different contexts the solutions to the problems addressed by this law indicate the need for more than one molecular model. Hence there exists a "looseness of fit" between the phenomenological and theoretical structures that cannot be accommodated on the model/sub-model account.

Friedman contrasts the van der Waals law, $(p + a/V^2)(V - b) = RT$, with the Boyle-Charles law, $pV = RT$, claiming that the latter is false, whereas the former presents a more accurate account of real gases (1983, p. 239). He goes on to claim that it is the structure of the kinetic theory that supplies us the properties and relations that enable us to formulate the more accurate law. As a point of historical interest, if we look at the details of this "deduction" we can see that the van der Waals method for deriving his equation of state departed from the kinetic principles illustrated by the virial theorem, and as such his equation was unsatisfactory as a deduction from the kinetic theory.[10] It is also interesting to note that although the van der Waals theory suggested the possibility of explaining the gas–liquid transition in terms of intermolecular forces, it was not really an application of statistical mechanics (see the Appendix to this chapter). The first and simplest example of a phase transition derivable from statistical mechanics was the famous condensation of an Einstein-Bose gas at very low temperatures. And although that discovery was made in 1924, its physical significance was not appreciated until almost 10 years later (Brush 1983).

Historical points aside, the important issue here is the possibility of reconstructing this example according to the model/sub-model strategy. In order to do this, we require, minimally, that the molecular assumptions required for the corrected van der Waals law to hold not contradict those required of the Boyle-Charles law. Not only is this condition not met, but also when we move to further refinement of the gas laws we need additional assumptions that conflict with those initially postulated. Despite its experimental corroboration, when applied to cases of greater than first-order deviation from Boyle's law, the molecular model suggested by the van der Waals approach was seen to be insufficient (Tabor 1979; Jeans 1917, esp. ch. 3). Basically, the model overlooked the fact that when cohesive forces exist between the molecules, some molecules never reach the boundary (the wall of the container). As a result, van der Waals assumed that those molecules exerted a negative pressure, an assumption that implied negative values for p. Because an examination of physical conditions showed that the true value for p had to be positive, an alternative formulation and molecular model were proposed by Dieterici: $p(V - b) = RT \exp(-a/RTV)$ (Tabor 1979; Jeans 1917). That model assumed a constant temperature of the gas molecules, so that the total energy distribution applied to molecules striking the wall as well as those that did not. Although both equations imply the existence of what is termed a critical point, a point where the liquid, gaseous and vapour states meet,[11] they make different predictions as to the existence of this point; with the Dieterici equation appearing to be more accurate for heavier and more complex gases. Generally speaking, however, neither one comes particularly close to actual observations of critical data.[12] The reason for the discrepancy is that both equations are true only when deviations from Boyle's law are small, with the critical point representing a rather large deviation.[13]

Various attempts have been made to improve the van der Waals equation by the introduction of more readily adjustable constants to supplement a and b, constants referring to molecular size and forces that can be chosen to make the equation's results agree more closely with experiment. One approach introduced a term a' to replace a. Because a' specified that a vary inversely as the temperature for some gases, it provided a better fit with the observations than did the original van der Waals equation.

The overall difficulty seems to be one of specifying a molecular model and an equation of state that can accurately and literally describe the behaviour of gases. We use different representations for different purposes: the billiard-ball model is used for deriving the perfect-gas law, the weakly attracting rigid-sphere model is used for the van der Waals equation and a model representing molecules as point centres of inverse-power repulsion is used for facilitating transport equations. What the examples illustrate is that in order to achieve reasonably successful results we must vary the properties of the model in a way that precludes the kind of literal account that Friedman prescribes (an account that assumes that our model is a literally true description of reality). Instead, an explanation of the behaviour of real gases (something the van der Waals law is designed to explain) requires many different laws and incompatible models. And there are other difficulties with the accuracy or so-called truth of the van der Waals law. In general, the equation tends to smooth out the differences between individual substances and predicts that they will behave more uniformly than they do.[14] In fact, very accurate experiments have shown quantitative discrepancies from the results predicted by the van der Waals law. In calculations of the difference in density between the liquid and gaseous phases, the equation predicts that the difference should go to zero as the square root of the difference between the temperature and T_c. In reality, that difference varies nearly as the cube root, a result that suggests differences in the microstructure of fluids. In what sense, then, can we link the van der Waals gas law with a molecular model that truly describes or can be identified with the behaviour of gases at the phenomenological level?

If the relationship between the behaviour of gases and their molecular model is one of sub-model to model, then the same relations and properties that hold in the latter must hold in the former (with the sub-model being a restriction of the relations in the model). So if the van der Waals equation requires a specific molecular model to establish its results, and the Dieterici equation requires a different model, it seems that we are unable to claim that either provides a literally true account of molecular structure. The so-called derivation of the van der Waals law can be achieved using a particular model that we know to be inapplicable in other contexts. Hence it appears that the uniqueness of the mapping in the model/sub-model account is actually a drawback rather than an advantage.

The fact that the observable behaviour of gases requires more than one model for its explanation again seems to favour the embedding approach over the sub-model account as a way of understanding physical theory.[15] The advantage here is that one is not committed to the literal truth of the theoretical claims, something that allows for an evolution in our views about the nature of physical systems without falling into logical and semantic difficulties over interpretations of truth. Because the embedding account does not require commitment to a strictly defined and unique identity relation holding between observable properties of physical systems and their molecular configurations, we can accept many different possible interpretations of theoretical structure at the same time. As pointed out earlier, this does not rule out a literal interpretation of the molecular structure postulated by the kinetic theory; we need not become instrumentalists in the way Friedman suggests. The sceptical problem regarding our ability to assign a particular truth value to the theory or model in question is an epistemological issue, not a semantical issue. Whether or not we have sufficient evidence or justification to claim that our theory is true is a separate concern, and an answer is not required for a literal interpretation of the theory's assumptions.[16]

