1987

Dimensions Of Theory Acceptance: Methodology And Experiments

Margaret Catherine Morrison

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ MICROFILMÉE TELLE QUE NOUS L'AVONS REÇUE

DIMENSIONS OF THEORY ACCEPTANCE: METHODOLOGY AND EXPERIMENTS

by

Margaret Catherine Morrison

Department of Philosophy

Submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario

June, 1987

ISBN 0-315-36589-7

THE UNIVERSITY OF WESTERN ONTARIO
FACULTY OF GRADUATE STUDIES

CERTIFICATE OF EXAMINATION

Chief Advisor

L. E. Butts

Advisory Committee

J. Nicholas

Examining Board

R. B.

William J. Hayes

J. Nicholas

The thesis by
Margaret C. Morrison

entitled
DIMENSIONS OF THEORY ACCEPTANCE:
METHODOLOGY AND EXPERIMENTS

is accepted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Date 17 July 1987

Thomas K. Brown

Chairman of Examining Board

Thesis Abstract

Recent arguments for scientific realism have emphasized the importance of both methodological factors, such as theoretical unification [Friedman 1983], and experiments [Cartwright 1983 and Hacking 1983], as evidence for a realistic view of certain aspects of theoretical structure (and entities). Throughout this dissertation I argue that neither strategy is sufficient as a defense of realism.

Chapter one consists of a discussion of Friedman's argument for realism as outlined in his **Foundations of Space-Time Theories** (Chapter VII). I argue that his reliance on theoretical unification and conjunction as grounds for a selective brand of realism is unsatisfactory for a variety of reasons. Not only does his structure-substructure reductivist model of theories and the emphasis on conjunction fail to fit actual cases of theory evolution, but the historical relativity built into the approach is undesirable even from a realist perspective. The relationship between conjunctive inference and unification is discussed and it is suggested that although theoretical unification is desirable at some stages of theory development it contributes nothing to the justification or evidential warrant of theories.

In chapter two my claims regarding unification are applied to the development of Maxwell's electromagnetic theory. This case was chosen because it provides, in my opinion, the strongest case for the kind of straightforward

reduction/unification that is characterized by Friedman's model. There were few corrections to the laws of the theories involved (save for some changes in the theory of physical optics) unlike the case of universal gravitation, and little or no semantical change in the theory itself. I discuss the difficulties in specifying a workable model of the aether as the unifying theoretical structure as well as the importance of models and analogies in Maxwell's work. It is argued that this obvious unification of electromagnetism and optics played virtually no evidential role in the final acceptance of the theory. It was not until the completion of Hertz's experiments that Maxwell's theory was actually vindicated.

Chapter three focuses on the experimental difficulties associated with Maxwell's theory. I discuss the relationship between Hertz's experiments and his theoretical interpretation of electromagnetism. The historical discussion raises some interesting philosophical issues regarding the way in which experimental evidence bears on theory.

In chapter four I summarize some of the general difficulties associated with the notion of independent evidence as it pertains to experimental results. In addition to the discussion of Hertz's experiments some current examples from high energy physics are briefly discussed. Some problems for Hacking's account of entity realism are raised. I conclude by arguing that although experiment provides a more persuasive case for realism than methodological factors, in many instances it fails to provide the kind of epistemological justification required for the realist assumptions about the ontology of scientific theories.

Acknowledgements

I would like to thank my teachers, particularly Kathleen Okruhlik and James Brown who first aroused my interest in Philosophy, as well as the members of my advisory committee; all of whom have been a constant source of encouragement. A very special thank you goes to my supervisor Robert Butts, his generosity, support and friendship have been invaluable to me. Thanks also go to Phillip Catton for stimulating discussion and helpful comments on earlier drafts and to Bas van Fraassen for reading and talking with me about what were the very rough beginnings of chapters one and two. I have benefited from Paul Forster's keen philosophical insight in ways too numerous to mention. His love, patience and sense of humour have sustained me throughout some of the more difficult times. Finally a thank you to my parents for their faith, love and support. They have been an unending source of strength when I needed it most. Support of research by the Social Sciences and Humanities Research Council of Canada and the province of Ontario in the form of a Queen Elizabeth II Scholarship is gratefully acknowledged.

TABLE OF CONTENTS

Chapter One

The Problem of Theoretical Inference: A Critique of Friedman's Approach

<u>1. Introduction.</u> Realist philosophers of science have typically sanctioned the move from asserting that a theory T is the best explanation of a phenomenon P to asserting the truth of T; or, more modestly, that explanatory power provides good evidential warrant for the truth of T. Criticism of this practice, known as inference to the best explanation, can be found in the writings of van Fraassen [1980], Cartwright [1983] and Friedman [1983]. The arguments are varied, some emphasize the fact that explanation has to do with providing answers to questions, organizing and systematizing our knowledge; pragmatic features that do not provide evidence for the literal truth of the background theory assumed in explanatory contexts. But, even for those that disagree about the pragmatic status of explanation, the best available explanation may not be the one that we would want to accept, even provisionally. Friedman opposes inference to the best explanation 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 <u>ever</u> be taken literally. 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 a vis the observable realm. Friedman's

solution to this problem consists not in giving up this method of inference but rather of restricting its applicability. He argues that theoretical inference can be sanctioned when accompanied by unification. 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.

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. Friedman thinks that any theoretical structure not participating in unification can be treated as purely representational without any adverse consequences for the theory in question.