How about the right model

There is, however, another line of defence that the realist could use to vindicate the position. One could claim in this case that the van der Waals example shows that a single *general* molecular theory that does *not* incorporate any *specific* assumptions about when a fluid is gaseous and when it is liquid could be used to explain the transition from one state to the other. The idea of continuity between the gaseous and liquid states of matter supplies a kind of ontological unity that forms the theoretical basis for the van der Waals equation. One could then interpret this as suggesting that there is one overarching model of molecular theory, the details of which change over time, allowing for corrections and revisions to the same basic structure. If an account like that could be motivated, then one could claim that despite the changes in some properties, the core of the model would remain unchanged. Such a view clearly would go some way toward vindicating Friedman's model/sub-model interpretation of theories, and it would solve the problem of incorporating the substantial changes that took place within the framework of the kinetic theory over a period of time. For instance, in addition to the fact that the van der Waals law and the properties of real gases require a number of different models for their explanation, several changes in the structural presuppositions of the theory were necessary to account for the problem of specific heats at low temperatures. Because there occurred a "falling off" of specific heat for diatomic gases such as hydrogen, a phenomenon that could not be attributed to the fact that hydrogen ceases to be a perfect gas at low temperature, the kinetic theory needed to be modified in the domain of idealized perfect gases. The theory was subsequently improved by reinterpreting it in terms of the mechanics of relativity. The foundations were left unchanged (i.e., the Gibbs theorems on the conservation of extension and density in phase, and the ergodic hypothesis), but the law of equipartition was rejected and replaced by a different law of partition. It was soon noticed that those refinements were not applicable when the motions of

molecules were slow, a situation that arises at low temperatures when relativistic theory merges into the classic theory. The required modification was furnished by the quantum theory and the statistics of Fermi and Bose.

The difficulty with adopting the kind of restrictive yet generalized realism described above is that it requires a separation of the entities like molecular structure and the properties that define that structure, properties that supposedly give rise to many of the empirical phenomena we are concerned to explain. The idea is that there exists a molecular structure that forms the basic core of the model, but we have no stable, realistic account of the details of the structural properties, because our account of them changes with shifts in theoretical knowledge. However, it is exactly these properties that figure importantly in the derivations of phenomenological laws from the higher-level theoretical structure. Postulating the existence of a molecular structure defined by a model that is devoid of specific properties and relations allows us to maintain our structure over time, but in return provides none of the advantages the model was designed to create. On the other hand, a molecular model endowed with specific features cannot be interpreted as a literally *true* description of theoretical structure, because such an interpretation would provide no mechanism for changes in the model over time and no account of the nature of incompatible models – situations that are necessary for a realization of the various contexts and possibilities envisioned by the theory. Hence, neither option is a possibility. Because the embedding approach allows for a variety of models of the phenomena, it seems closer to actual scientific practice. Not only is Friedman's model/sub-model account too restrictive in its requirement that the model be literally true, but also the relationship between the phenomenological structure and theoretical structure cannot be accurately depicted by the kind of stringent logical requirements (furnished by the identity map) that hold between models and their sub-models.[17]

The embedding approach is a significant feature of van Fraassen's semantic view of theories. The importance of the semantic view for our discussion rests not only with the details discussed earlier but also with the way models function as part of scientific practice. According to the semantic view, a theory is simply a family of models that represent actual phenomena in a more or less idealized way; so when we talk of embedding we refer to the process of mapping the phenomena into one of the many models that might be used to account for them. For example, there are many models that can describe a quantum-mechanical particle in a potential well, and we can model the pendulum as an undamped or a damped harmonic oscillator, depending on how "realistic" we need our model to be. The laws of Newtonian mechanics tell us how to add corrections to the model of the undamped pendulum in order to make it more like the physical apparatus, and in that sense the resulting model can be seen as a model of the theory, an application of the theory's laws in a specific context as well as a model of the physical pendulum. We don't simply apply the theory directly to nature; instead, we construct a model that resembles some aspects of a natural phenomenon to see how the laws apply to an object so defined. The damped oscillator gives predictions that more accurately approximate the phenomena, but sometimes that level of accuracy is not needed to solve the problem at hand. The important point here is that the semantic view seems better equipped to handle the diversity of models in scientific practice and their role of providing approximations to physical systems. In other words, scientific models are never exact replicas of either the theories or the phenomena. They can approximate either the theories or the phenomena or both. In that sense it would be misleading to claim, as the model/sub-model account does, that there is a literal identification between the elements of a model (in that case understood as the observable phenomena) and the theoretical structure. Neither the natural sciences nor the social sciences view models in that way. Not only are there many ways the phenomena can be modelled, but also theoretical structure, by its very nature, will always present a more abstract picture than the phenomena themselves.[18]

Where does that leave us with respect to achieving or defining unity? I shall have more to say about this in later chapters, particularly on the connection between explanation and unification, but for now let me briefly recap the basic strategy of Friedman's argument. On his view, we achieve theoretical unity through a process of reduction and conjunction. The observational structure of a theory is reduced to its theoretical counterpart by identifying the two in the appropriate sort of way. This process can be represented in logical terms by means of the identity map that relates a sub-model to its model. Once the reduction and identification are complete, the theoretical structure can then be conjoined to other structures to achieve a unified theoretical account of diverse phenomena. What I have tried to show is that neither the theories that incorporate observational and theoretical structure nor the phenomena themselves are literally reducible in the way the model/sub-model approach suggests. The use of the reductivist strategy as well as the use of conjunction as important features in unification become either simply inapplicable or at best questionable. The apparatus of formal model theory is simply too rigid to capture the rather messy relations that are part of the modelling of scientific phenomena.

If we look at the pattern of unification suggested by William Whewell under the title "consilience of inductions", an approach that Friedman likens to his own account (1983, p. 242, n. 14), we quickly see that neither conjunction nor reduction plays a role in the unifying process. As we saw in Chapter 1, we find instead a rather complex process of reinterpretation of basic aspects of theoretical laws and structures, a reinterpretation that extends far beyond the product of the simple conjunction of existing theories. Although Whewell's account of the process squares better with scientific practice, he also sees unity as a "stamp of truth", something that, I want to claim, need not be the case.