Much of Friedman's discussion focuses on the role of conjunctive inference in scientific methodology and in facilitating this unification. He embraces the traditional Putnam-Boyd objections to anti-realism by emphasizing its inadequacy in accounting for theory conjunction. This is seen as a problem because Friedman claims that theories evolve and are unified by a conjunctive process. In the context of this discussion it is important to point out that this so-called conjunction objection has two dimensions and as a logical point about the nature of conjunctive inference it should be distinguished from the methodological claim that theories in fact evolve and are unified by a conjunctive process. Because this distinction becomes somewhat blurred in Friedman's discussion part of my critique will be an attempt to clarify the issue as it arises in scientific contexts.

The reductivist approach outlined by Friedman raises several issues that are frequently conflated in philosophical discussions of theoretical unification or consilience. Perhaps the most important of these is the distinction between the mechanisms involved in unification and the more straightforward operation of theory conjunction. Although Friedman emphasizes the role of unification he fails to distinguish between a true consilience of inductions in the traditional Whewellian sense, and the kind of broad sweeping theory one gets as a result of conjoining several individual hypotheses.

If Friedman's approach is to be a viable one some account of reduction that can incorporate the semantic and mathematical changes necessitated by a consilience/unification is required. One of Friedman's examples focuses on the kinetic theory of gases and the literal reduction of the observable properties of gases to their molecular configurations. The relationship between these two levels is characterized as that of model to submodel. However, it is important to point out that as the kinetic theory evolved the molecular models underwent drastic changes with the behavior of gases often being explained by a variety of different and seemingly incompatible models of the phenomena. In light of this the status of Friedman's truth-preserving reductions becomes questionable. If the observable structure of our theories is to be construed as a submodel of the larger theoretical structure (model) as 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. Moreover, it is frequently the case that a particular empirical law and its approximations (van der Waal's law for example) requires different theoretical models depending on its application. If we interpret our model as providing a literally true description of the world then it becomes difficult to envision how these aspects of theory are to be accounted for. Throughout the paper I will attempt to show that the model-submodel approach is an ineffective way to characterize scientific theories; that the logical constraints on the model submodel relationship are too tight to allow for the kind of looseness of fit that exists between the theoretical and observable structures of theories.

To fill out some of these rather sketchy points I begin with a brief summary of Friedman's argument as applied to the kinetic theory of gases. From there I will discuss his account as it relates to some of the traditional views on theory reduction, conjunction, unification and consilience in an effort to isolate what I perceive to be the difficulties with the overall approach.

<u>2. The Friedman Model.</u> We characterize a typical scientific explanation in the following manner: We postulate a theoretical structure $\mathfrak{A} = \langle A, R_1, \ldots R_n \rangle$ (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 $\mathfrak{B} = \langle B, R'_1, \ldots R'_m \rangle$ $(m \leq n)$. $\mathfrak{A}$ functions as an explanation or reduction of the properties of $\mathfrak{B}$. Using the kinetic theory we can explain the observable properties of gases characterized by $\mathfrak{B}$ by embedding them in $\mathfrak{A}$ where $\mathfrak{A}$ is literally construed as the world of molecular theory. This enables us to account for the behavior of gases by identifying them with large configurations of molecules that interact according to the laws of Newtonian mechanics. Due to the properties and relations provided by the theoretical structure we can derive laws that govern the behavior of observable objects. By contrast, if we remained strictly on the phenomenological level we would not be able to accurately formulate a law like 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 $\mathfrak{A}$ and $\mathfrak{B}$ as that of model to submodel where $B \subseteq A$ and $R'_i = R_i$ $(i \leq m)$. This characterization affords us a literal identification of the elements in $\mathfrak{A}$ and $\mathfrak{B}$ 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 (cf. p.240). Contrast this with what Friedman terms the representational account. On this view we do not interpret $\mathcal{A}$ literally (as the molecular world), rather, it is construed as a mere representation. Instead of asserting that $\mathcal{B}$ is a submodel of $\mathcal{A}$ we claim only that $\mathcal{B}$ is embeddable into $\mathcal{A}$; there exists a one-one map ϕ :B --> A such that $\phi(R'_i) = R_i / \phi(\mathcal{B})$ $(i \leq m)$. $\mathcal{A}$ does not "induce" the necessary theoretical properties on objects in $\mathcal{B}$ unless of course these 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 submodel interpretation due to 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 <u>conjoin</u> 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) which it "transfers" to the phenomenological description. Without this transfer of confirmation the phenomenological description receives confirmation only from the behavior 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 <u>ad hoc</u> (p.241).

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. For example where Δ_1 and Δ_2 are classes of models a reduction facilitates the inference

$$\begin{array}{l}
\langle B, R_1 \rangle \subseteq \mathcal{A} \text{ and } \mathcal{A} \in \Delta_1 \\
\langle B, R_2 \rangle \subseteq \mathcal{A} \text{ and } \mathcal{A} \in \Delta_2 \\
\hline
\langle B, R_1, R_2 \rangle \subseteq \mathcal{A} \text{ and } \mathcal{A} \in \Delta_1 \cap \Delta_2
\end{array}$$

whereas the nonliteral interpretation restricts the inference to the following form:

$$\begin{array}{l}
\exists \mathcal{A}\, \exists \varphi : \langle B, R_1 \rangle \rightarrow \mathcal{A} \text{ and } \mathcal{A} \in \Delta_1 \\
\exists \mathcal{A}'\, \exists \psi : \langle B, R_2 \rangle \rightarrow \mathcal{A}' \text{ and } \mathcal{A}' \in \Delta_2 \\
\hline
\exists \mathcal{A}''\, \exists \chi : \langle B, R_1, R_2 \rangle \rightarrow \mathcal{A}'' \text{ and } \mathcal{A}'' \in \Delta_1 \cap \Delta_2
\end{array}$$

This latter inference is invalid because not only are $\mathcal{A}$ and $\mathcal{A}'$ 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 individual boosts in confirmation at times 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 ($\exists\varphi(A), \exists\psi(B)$) 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.