2.4. Consilience and Unification

According to Whewell, a consilience of inductions is said to occur when a hypothesis or theory is capable of explaining two or more classes of known facts, when it can predict cases that are different from those the hypothesis was designed to explain/predict, or when it can predict/explain unexpected phenomena.[19] In each case a consilience of inductions results in the unification or simplification of our theories or hypotheses by reducing two or more classes of phenomena that were thought to be distinct to one general kind, or by showing their behaviours to be describable by one theory. In addition, this unification results in a reduction in the amount of theoretical structure required to account for the phenomena. Reduction is an important element, but not the kind of reduction characteristic of the model/sub-model approach. Although the deductive-entailment content of consilient theories is very high, often that virtue is achieved only through a process of *reinterpretation* of the laws and key terms:

> When we say that the more general proposition includes the several more particular ones . . . these particulars form the general truth not by being merely enumerated and added together but by being seen in a new light. (Butts 1968, pp. 169–70)

Whewell goes on to point out that in a consilience of inductions

> there is always a new conception, a principle of connexion and unity, supplied by the mind, and superinduced upon the particulars. There is not merely a juxta-position of materials, by which the new proposition contains all that its component parts contained; but also a formative act exerted by the understanding, so that these materials are contained in a new shape. (Butts 1968, p. 163)

Perhaps the most frequently cited example of a consilience (both by Whewell and by contemporary philosophers of science) is the unification of Kepler's and Galileo's laws under the inverse-square law. Newton's theory of universal gravitation could explain terrestrial phenomena like the motions of the tides and celestial phenomena like the precession of the equinoxes, classes of facts that were thought to be disjoint. The corrections applied to lower-level laws such as Galileo's laws of falling bodies and Kepler's third law of planetary orbits were motivated strictly on the basis of the overarching theory, rather than as generalizations from phenomena, as was the case in their initial formulations. Alternatively, one could say that Newton's theory showed how terrestrial and celestial phenomena were the same kind of entity (i.e., gravitating bodies).

Traditionally, many philosophers of science (including Duhem, Hesse, Laudan and Butts) have argued that the relationship between Newton's theory and Kepler's laws (as well as Galileo's laws) is not one of conjunction or even entailment, because universal gravitation *contradicts* the conclusions that those individual sets of laws provide. More recently, Malcolm Forster (1988, pp. 88–91) has pointed out that Newton himself was explicit in denying that view, maintaining that Kepler's third law held exactly, because all observed deviations from the law could be ascribed to other causes (which he afterward explained as the result of other gravitating masses). Forster claims that Newton proves, in the *Principia*, that Kepler's three laws for the earth's motion around the sun or the moon's motion around the earth can be expressed as instances of the inverse-square law, with the equation for each of the respective motions representing the Keplerian component of the motion as described by the three laws; that is, the solutions for the equations are elliptic orbits satisfying the area and harmonic laws. Conversely, Kepler's area law implies that the acceleration of each body is toward the other, and the ellipticity of the paths proves that the acceleration is inversely proportional to the square of the distance between the two bodies. In that sense, Kepler's laws can be seen to entail instances of the inverse-square law, suggesting the conclusion that Newton's theory of gravitation *entails* that Kepler's laws are true. Although they do not provide a complete description of the phenomena, they do give a "description in full agreement with Newton's theory" (Forster 1988, p. 89) and play an essential part in Newton's reasoning.

It is important, however, in this context to look beyond the mathematics to the physical interpretation that each theory furnishes. As we saw in the statistical-mechanics examples discussed earlier, we often need to supplement abstract mathematical laws and models with a qualitative understanding of the theory in order to appreciate the implications for concrete physical systems. In the case of Newtonian mechanics and Kepler's third law, the latter states that the cube of the mean distance of a planet from the sun divided by the square of the period of revolution is a constant for all planets ($a^3/T^2 =$ constant). [for all planets] The Newtonian version of the law states that $a^3/T^2 = m + m'$, where m is the mass of the sun, and m' is the mass of the planet in question. By ignoring m' on the ground that it is much smaller than m (at least for our solar system), we can assume that the two laws are roughly the same. In what sense is the Keplerian formulation true? It is true only if we ignore the fundamental qualitative aspects of Newton's theory that serve to differentiate it from Kepler's account of celestial mechanics. If we consistently ignore m', it becomes impossible to apply Newton's theory, because there is no gravitational force on a body with zero rest mass. On the other hand, if we make the simplifying assumption that m' is the same for all planets, we can then apply Newton's theory to get Kepler's laws within all observational accuracy. But here again the point is that the masses of the planets *are* different, and it is one of the benefits of Newtonian mechanics that we are able to calculate planetary orbits based on that information.

Nice example of full [illegible]

A similar situation holds for the case of Kepler's first and second laws. Newton himself remarks (*Principia*, bk. 1, prop. lxv, theorem xxv) that in cases where more than two bodies interact, Kepler's first two laws will be approximately valid at best, and even then only in very special cases. Although it is true that in some instances we can arrive at the same numerical values using both Kepler's law and Newton's law, the relationship between them, within the context of an overall mechanical theory, is not one of entailment. Nor is the latter the result of a conjunction of the former plus Galileo's laws.[20] From neither of these two groups of laws taken separately is there any indication of how they can be conjoined to produce a theory like Newtonian mechanics.

Cross terms again

What about other instances of unification, such as relativity theory and Maxwell's electromagnetism? Is it reasonable to assume that there has been significant

conjunction in each case? In the case of electromagnetic theory, there was a relatively straightforward identification of light waves and electromagnetic waves that resulted from Maxwell's introduction of the displacement current, a phenomenon for which there was no experimental/physical justification. That led to modification and reinterpretation of the physics behind Ampere's law, which described the relationship between an electric current and the corresponding magnetic field, a law that had been central to the earlier theories of electromagnetism. In addition to those changes, there were laws of physical optics that could not be accounted for by the new field theory, most notably reflection and refraction (dispersion). The sense in which optics and electromagnetism were simply conjoined and corrected is remote at best. The same is true of special relativity. Einstein left the mathematical form of Maxwell's equations virtually intact in showing them to be Lorentz-covariant. There were, however, several changes made to Newtonian kinematics and dynamics. And although for cases of low velocity Newtonian mechanics gives completely accurate quantitative results that are indistinguishable from those of relativity, the unification of Newtonian mechanics and electrodynamics involved substantial reinterpretation of classical physical magnitudes, including the nature of space and time.

Those instances, as well as the unification achieved in general relativity and our current attempts at bringing together quantum mechanics and relativity, exemplify a much more complicated process than simple conjunction and correction. In each case there is what Whewell calls the "introduction of a new mental element" (Butts 1968, p. 170). Whewell is careful to point out that the inductive truth is never merely the "sum of the facts"; instead, it depends on the "*suggestion* of a conception not before apparent" (Butts 1968, p. 170). The generality of the new law or theory is constituted by this new conception. This issue speaks not only to the unification of specific theories but also to the wider issue of convergence and unity in physics as a whole. The presence of limiting cases (e.g., Newtonian mechanics as a limiting case of relativity for low velocities) does not, in and of itself, suggest unity or convergence, any more than conceptual reorganization suggests incommensurability. It seems reasonable to expect that many low-level laws describing the behaviour of macroscopic objects will remain relatively unchanged in the development of scientific theories. Where significant change will occur is in our understanding of fundamental aspects of nature. This is the point at which our concepts and theories are becoming drastically altered, often with a complete reorientation as to what counts as an explanation of a natural process. The case for unity and convergence across the history of physics is a difficult one to make, most likely because it simply can't be made in anything but a trivial way.

Cross them

agree

Whewell's own work indicates a lack of unity across the sciences, but argues strongly for unity within each one. However, the *kinds of arguments* he offers for that unity have limited, if any, application to contemporary practice. Contemporary realist approaches take after Whewell in attributing a higher degree of confirmation to the unifying theory on the assumption that we can explain and possibly predict a variety of phenomena. The criterion of diversity emphasized by Whewell is the key to understanding his notion of consilience. When a theory was found to be applicable to a body of data other than that for which it was designed, the additional data were seen as providing independent evidence for the theory. So, for example, in the development of the theory of electromagnetism by Maxwell, many changes were made to existing notions of electromagnetic-wave propagation. But in addition, Maxwell found that the velocity of wave propagation coincided with the velocity of light and that the theory was able to account for various optical phenomena. According to Whewell (but not necessarily Maxwell himself, as we shall see in the next chapter), that would constitute independent evidence for the theory. A consilience is similar to the testimony of two witnesses on behalf of a hypothesis:

> . . . and in proportion as these two witnesses are separate and independent the conviction produced by their agreement is more and more complete. When the explanation of two kinds of phenomena, distinct and not apparently connected leads to the same cause such a coincidence does give a reality to the cause, which it has not while it merely accounts for those appearances which suggested the supposition. This coincidence of propositions is . . . one of the most decisive characteristics of a true theory . . . a consilience of inductions. (Whewell 1847, II:285)

When two different classes of facts lead to the same hypothesis, we can assume that we have discovered a *vera causa*.

We must, however, be cautious when characterizing the so-called independent evidence that is cited on behalf of a consilient theory. Although the evidence may be drawn from a variety of different and supposedly independent domains, there may nevertheless be no independent evidence for the theory itself or its particular unifying structure other than its ability to present a unified account of disparate phenomena. Again, Maxwell's electromagnetic theory was highly successful in unifying electromagnetic and optical phenomena. However, the mechanical model of the aether and the displacement current that initially facilitated that unification could not be justified on experimental or independent grounds. It was the failure of the aether model that led Maxwell to reformulate the theory using the abstract dynamics of Lagrange, with electrical and mechanical concepts occupying merely illustrative roles. As mentioned earlier, a crucial component in Maxwell's aether model and in the unification of electricity and magnetism was a phenomenon known as the displacement current. It was introduced to augment the usual conduction current and to create a field-theoretic explanation of propagation, which further enabled him to calculate the velocity of electromagnetic waves travelling through the hypothetical medium. When the early aether models were abandoned, electric displacement remained as a designated quantity, yet he remained agnostic about any qualitative account that might be given of either its nature or operation. Although the theory unified a great many phenomena, there was no evidence for the existence of electromagnetic waves themselves, nor any explanation of how they could be propagated.

Many of Maxwell's contemporaries, including Kelvin, thought of displacement as little more than an *ad hoc* postulation. And as the history of the period reveals, Maxwell's theory was not well received, despite the rather remarkable unification it achieved; nor was there increased support for the displacement current as a crucial theoretical structure that facilitated the unification. Maxwell's final account of electromagnetism, presented in his *Treatise on Electricity and Magnetism*, was a theory that achieved its unifying power from the Lagrangian formalism; a slightly revised version was fully embraced by the scientific community some 13 years later, after Hertz's famous experiments on electromagnetic waves. The vindication of the theory came not as a result of its ability to explain or unify a variety of independent phenomena through the postulation of theoretical structure, but from independent experimental tests showing the existence of electromagnetic waves and the use of the field equations to describe their propagation. So although these examples seem to follow the *process* of unification described by Whewell, there is little historical evidence to suggest that the epistemological aspects of his views were compelling or that people saw unity as a stamp of truth or as a reason to accept a particular theory.

?

On Friedman's account of unification he repeatedly emphasizes the importance of the confirmation that results from the ability to predict and explain a variety of phenomena. To that extent his views are similar to Whewell's and to traditional philosophical accounts of confirmation. However, according to Friedman this confirmation results from the fact that our theories evolve by conjunction. Although he claims that it is unifying power that serves as the criterion for a realist interpretation of theoretical structure, his case involves a more complicated methodology. The persistence or stability of particular structures through time enables our theories to evolve through a conjunctive process, thereby increasing their confirmation value (Friedman 1983, p. 245).

This approach is, however, at odds with most cases of consilience/unification, in which significant changes have been made to the laws and terms of the hypotheses involved, with the addition of new structures and entities. Although simple conjunction of hypotheses does not occur in cases of consilience, there *is* a reduction of entities, structures or hypotheses under the umbrella of one theory. This reduction obviates many of the traditional problems associated with Friedman's model/submodel account because it involves the unification of two groups of phenomena by means of a mechanism that involves theoretical, mathematical and semantical changes. No account of unification can be complete without recognition of the need for and the implications of this conceptual reshuffling. Consequently, much of the discussion of conjunction, together with the model/sub-model account of theoretical and observational structure, cannot be used to motivate a philosophical account of how unification, confirmation and realism are connected.