Although there is little consensus about the advantage of predictive power over explanatory power (cf. Horwich [1980]) Friedman is nonetheless quite right about the shortcomings of the representationalist/instrumentalist programmes in accounting for the role of conjunctive inference. Quite independently of whether theory evolution or unification actually relies on conjunction the logical point, first raised by Putnam in "Explanation and Understanding", still stands. That is, accounts of truth that appeal to notions of warranted assertability, successful prediction and other instrumental values cannot account for the logical principle that the conjunction of a true theory T_1 and a true theory T_2 will yield a true theory T_3. In our acceptance of logical rules of inference we acknowledge that certain moves will be truth preserving; so, it would seem that ruling out the practice of conjunctive inference results in the inapplicability of particular logical rules to scientific theories. Prima facie this seems absurd, why should accepted principles of reasoning be inappropriate in these contexts; ergo realism is our only option if we are to make sense of our inferential practices in scientific contexts. Perhaps this argument can vindicate a particular brand of metaphysical or semantical realism that advocates the correspondence theory of truth but it is far from clear that scientific realism (a realism that advocates true belief in theoretical entities and in the truth or approximate truth of our theories) has been swept along in the wake.

Consider the claim being advanced by Putnam (a version of which is supported by Friedman). According to Putnam [1975] when a scientist accepts a theory he believes it to be true; it is only by having such a belief that we are able to perform the appropriate conjunctions. In other words, because theory conjunction is a desirable virtue our epistemic attitudes must be such as to allow for this practice. Friedman takes the argument one step further and claims that the products of these conjunctions, the part of the theoretical structure that unifies the other parts, is what is to be believed or interpreted literally. However, since 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. So, 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. We can't simply limit belief to the unifying part of the theory, we need a stronger form of realism to motivate this model of unification. However, this assumption of truth and literal reduction is just what is at issue as far as the scientific anti-realist is concerned.

Nor can we get around the problem by using approximate truth. 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.[1]

As I mentioned above it is possible to retain a version of realism that preserves the semantical claim that our theories are either true or false, and reserves the right to appeal to the conjunction rule in cases where we have true theories. But, nothing follows from this brand of realism about the actual truth of the theory(s) in question. This kind of realism denies the tenets of classical instrumentalism by interpreting theoretical structure literally. Hence we are able to appeal to theoretical structure for derivations and entailments of phenomenological laws and need not rely solely on phenomenological properties to provide us with the impetus for theoretical explanation. A semantical realism of this sort constitutes part of what van Frassen has termed "Constructive Empiricism". Although the theory is interpreted literally there is no requirement that every part of the theory has a counterpart in reality. The important subtlety in van Fraassen's position that guards against the difficulties of instrumentalism, the kinds of difficulties Friedman alludes to, is the distinction between theoretical structure being the kind of thing that can't be interpreted literally (and should only be used instrumentally) and theoretical structure as something capable of existing; something about which true or false claims can in principle be made without our being in a position to justifiably make such claims." Instead on a view like van Fraassen's we have an "in principle" form of realism whose demands cannot be answered in practice. If our theories are true then we can successfully conjoin them, however, this knowledge of truth conditions is not something we can acquire.

Independent of these logical issues about truth and inference is the question of whether the actual practice of science and evolution of theories can be modelled on the kind of approach described above; whether in fact science proceeds in accordance with the kind of logical rigour that philosophers are fond of imposing on it. It is important however to distinguish between the "in principle" use of conjunctive inference (the logical issue) and its legitimation in specific instances. No one objects to the use of conjunction as a logical rule that guarantees truth, if we begin with true conjuncts. But, 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 motivation behind the initial formulations of the conjunction objection (Putnam/Boyd) was the belief that our theories were <u>true</u>. But, as many realists conceded, this was to demand too much from science and the theories it produces. If we interpret truth in the traditional sense (the way most realists do)—thinking of it as a timeless property, then it becomes difficult to envision a scientific theory as something to which we would ascribe the property "true". The problem with the realist position is that anything less than truth cannot comply with the requirements of the conjunction rule. In reply the anti-realist simply denies that we have knowledge of the truth values of our theories, the kind of knowledge that is required to validate or guarantee the outcome of our conjunctive practice.

However, the applicability of conjunctive inference involves not only this logical issue but the validation of scientific methodology as well. Friedman claims that our theories evolve by conjunction while Putnam maintains that conjunction is a move that scientists frequently make and one that is central to their practice if scientific inquiry is to have any cumulative character. In order to use these methodological considerations as an argument for believing in the truth of theories one must show the indespensibility of conjunction for theory evolution and the validation of scientific methodology. In what follows I shall try to show the difficulties involved in establishing such a position. But first, recall that Friedman's programme stresses the importance of reduction in facilitating conjunctive inference. In order to deal adequately with the issue of conjunction some examination of the assumptions involved in reduction is required. By addressing actual cases of theory reduction we can perhaps get a better understanding of the nature of the logical models of reduction and their role in scientific methodology. This discussion of reduction is followed by a detailed account of conjunction and its relationship to theory unification and evolution.