2.5. Unification as an Evidential or Epistemic Virtue

Earlier I claimed that as a matter of historical fact unity seemed to provide little in the way of empirical support for theories. The cases detailed in the following chapters will bear out that conclusion in what I hope is an unambiguous way. But my analysis here would be incomplete without some discussion of the philosophical status of unification itself as a justification for realism. Regardless of whether or not one adheres to the conjunction model of unification, there seem to be independent reasons for thinking that unification can sustain or motivate only a contextually based form of realism. Let me briefly explain why I take this to be so.

In discussing consilient theories it is important to note that a theory becomes consilient when it shows that phenomena originally thought to be of different kinds are in fact the same kind. This occurs only in relation to some other theory or set of particular beliefs and background knowledge about the phenomena, conditions that usually take the form of the currently accepted theory. For example, Newton's theory was consilient at the time of its emergence because celestial and terrestrial phenomena were regarded as distinct types. Had universal gravitation been proposed within the context of a Cartesian system, it would not have been considered consilient, because Descartes regarded both kinds of phenomena as due to the actions of similar types of vortices.[21] Friedman himself emphasizes that unifying power is a relative notion (1983, p. 249). He points out that absolute rest had no unifying power in the context of Newtonian gravitation theory. However, in the context of classical electrodynamics, absolute rest did have unifying power and therefore should be interpreted literally; but because we now no longer subscribe to classical electrodynamics, we can assign absolute rest a purely representative status.

This issue of historical relativism is perhaps the most important difficulty for a realist account of confirmation like Friedman's (one that equates consilient or unified theories with reasons for true belief or ontological commitment). Because the question of whether or not a particular piece of theoretical structure plays a unifying role becomes relativized to a specific context, it is difficult to see how this kind of virtue could be taken as evidence for truth or realism (unless, of course, both are construed in a purely local context). Specific ontological claims are legitimated on the basis of unifying power, but because these entities/structures perform a unifying role in some contexts and not in others, our beliefs become dictated solely on the basis of the historical contingencies involved in the unifying process. However, to talk of ontology and truth from within the confines of a particular theory is to collapse talk of truth and ontology into talk about the theory.[22] A strategy of that kind simply rejects what is right about realism, namely, the search for theory-neutral facts that can act as arbiters in theory choice and remain relatively stable in the face of theoretical restructuring. This notion of independence or neutrality seems to be sacrificed in an account that uses unification (in the context-dependent way Friedman describes) to motivate realism. Indeed, it is difficult to see how any distinction between the real and the representational at this local level could be extended to a global epistemology of science. Because a realism based on criteria like explanatory unification provides no ontological stability over time, it would seem that as a justification for *belief*, even at the local level of scientific practice, it proves unsatisfactory for realists and anti-realists alike. To the extent that theory confirmation depends on beliefs about the relative merit of the new theory as compared

Not for us, except for quasi ash[?]

That's semantic

Why stuck?

Certain is [illegible]

with its predecessor, it may always involve a certain amount of historical localization; but that is a different issue from the kind of variability that results from isolating a methodological process like unification as the criterion for a realistic interpretation of theoretical structure/entities. In the latter context, our commitments and beliefs become doubly abstracted; they are not simply relativized to a particular epistemic community, but to the domain of a unifying theory.

As mentioned earlier, one of Friedman's justifications for a literal construal of theoretical structure is that it allows for a persistence of that structure over time. But given the historical variability present in his own account, it is unlikely that such persistence could be guaranteed independently of a particular context.[23] Thus far I have tried to show some practical and philosophical difficulties with Friedman's particular account of unification. Not only is conjunction and the model/submodel approach ill-equipped to handle the ways in which theories and models function in practical contexts, but also the general strategy seems unable to support the kind of realism it was designed to defend. And once we acknowledge the relation between unification and particular theoretical contexts, it becomes difficult to provide an argument that links unity, truth and realism.

I have not yet said much about explanation, except to criticize Friedman's "inference to the unifying explanation". I have, however, suggested that unification is not the kind of criterion on which to base arguments for realism, and I have also hinted that it may not be important for theory acceptance either. In order to substantiate that argument with empirical evidence, I want to examine some specific and paradigmatic cases of theory unification in both the physical and biological sciences. One of my claims is that unification typically was not considered to be a crucial methodological factor in either the development or confirmation of the physical theories. And even in cases where it was a motivating factor, such as the unification of Darwinian evolution and Mendelian genetics, the kind of unity that was produced could not be identified with the theory's ability to explain specific phenomena. This is not to say that there is no evidence for unification in science or that unity is largely a myth, as some contemporary writers on disunity suggest (e.g., Dupré 1993, 1996). Rather, I want to demonstrate that the ways in which theory unification takes place and the role it plays in scientific contexts have little to ~~do~~ with how it has been characterized in traditional philosophical debates.[24] Once we have a clearer understanding of the unifying process, we can begin to see where its importance lies, what its connection is, if any, to explanation and the way unity functions in particular domains as well as in the broader context of scientific inquiry.

Appendix

Derivation of the van der Waals Law: Historical Details

The van der Waals law was originally formulated as a response to the idealizing assumptions of the Boyle-Charles law, which maintained that the sizes of molecules and the forces between them were negligible.[25] Although the molecules of real gases are assumed to have finite sizes and exert intermolecular forces at ordinary temperatures and pressures, real gases behave very much like ideal gases and thus in some situations obey the Boyle-Charles law. However, at sufficiently high and low temperatures, real gases can become liquified, thereby invalidating the application of ideal-gas laws to real gases. The foundations of the kinetic theory disregard the volumes and mutual attractions of the molecules, yet it is these attractions that account for such phenomena as cohesion, surface tension, the existence of a critical point and the phase transition (condensation).