3. Reduction vs. Representation.

3.I. Is Reduction a viable approach? The notion of reduction in physical science encompasses a variety of relationships between theories, their concepts and ontologies. Reductions are usually characterized as being heterogenous (domain-combining) or homogenous (domain-preserving). The former achieves postulational and ontological economy by reducing one level of phenomena to another. This typically involves a derivational reduction and provides an explanation of one theory by another. Dupré [198[illegible]] calls, this synchronic reductionism; a reduction that deals with the relations between co-existing theories that each address different levels of organization. Homogenous reduction does not normally achieve this kind of economy nor does it supply deductive explanations of predecessor or lower level theory by a successor or higher level one. In the context of these homogenous reductions the original theory is usually retained as a limiting case of the successor; as in the case of special relativity and classical mechanics, for example. Synchronic reductions involve the identification of distinct entities such as light and electromagnetic radiation as well as the reduction of particular phenomena like the behavior of gases to their ultimate constituents or causal properties, namely configurations of molecules. These latter kinds of reduction also encompass the reduction or derivability of laws, for example, the derivability of Kepler's and Galileo's laws from Newton's law of universal gravitation. In this case the relationship between these laws provides a further ontological reduction or unification of celestial and terrestial forces. A similar type of situation occurred with Maxwell's discovery of the electromagnetic theory of light. Maxwell's first derivation of the theory relied on his mechanical model of the aether as an elastic solid and the relationship between electromotive force and the displacement current.[3] He calculated the velocity of propogation of transverse disturbances through the aether and found that the value agreed with the value for the velocity of light; a discovery that resulted in the inference that light consisted in the transverse undulations of the same medium which is the cause of electric or magnetic phenomena.[4]

The difficulties with these reductions is that they often involve substantial changes in the meanings of the terms involved, corrections to the laws, addition of auxilliaries and quite simply, problems identifying the reduced phenomena in the terms of the reducing theory. In other words, the straightforward "reductions" discussed in abstraction are rarely, if ever, realized in practice. Given these difficulties the problem becomes one of understanding the conjunctive process facilitated by this reductivist model. An additional and perhaps more important consideration is whether we can uphold the logical principle of conjunction (considered as a methodological tool) independently of the success or failure of the reductivist strategy.

A classic example of identifying the properties in one theory with those in another arises in the reduction of thermodynamics to statistical mechanics.[7] Thermodynamics involves many concepts, and general laws that are also employed in mechanics; for example, the notions of volume, weight, pressure, as well as the laws governing the lever and Hooke's law. However, there are other concepts such as entropy, temperature and heat that are specific to thermodynamics but are not easily incorporated into the statistical framework. Although thermodynamics can be understood without the relationship to mechanical systems and a reliance on microscopic structures of thermal systems, nineteenth century work on the kinetic theory of gases revealed an intimate connection between thermal and mechanical phenomena. On the basis of mechanical assumptions about the molecular constitution of ideal gases Maxwell and Boltzmann were able to derive the Boyle-Charles law in a straightforward manner. Similarly, Boltzmann was able to interpret the entropy principle as an expression of the statistical regularity that characterized the aggregate mechanical behavior of molecules. The difficulty involved in the reduction is the identification of temperature with mean kinetic energy; an assumption that plays a crucial role in the derivation of the Boyle-Charles law.

The problem arises in the following way: Assume we have an ideal gas in a container having perfectly elastic walls and a volume V. The gas is said to be composed of a large number of perfectly elastic spherical molecules with equal masses and volumes whose dimensions are small relative to the average distances between them. The molecules are in constant relative motion and subject only to the mechanical forces of impact between themselves and the walls of the container. The problem is to calculate the the relationship of the other features of their motions to the pressure exerted by the molecules on the walls of the container. Classical mechanics is inapplicable here because the state coordinates of individual molecules are unknown. Instead we must introduce a statistical assumption regarding the positions and momenta of the molecules. The volume of the gas is divided into a number of equal smaller volumes whose dimensions are large relative to the diameters of the molecules. In addition, the maximum range of velocities that the molecules may possess is divided into a large number of equal intervals; with all possible velocity intervals being associated with each small volume. Each volume associated with a velocity interval is termed a "phase-cell". The statistical assumption is that the probability of a molecule's occupying an assigned phase-cell is the same for all molecules and is equal to the probability of its occupying any other cell. In addition the probability that one molecule occupies a cell is independent of whether any other molecule does. Together with these assumptions we stipulate that the pressure p exerted by the molecules at time t against the walls of the conatiner is the average of the instantaneous momenta transferred from the molecules to the wall. From this information we can deduce that p is related to the mean kinetic energy E of the molecules in the following way:

$$p = 2E/3V \text{ or } pV = 2E/3 \qquad (1)$$

If we compare this to the Boyle-Charles law (which states that when T equals absolute temperature and k is a constant for a given mass of gas $pV = kT$) we see that we can only deduce this law from our assumptions if temperature is related to the mean kinetic energy of molecular motion. In other words, we must introduce the postulate that the absolute temperature of a gas is proportional to the mean kinetic energy of the molecules that constitute it, i.e.

$$2E/3 = kT. \qquad (2)$$

As Sklar [1967] and others have pointed out the identification of temperature with mean kinetic energy is less than straightforward.[6] Within the parameters of the statistical theory we are provided with alternative concepts of mean kinetic energy based on different ways of averaging; for instance, we can use ensemble averages, time averages or a variety of other approaches. And, although the theoretical concepts are different the various values of mean kinetic energy are the same. The difficulty is deciding which statistical concept to identify with temperature. Each alternative results in different patterns of resemblances between the two theories. Attempts to identify the new concepts like heat, temperature and entropy with their traditional counterparts results in similarities that are contingent on our choice of identifications.