It was Rudolf Clausius who initially suggested that the intermolecular forces that account for cohesion of the liquid phase must act throughout the range of temperatures and pressures. Their effects should be appreciable even in the gaseous phase, when molecules closely approach one another in collisions. In other words, because the nature of the substance was defined by its molecular model, the properties of the model should be present under all temperatures and pressures.

Van der Waals came upon this idea of continuity of the liquid phase and gaseous phase as a result of the work of Clausius on the virial theorem, which was an attempt to reduce the second law of thermodynamics to a purely mechanical form.[26] The theorem states that for a system of material points in which the coordinates and velocities of all the particles are bounded, the average (taken over long times) of the total kinetic energy is equal to the average of the virial:

$$\left\langle \sum \left(\frac{1}{2} m_i \mathbf{u}_i^2 \right) \right\rangle = \left\langle -\frac{1}{2} \sum \mathbf{F}_i \cdot \mathbf{r}_i \right\rangle \qquad \text{(A2.1)}$$

The brackets denote time averages, and the quantity on the right-hand side is what Clausius defined as the virial. The $\mathbf{r}_i$ denotes the coordinates of the ith particle, whose mass is m_i and whose velocity is $\mathbf{u}_i$, and $\mathbf{F}_i$ is the resultant force acting on it. The object was to trace the actual motions of the molecules that constituted the heat and to show that the effective force of heat was proportional to absolute temperature. The problem remained a purely mechanical one, with no appeal to probabilistic arguments to investigate the motions. Quite simply, the theorem states that for a system of material points in which the coordinates and velocities of all the particles are bounded, the mean kinetic energy of the system is equal to its virial. If the forces on the particles confined in a volume V can be divided into a uniform external pressure (from the container walls) and the central forces $\phi(\mathbf{r}_{ij})$ acting between particles, the theorem will take the following form:

$$\left\langle \sum \left(\frac{1}{2} m_i \mathbf{u}_i^2 \right) \right\rangle = \frac{3}{2} P V + \left\langle \frac{1}{2} \sum_{ij(i>j)} \mathbf{r}_{ij} \phi(\mathbf{r}_{ij}) \right\rangle \qquad \text{(A2.2)}$$

where $\mathbf{r}_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. Using the theorem, Clausius was able to derive a form for the second law only in the case of reversible processes; van der Waals, on the other hand, saw in it implications for the properties of matter.[27]

Although the virial theorem incorporated the possibilities of both the static and kinetic molecular theories, van der Waals did not calculate the molecular pressure P' from the average virial of intermolecular forces. Instead, he used a distinctly different approach to discuss the effects of the extended molecular volume and intermolecular attraction. Using a

series of assumptions about the mean free path,[28] he allowed for the fact that the molecules were of finite size.[29] In calculating the value of P, van der Waals argued that the effective force on a unit area of surface, arising from attractive forces between molecules, was the result of a thin layer of molecules below the surface. That followed from continuity considerations, which implied that the attractive forces acted over only a very short range.[30] The final step toward the completed equation of state involved replacing the average kinetic energy of the fluid by the expression proportional to the absolute temperature for one mole of ideal gas:

$$\left\langle \sum \frac{1}{2} m \mathbf{u}^2 \right\rangle = \frac{3}{2} RT \qquad \text{(A2.3)}$$

an assumption that could be argued for only on the basis of plausibility considerations. The equation in its final form was

$$(P + a/V^2)(V - b) = RT \qquad \text{(A2.4)}$$

The volume-correction term was only an approximation and was not valid at high compressions, a difficulty that van der Waals was unaware of. In addition, the pressure calculations were not completely satisfactory. The correction a arises from forces that the molecules exert on one another when reasonably near to one another. The correction b arises from forces that the molecules exert on one another when their centres are some distance apart. However, we cannot suppose that the forces acting on natural molecules can be divided up into two distinct types; they must change continuously with the distance. As a result, the a and b of the van der Waals equation ought to be different contributions from a more general correction, and so ought to be additive. The equation itself allows for no such correction. However, once a and b were determined experimentally, the isotherms (T) could be predicted for all P, V. The values arrived at using the van der Waals equation were confirmed experimentally in tests carried out by Andrews[31] on isotherms for carbon dioxide.

Conclusions

One of the themes I would like the reader to take away from this discussion is that unity, either in the broader context of science itself or in the more localized settings of specific theories, cannot be uniquely characterized. More specifically, no single account of theory unification can be given. A philosophical consequence of that claim is that unity should not be linked to truth or increased likelihood of truth; unification cannot function as an inference ticket. What exactly is the connection here? Although I have tried to explicate the features necessary for distinguishing a truly unified theory from a mere combination or conjunction, even within that framework different kinds of unity can emerge. Not only can unification exhibit a variety of ontological patterns (e.g., reductive and synthetic), but the way in which the unification is achieved can have a significant impact on whether or not there is an accompanying theoretical story, an account that can be claimed to provide a plausible representation of the phenomena. For instance, in neither the early version nor the late version of Maxwell's theory was there a viable physical interpretation of the electromagnetic field; hence, any affirmation of truth with respect to the theory would need to be severely limited in its content. Even if we consider Hertz's famous claim that Maxwell's theory was Maxwell's equations, prior to 1888 there was no guarantee that those equations were descriptively accurate, because there was no proof that electromagnetic waves existed. In what sense, then, would we be justified in claiming that Maxwell's theory, in virtue of its unifying power, was more likely to be true than not, or more likely than its rivals?

The second significant issue is that of empirical support. Intuitively one might want to claim that a unified theory has greater empirical support than two disparate theories simply because it can account for a larger number of phenomena with supposedly fewer assumptions. But here again we need to be mindful of the way in which the unification has been achieved. Obviously we want to minimize the number of free parameters in the theory, or, to use Weinberg's phrase, we want to increase the theory's rigidity. The goal is to have a tightly constrained unifying core applicable across a wide range of different phenomena. The electroweak theory has only one free parameter, the Weinberg angle, and although that parameter is representative of the theory's unity, its value is not derivable from the theoretical structure itself in the way that, for example, the value for the velocity of electromagnetic waves is a consequence of Maxwell's theory. Nevertheless, taken

by itself, the electroweak theory certainly qualifies as a unified theory in the sense I have described, unlike, say, the standard model, which has no fewer than 18 arbitrary parameters, including particle masses, the number of charged leptons and so forth. These parameters need to be inserted into the calculations at the appropriate moments in order for the theory to account for the various particle interactions and to determine the forces between the various quarks and leptons. But these parameters are more than just calculational tools; they provide the foundation on which current particle physics is based. In fact, like the Weinberg angle, it is these parameters that function as the unifying core of the standard model; yet because they are "free" parameters, there is no apparent or discernible connection between them. And therein lies the problem.

In order for the theory to be truly unified, some mechanism linking these parameters must exist. Moreover, as long as there is no understanding of why the parameters have the values they do, the theory cannot explain at a fundamental level even why atoms and molecules exist. In other words, an explanation of the origin of these parameters is crucial for a basic understanding of the physical world. We see, then, that it isn't simply the number of parameters that creates a problem for unification; the mysterious nature of the parameters creates a corresponding difficulty for explanatory power. In the electroweak case we have only one arbitrary parameter, and, as we have seen, its status creates similar problems for the explanation of electroweak coupling, problems traceable to the mass of the Higgs boson. Although it and the standard model enjoy a great deal of empirical support, the unity provided by the standard model as a whole is surely suspect.

This issue of empirical support has another dimension that has consistently reappeared as a theme throughout the preceding chapters. In each of the examples I have discussed, a general and rather abstract mathematical framework has played an essential role in the unification of disparate phenomena. Even in Fisher's work his statistical analysis of genetics was based on a model from gas theory that incorporated only general assumptions about populations. The significance of these frameworks is that their generality allows one to ignore specific features of the systems/phenomena in question, thereby enhancing the potential for unification. The more general the structure, the more phenomena it is able to incorporate, and hence the greater the empirical support. But typically this is achieved at the expense of a detailed theoretical account that explains the mechanisms and processes associated with the phenomena. Hence the empirical support is limited to quantitative predictions. But philosophers of science know all too well the arguments against linking predictive power with increased likelihood of truth.

Given my arguments against the alliance of unification, truth and explanatory power, some might be tempted to associate my position with those of the "disunifiers", philosophers, historians and sociologists who argue against the unity of science on various grounds – metaphysical, epistemological, political. Although I am sympathetic toward arguments against a general unity of science, and I oppose arguments that attempt to infer a unity in nature on the basis of unified theories, obviously I do not want to deny that unity does exist within the confines of particular disciplines and theories. My goal has been to uncover the nature of that unity, the ways in which it is achieved and the kinds of consequences it can support with respect to the ontological status of theories. There can be no doubt that unity exists in science, that unified theories have been enormously successful and that unity is a goal pursued by many practicing scientists in a variety of fields. But nothing about a corresponding unity in nature follows from those facts; that is, we cannot, simply on the basis of the existence or success of unified theories, infer a unified world. In order for this negative thesis to be persuasive, one must first expose the foundations and structures of unified theories. Once that is done, it becomes possible to see the gap between the simple conceptual design of the theory and the complexity of the system it represents. The unification of electromagnetism with the weak and strong forces involves energies on a scale trillions of times beyond the reach of current instrumentation. A rather daunting problem indeed – one that led Sheldon Glashow and his colleague Paul Ginsparg to write that "for the first time since the dark ages we can see how our noble search may end, with faith replacing science once again". In that context the disunity arguments become extremely persuasive.

On the other hand, in contexts less ambitious than grand unification, where physics is somewhat less theological, it becomes difficult to dismiss undisputed facts such as the claim that light is a form of electromagnetic radiation. Although our understanding of electromagnetism has changed since the formulation of Maxwell's equations, it would be remiss to suggest that somehow the unity had not been preserved. However, as we saw in the preceding chapters, the unity produced by Maxwell's electrodynamics is in many respects unique. It effected a kind of ontological reduction that few, if any, subsequent theories have been able to match. However, ontological reduction in a case like Maxwell's theory needs to be distinguished from the kind of reduction that occurs when one theory supersedes another. It is common to argue that because newly constructed theories typically embody, as limiting cases, the things that were correct in their predecessors, we should see this as evidence for some form of convergence or unity in science. However, rather than contributing to scientific unity, such cases, on closer inspection, tend to reveal inconsistencies that actually speak against unity and convergence. Consider, for example, the replacement of Newtonian gravitation theory with general relativity. The fundamental ontology of the latter is curved space, rather than gravitational force; yet in the Newtonian limit the qualitative interpretation of the mathematics yields a flat space and gravitational forces. The problem is not simply that we want to retain the Newtonian ontology for particular purposes and ease of calculation, but rather that general relativity itself embodies two distinct ontologies that say different things about the constitution of the physical world. In that sense, then, the subsumption of old theories as limiting cases of new ones affirms the kind of disunity we expect from a science marked by periods of sweeping change. Nevertheless, the unity achieved in cases like Maxwell's theory, special relativity and evolutionary genetics requires us to take seriously the role of unification in scientific discourse and practice.

Hence, my position falls squarely between the two camps; arguments for unity and those for disunity have important lessons to teach us. However, viewed as a metaphysical thesis, the problem takes on a structure resembling a Kantian antinomy, where the truth of one of the disjuncts (unity or disunity) supposedly implies the falsity of the other. The resolution consists in showing that the thesis and antithesis (in this case, unity and disunity) are mutually compatible, because neither refers to some ultimate way in which the world is constructed. As empirical theses about the physical world, both are right, but as metaphysical theses, both are wrong, simply because the evidence, by its very nature, is inconclusive. To that extent the problem can be seen as presenting a false dichotomy. Neither position is adequate for providing a descriptive, explanatory or metaphysical account of nature; yet in certain contexts and for certain purposes, either may be sufficient to capture the phenomena at hand. As we saw in the discussion of specific theories, different kinds of unity were appropriate for different situations, and unity sometimes was coupled with an element of disunity.