If we look at temperature ratios within thermodynamics we can see that they are defined by reference to reversible processes of operating between two levels L and L'; each of which is characterized by the same temperature. The ratio of the temperatures with the ratio between the amount of heat absorbed at the higher level and the amount rejected at the lower lever is defined by

$$T:T' = Q:Q'.$$

(3)

The concept of temperature (the established usage) requires that it be a unique value and independent of the material of the substance chosen for the cycle. Taken together these conditions imply the second law of thermodynamics in its phenomenological form; hence, an application of the thermodynamic concept of temperature to concrete situations entails the non-statistical second law. This concept is not to be found within the domain of the kinetic theory; no dynamical concept possesses the requisite property. The statistical theory allows for fluxuations of heat back and forth between two levels of temperature in opposition to one of the laws inherent in the thermodynamic concept of temperature.[7]

A similar difficulty arises in specifying the relationship between the thermodynamic concept of entropy and its statistical counterpart. The former is measurable by infinitely slow reversible processes only while the latter allows for a broader application.[8]

The problems associated with the reduction and identification of the domains of thermodynamics and statistical mechanics dealt with the identification of concepts. However, the idealized nature of the assumptions concerning the probability of a molecule's being at a certain phase cell (as well as the independence condition) expose a more serious problem, namely, that the global properties of the statistical system are not simply the result of the collective properties of the individual parts. In other words a literal reduction of the whole to its parts is impossible. Consider the following case. 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 is a function of these n variables. $E_1 = E_1(x_1 \ldots x_k)$ and depends on some of the dynamical coordinates while $E_2 = E_2(x_{k+1} \ldots x_n)$ 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.* 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. These components characterize pure energy (in the sense given above by the definition of a component). For example, consider a system of one material particle with components of velocity and mass being u,v,w,m; if its energy E reduces to kinetic energy 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. But, because each component is a group of dynamic coordinates and has a definite energy it has its own phase space; with 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 each of which, taken together, determine the structure function of the given system. Indeed the law governing composition of structure function is one of the most important formulas in statistical mechanics.

An additional problem is the methodological paradox that arises from the decomposition of the system into components is the 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 total energy of a gas is expressed as the independent energies of the two components (the energies of the molecules), 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 representing the Hamiltonian of this particle), then the entire system of equations governing the motion of the system[10] splits into component systems. Each component system describes the motion of some separate particle and is not connected to any other particle.[11] As a result the energy of each particle expressed by its Hamiltonian function appears as an integral of equations of motion and remains constant.[12] From the fact that the particles are independent and the sum of the energies is constant it follows that the individual energies must be constant as well. But, since 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.

The difficulty is resolved by idealizing assumptions that consider particles of matter as <u>approximately</u> isolated energetical components. Although the precise characterization of energy contains terms that depend simultaneously on the energy of several particles, and allow 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 to 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 the calculation of the Boyle-Charles laws,* 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 since they provide the basis for an understanding of energy exchange between particles; the very core of statistical mechanics.

What I think these examples illustrate is that the reduction of systems to their parts (at least in statistical mechanics) raises fairly obvious difficulties for the kind of literal reductivist approach outlined by Friedman. Even if we disregard the problem of correlating temperature and mean kinetic energy across theoretical boundaries, a more significant difficulty arises in the case of identifying the constituents of the systems 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 this 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 (p. 241). But the motivation for Friedman's account is to achieve a literal interpretation of theoretical structure, which in turn yields greater confirmation of hypotheses; something he sees as guaranteed by the model-submodel approach.

If we recall what constraints are involved in the relationship between a model and its submodel we see that they are structurally similar insofar as the interpretation of each relation, function and constant symbol in the submodel $\mathcal{B}$ is the restriction of the corresponding interpretation in the model $\mathcal{A}$. Equivalently for every atomic formula φ and assignment [s] in $\mathcal{B}$, $\mathcal{B} \models \varphi$ [s] iff $\mathcal{A} \models \varphi$ [s]. Applied to our physical example we see that a <u>literal</u> identification of the properties of individuals of the thermodynamic system ($\mathcal{B}$) cannot be accomplished given the structural constraints on ($\mathcal{A}$), 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 its 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 <u>identification</u> 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 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 so-called "reduced" entities/theory are correlated with the reducing theory or model. Given the logical properties of the model-submodel relationship we should expect that the relations and functions specified by the identity map would be preserved in the way we think of inference rules as truth preserving. A further benefit of the representational approach is that it allows the relationship between $\mathcal{B}$ and $\mathcal{A}$ to change over time, something that is _prima facie_ ruled out by a _literal identification_ of their corresponding elements.

A possible objection to my argument against the reductivist approach is that I have addressed only the problem of theory reduction rather than the simple reduction of observable entities to their theoretical counterparts. Because theory reduction presents a series of complicated issues it isn't surprising that we are unable to provide a nicely circumscribed model to illustrate its features. Nevertheless, one might claim that reduction is still a viable, if not preferable, way to describe the relationship between the observable and theoretical properties/entities of our theories. Although this issue was partially addressed in the discussion of statistical mechanics in the following section I will try to show in more detail the problems associated with entity reduction.