My argument has focused on the rather narrowly defined issue of theory unification, but there are other forms of unity that may exist to greater or lesser degrees at the general level of scientific practice. The logical-empiricist goal of finding a common language that would bring unity to the sciences has not been and cannot be achieved; nevertheless, one might claim that there is a kind of unity that exists within specialized disciplines and sometimes even across disciplines – a unity produced by practices involving commonality of mathematical methods, instrumentation, measuring techniques and so forth. All of these function to produce a scientific community of the sort described by Galison (1998) in his metaphor of the "trading zone". Those unities are different in kind from the unity I have addressed in the preceding chapters. The issue of unity in nature need not arise in discussing unity of practice. And although such practical concerns are important for understanding the culture of science, they bypass the fundamental question that motivates much of the debate on scientific unity, specifically the ontological problem of whether or not nature is itself a unified whole. I have tried to address that particular issue by showing that the ontological question is, in most cases, distinct from the ability to construct a unified theory. In fact, my claim is that it is this lack of ontological unity that is partially responsible for the separation of unity and explanation at the theoretical level.

Throughout this book I have maintained that unification and explanation are often at odds with each other. Distinguishing between unity and explanation is important not only for a better understanding of the nature of theory unification but also to expose the fact that inferential practices grounded in the explanatory power of a theory often are simply inapplicable to the theory's unifying power. Initially that may seem rather odd, because it is usual to assume that the more phenomena that are unified within a theoretical framework, the more phenomena will be explained. We have seen, however, why that doesn't follow. In my critical discussion of the division between explanation and unification, little attention was paid to exactly what a scientific explanation should consist in. I spoke briefly about the desire to know the hows and whys of physical systems, but I did not specify a form that good explanations must instantiate. I have remained deliberately silent on the subject primarily because it is, in itself, the topic of a separate book and one that many philosophers of science have already addressed in detail. That said, there is a feature of explanation that everyone would agree on, namely, the desire to impart understanding. Yet in many of the theories I have discussed, the information required for understanding the fundamental processes that figured in the unification was simply unavailable. The electroweak theory provides a "how possibly" story that rests on the discovery of the Higgs boson. Minkowski space-time enhances the unifying power of special relativity, and yet it provides no understanding of its relativistic dynamics over and above what had already been explained by Einstein. But here again we have a theory that is itself capable of explaining and unifying diverse phenomena, and yet much is left unexplained at the level of its own fundamental postulates. In that sense, then, when speaking about explanatory power we must distinguish between the phenomena that the theory "accounts for" and a physical explanation of the theory's fundamental structure. Although all of the theories I have discussed followed similar patterns for producing unity, the unification manifested itself differently in each case; different levels of reduction, synthesis and integration yielded distinct ways in which phenomena could be united under a common theoretical framework. Accompanying those disparate representations of theoretic unity were diverse levels of explanation; yet what was most significant in each of the cases was that the processes associated with the piece of theoretical structure necessary for the unification were themselves left largely unexplained. To put it differently, the mechanism responsible for unity contributed nothing to the explanatory power of the theory.

It is important to understand that this is not simply an instance in which the request for explanation must terminate, as, for example, in the search for certain kinds of explanations of quantum systems. Nor is it a case of contextual opacity with respect to the information that is being requested or transmitted. In each of the examples I have discussed, a fundamental process has been left unexplained in the face of a remarkable feat of unification; consequently there is a basic lack of understanding as to how (or even whether or not) the unifying process actually takes place in nature. In some cases this speaks directly to the limits of science, imposed as a consequence of different levels of technological advance. We are simply unable to probe the fundamental building blocks of nature with the energies currently available in particle accelerators. However, in other cases the problem is not one of instrumentation, but rather lack of the theoretical knowledge needed to fill in the missing details. Natural selection and genetics provide a case in point; there are still debates as to whether or not and to what extent selection functions in certain contexts. Similarly, in Maxwell's electrodynamics there was a good deal of uncertainty about the nature of the field, a controversy that remains in the transition from classical theory to quantum field theory, and the relationships between fields and particles. In each of these cases the stumbling blocks to unity and

understanding are empirical, grounded in science itself rather than in philosophical objections about the proper form for unification and questions about whether or not that has been achieved.

To conclude, let me summarize the connection between the philosophical and empirical parts of my argument. The motivating question is whether or not theory unification provides evidence for unity in nature. In order to answer that question, we first need to know what unification consists in. More specifically, is it connected with explanatory power, and if so, what is the link? Finally, how do these issues bear on the metaphysical thesis regarding a unified physical world? To put it slightly differently, how should we understand the unity that science has achieved? Has it achieved unity at all? And are there philosophical problems associated with this unity? These kinds of questions involve a philosophical dimension and an empirical dimension; in other words, we must begin by telling an empirical story about unity that is grounded in historical documentation before we can draw philosophical conclusions. My general conclusion is that it is a mistake to structure the unity/disunity debate in metaphysical terms; that is, one should not view success in constructing unified theories as evidence for a global metaphysics of unity or disunity. Whether the physical world is unified or disunified is an empirical question, the answer to which is both yes and no, depending on the kind of evidence we have and the type of phenomena we are dealing with. In some contexts there is evidence for unity, whereas in others there is not. But nothing about that evidence warrants any broad, sweeping statements about nature "in general" or how its ultimate constituents are assembled. The construction of theories that unify phenomena is a practical problem for which we only sometimes have the appropriate resources – the phenomena themselves may simply not be amenable to a unified treatment. Seeing unity in metaphysical terms shifts the focus away from practical issues relevant to localized contexts toward a more global problem – erecting a unified physics or science – one that in principle we may not be able to solve because nature may not wish to cooperate. The methodological point I want to stress is that eliminating metaphysics from the unity/disunity debate need not entail a corresponding dismissal of philosophical analysis. Philosophical methods and concepts are important for analysing the structure and implications of the evidence for both unity and disunity, but they should not be used as means for extending the domain of that evidence to claims about scientific practices and goals that are merely promissory. Debates about unity and disunity become productive when they take as their starting point an analysis of the accomplishments of empirical research. Science undoubtedly has achieved a certain level of unity, but in order to understand its limits, we need to know its nature. Uncovering the *character* of unity and disunity is a philosophical task, one that will contribute to the broader goal of better understanding the practice of science itself. And given the progress and methods of empirical science, there is perhaps nowhere that metaphysics is less helpful.