In the discussion of statistical mechanics I suggested that there may be several models of the theory each specifying different features that would be applicable in a particular context. The idea that the theory has many models nicely solves the apparent incompatibility of identifying individual particles in some contexts and not in others. The demand for different models to account for the same phenomena also arises in the more narrowly defined context of entity reduction; a situation that poses obvious problems for the model-submodel approach and the accompanying idea that we can correlate the elements in each model by means of an identity map.

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 and go on to derive 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 the so-called derivation of this law and its relationship to the kinetic theory is more complicated than Friedman would have us think. In order to properly expose these complexities I will begin with some historical remarks on the actual formulation of van der Waals law. From there I go on to point out that in different contexts the solutions to the problems addressed by this law indicate the need for more than one molecular model.

As I mentioned above, the problem of deriving the law from theoretical structure does not arise on an account like van Fraassen's. However, Friedman's argument addresses a somewhat more complicated issue. Even if we acknowledge a literal interpretation of theoretical structure it doesn't follow that the model-submodel approach can be vindicated. But, in order to show this we must show that the relationship between the phenomenological and theoretical structures is characterized by a "looseness of fit" that is inappropriate to the model-submodel account.

A. The Derivation of van der Waals Law - Historical Details.

Van der Waals law was originally formulated as a response to the idealizing assumptions of the Boyle-Charles law which maintained that the size of molecules and the forces between them were negligible.[13] Although the molecules of real gases are assumed to have a finite size and exert an intermolecular force at ordinary temperatures and pressures, real gases behave very much like ideal ones and thus in some situations obey the Boyle-Charles law. However, at sufficiently high and low temperatures real gases may 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 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 approached each other in collisions. In other words, because the nature of the substance was defined by its molecular model, properties of the model should be present under all temperatures and pressures.

Van der Waals came upon this idea of continuity of the liquid and gaseous phase as a result of Clausius' work on the virial theorem, which was an attempt to reduce the second law of thermodynamics to a purely mechanical form.[14] The theorem connected the sum of the time average vis viva of all the molecules to the forces acting on those molecules

$$\sum \overline{m/2\, c^2} = -1/2 \sum \overline{X_x + Y_y + Z_z} \qquad (4)$$

where $(m/2)c^2$ was the vis viva of the molecules and X, Y and Z the components of the forces acting on the molecule at x, y and z. The object was to trace the actual motions of the molecules that constituted heat and to show that the effective force of heat is 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 central forces $\varphi(r_{ij})$ acting between particles, the theorem will take the following form:

$$\left\langle \sum \tfrac{1}{2} m_i v_i^2 \right\rangle = \tfrac{3}{2} PV + \left\langle \tfrac{1}{2} \sum_{ij(i>j)} r_{ij}\, \varphi(r_{ij}) \right\rangle \qquad (5)$$

with r_i denoting the coordinates of the ith particle whose mass and volume are m_i and u_i respectively and $r_{ij} = |r_i - r_j|$. 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.[13]

Although the virial theorem incorporated the possibility 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 <u>assumptions</u> about the mean free path[14] he allowed for the fact that the molecules had finite size[15]. 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. This followed from continuity considerations which implied that the attractive forces were very short range.[16] 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

$$< \sum 1/2\, mu^2 > = 3/2\, RT \qquad (6)$$

an assumption that could <u>only</u> be argued for on the basis of plausibility consideration. The equation in its final form was.

$$(P + a/v^2)\ (v-b) = RT. \qquad (7)$$

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 which the molecules exert on one another when reasonably near to one another. The correction b arises from forces which the molecules exert on one another when their centers 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[14] on isotherms for carbon dioxide.

B. Philosophical Concerns. On the basis of our historical digression we can see that van der Waals' methods for deriving his equation of state departed from the kinetic principles illustrated by the virial theorem and as such was unsatisfactory as a deduction from the kinetic theory.[20] It is interesting to note that although 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. 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. Although this discovery was made in 1924 its physical significance was not appreciated until almost ten years later.[21] Despite its experimental corroboration, when applied to cases of greater than first order deviation from Boyle's law the molecular model suggested by van der Waals' approach was seen to be insufficient.[22] 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 these molecules exert a negative pressure, an assumption that implied negative values for P. Because an examination of physical conditions shows that the true value for P must be positive an alternative formulation and molecular model was proposed by Dieterici.[23] This model assumed the constant temperature of the gas so that the total energy distribution applied to molecules striking the wall as well. as those that don't. Although both equations imply the existence of what is termed a critical point, a point where the liquid, gaseous and vapour states meet,[24] they make different predictions as to the existence of this point; with the Dieterici equation appearing more accurate for heavier and more complex gases. Generally speaking however neither one comes particularly near to actual observations of critical data.[25] The reason for the discrepancy is that both equations are true only when deviations from Boyle's law is small with the critical point representing a rather large deviation.[26]

Various attempts have been made to improve the van der Waals equation by the introduction of more adjustable constants to supplement a and b; constants which can be chosen so as to make the equation 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 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 behavior of gases. 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). Not only is there a problem in deriving the van der Waals equation from the assumptions of the kinetic theory but an explanation of the behavior of real gases (something the van der Waals law is designed to explain) requires many <u>different</u> laws and <u>incompatible</u> models. In general the van der Waals equation tends to smooth out the differences between individual substances and predicts that they behave more uniformly than they do.[*] In fact very accurate experiments have shown quantitative discrepancies from the results predicted by the van der Waals equation. In calculations of the difference in density between 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. . In reality this difference varies nearly as the cube root, a result which 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 behavior of gases at the phenomenological level?

If the relationship between the behavior of gases and their molecular model is one of model to submodel then the same relations and properties that hold in the latter must hold in the former (with the submodel 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 while the Deterici 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 van der Waals law can be achieved using a particular model which we know to be inapplicable in other contexts. Hence it appears that the uniqueness of the mapping in the model-submodel account is actually a drawback rather than an advantage.

The fact that the observable behavior of gases requires more than one model for its explanation again seems to favour the embedding approach over the submodel account as a way of understanding physical theory. As I 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 semantical issue of whether our theory is true or false remains an important consideration. The skeptical problem becomes an epistemological concern regarding our ability to assign a particular truth value to the theory in question. Whether we have sufficient evidence or justification to claim that our theory is true is a separate concern and not required for a literal interpretation of the theory's assumptions.[22]

There is however another line of defense the realist could use to vindicate the position. One could claim in this case that the van der Waals example showed that a single general molecular theory which did not incorporate any 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 state of matter supplied a kind of ontological unity that formed the theoretical basis for van der Waals equation. One could easily 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 this account could be motivated then one could claim that despite the changes in properties the core of the model remains unchanged. Such a view would clearly go some way toward vindicating Friedman's model-submodel 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, regardless of the fact that 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 law 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 (that is, 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 these refinements were not applicable when the motions of molecules were slow, a situation that arises at low temperatures when relativistic theory merges into the classical one. 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 and their properties; properties which supposedly give rise to many of the empirical phenomena we are concerned to explain. Moreover, it is exactly these properties that figure importantly in the derivations of phenomenological laws from the higher level theoretical structure. Postulating a molecular model 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 <u>true</u> description of theoretical structure because such an interpretation provides 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. Because the embedding approach allows for a variety of models of the phenomena it seems closer to actual scientific practice. Friedman's model-submodel account is not only too restrictive in its requirement that the model be literally true but the relationship between the phenomenological and theoretical structure cannot be accurately depicted by the kind of stringent logical requirements that hold between models and their submodels.

The disadvantage of the embedding approach is that the kind of epistemological skepticism characteristic of the representationalist strategy cannot sanction the kind of conjunctive inference Friedman sees as necessary for theory evolution. However, the application of the conjunction rule requires that we interpret our theories as literally true, a challenge that the reductivist programme cannot meet on the basis of its structural constraints and logical requirements. That is, although the programme claims to provide a literal reduction of one class of entities to another, the difficulties involved in achieving such a reduction renders it a failure by its own standards. As we saw in the discussion of statistical mechanics, a literal interpretation of the mechanisms involved (a decomposition of the system into components) contradicts one of the fundamental principles of the model of a statistical system. At best it presents an idealized, non-literal account of theory structure and evolution. In light of these difficulties the role of conjunction as a purely logical principle becomes questionable. If our theories are not literally reducible in the way the model-submodel approach suggests then the use of conjunction must be motivated on the basis of methodological considerations.

The discussion in II. outlined the relationship between the logical and methodological aspects of the conjunction rule. On the one hand if we claim that our theories are literally true then we can allow for conjunctive inference; a logical rule that scientists accept. The other aspect (the methodological one) states that, as a matter of fact, scientists not only accept this logical principle but use it in developing their theories. As an abstract logical principle it presupposes the truth of each conjunct (theory), something that the reductivist model cannot accommodate and something that is not required for a literal interpretation of theoretical structure. In the remainder of this chapter I shall argue that conjunction is relatively unimportant as a methodological principle. I begin with a look at some instances of conjunctive inference and how it operates in scientific contexts. If we distinguish between conjunction and theoretical unification or consilience it becomes clear that the former plays virtually no role in the latter; moreover, many cases of theoretical unification provide counter-examples to the kind of strict reductivist programme characterized by Friedman.

4. Conjunction. One of the important advantages of theory conjunction concerns our ability to predict and explain

phenomena that could not be accounted for on the basis of either conjunct taken separately. The realist model suggests that if we believe T_1 and T_2 to be true then we can a priori conjoin them and thereby acquire an increase in overall information and predictive power. The issue is whether this model presents an idealized version of the way theories actually evolve. We know that the truth of the conjunction can only be guaranteed if we know that each conjunct is true; but, unfortunately, to assume this knowledge is to simply assert what we want to conclude. The process of conjunction consists of bringing together two hitherto successful theories in the hopes that their conjunction will be successful. The minimum requirement for this operation is that the two theories are consistent; but this is simply a point about the logico-mathematical structure of the theory, knowledge that the anti-realist has as well. Although the anti-realist (by denying truth) has no guarantee that the conjunction will be successful, neither does the realist. Epistemic attitudes fail to guarantee the success of inferential practices. We initially conjoin our theories with an eye to further testing and successful prediction and only if the conjunction survives empirical tests and yields accurate predictions will it be believed or accepted. Our expectations for theory conjunction are based on predictive success, but to equate predictive success with truth yields the kind of instrumentalist theory that realists typically object to because it cannot account for the theory conjunction.

The realist might want to claim that successful conjunction provides evidence for or is an indicator of truth; or that only well established theories (ones that are more likely to be true) ought to be conjoined. The crucial feature of this type of argument is that it necessitates that some explanation of the connection between truth and success be given. This leaves realism in a rather unfortunate position because, as the history of science shows us, truth cannot explain the overwhelming success of many false theories. However, if we subscribe to a theory of truth that appeals to epistemic properties we cannot validate the conjunction rule. It would seem then that despite the criticisms of anti-realism on the basis of conjunction the realist is in no better position to anticipate the outcome of theoretical conjunctions than the anti-realist.

One instance where conjunction is used and does seem necessary is in the context of the relationship between theoretical hypotheses and auxilliary assumptions. In order to successfully explain or predict a phenomenon we often require a theory together with laws and assumptions taken from other domains. Many realists [cf. Boyd 1985a; 1985b] want to claim that this process requires the backdrop of realism in order to be successfully explained. We need to presuppose the truth of a "unity of science" principle according to which a variety of well-confirmed theories may be legitimately employed conjointly in making observational predictions. The unity of science principle itself presupposes some sort of univocality of theoretical entities and structure occurring in the theories in question. This kind of unity, claims Boyd, is required if the D-N model of explanation is to retain any degree of plausibility. Hoowever, the D-N model of explanation (as well as the H-D method of inference) tell us nothing about the truth of the explanans; rather, it tells us that when certain hypotheses are conjoined with other theoretical assumptions or laws we can derive the explanandum. Although we assume our auxilliaries to be well-confirmed, in the process they are employed as something akin to <u>ceteris paribus</u> clauses. We choose the relevant factors that we think will enable us to achieve successful predictions and manipulate the variables until we arrive at the desired outcome. But this context differs in important respects from the motivations inherent in cases of theoretical conjunction. In the latter case we simply assume that by conjoining true hypotheses with other true hypotheses we will get the explanation or theory that we desire, and although the same principles are presupposed (unity of science and univocality of theoretical entities and structure) we expect the conjunction to yield true predictions prior to experimental testing. It typically involves projections rather than legitimation through a posteriori justification.

Especially relevant to the discussion of conjunction is the issue of correction. In a review of van Fraassen [1980]

Demopolous [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.[29] It seems however, that by introducing the issue of correction the logical and methodological aspects of the conjunction objection have changed bringing them more into line with the role played by auxiliaries. The initial claim advanced on behalf of realism was that one could expect, independently of other considerations, that if T_1 and T_2 were true their conjunction would be true [Putnam 1963]. So, the point of the conjunction 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 couldn't expect true predictions what could possibly be the purpose of conjoining theories?[30] 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 been previously 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 a logical operation encompassing semantical and epistemological considerations.

The question that remains concerns the role of conjunction in theory unification or consilience, a process that involves much more than simply bringing together two successful hypotheses. If we want to achieve the kind of unification that Friedman alludes to then it is important to recognize that in cases of this sort the theoretical changes that take place are often so great as to render simple conjunction inapplicable. If this is in fact the case then the role of conjunction and Friedman's claim that theories evolve by conjunction [p.245] is typically of no importance for scientific methodology. To test my hypothesis I will begin with a few preliminary remarks about unification and how it relates to Whewell's notion of consilience.[31] It may be that unification can be seen as a justification for realism regardless of whether it proceeds according to the model outlines by Friedman.

<u>5. Consilience and Unification.</u> We know that Friedman's model of theory evolution consists of a reduction or identification of phenomenological properties with their literally interpreted theoretical counterparts which, when conjoined with other pieces of theoretical structure, form

the basis for a unified theory that can successfully predict a variety of phenomena. We also know that the justifiction for the literal interpretation of theoretical structure is motivated by the need for conjunctive inference in the context of theory evolution and unification. Consilience differs from this account insofar as the unified theory is not built up as a result of a conjunctive process. So, although the unification that one achieves in a consilience of inductions involves the reduction of disparate phenomena, I argue that the kind of reductivist model suggested by Friedman is, to some degree, incidental. Moreover, because consilience involves not only a correction of the laws involved in the theories but also a change or reinterpretation of the terms, the model-submodel approach presents too restrictive a programme to properly explicate the requisite kind of unification.

Traditional accounts of consilience have emphasized its relationship to what we commonly call unification. A consilience of inductions is said to occur when an 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.[34] 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 which were thought to be distinct to one general kind or theory. In addition, this unification results in the reduction of the amount of theoretical structure required to account for the phenomena. Although the deductive entailment content of consilient theories is very high this virtue is often achieved only at the cost 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.69-70].

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 juxtaposition 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 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 terrestial 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. Consider the following example: Neither Galileo's law of falling bodies (a body dropped from a tower falls a distance s in time t where s = f(t)) nor Kepler's second law (the radius vector from the sun to earth sweeps out equal distances in equal times) can be deduced from the inverse square law (any two masses in the universe attract each other with a force proportional to the product of the masses as the inverse square of the distances between them). In addition we require the assumptions that the earth and sun are large masses with the earth being smaller and that all bodies near the surface of the earth are relatively small masses. Given these conditions we then reinterpret Galileo's and Kepler's second law (respectively) in the following fashion:

> GL - The very small mass starting from rest in the neighbourhood of a large mass and relatively far from other large masses moves toward the large mass so as to describe distance s in time t where s = f(t).

> KL - The radius vector from one mass to another where these are relatively far from other large masses, sweeps out equal areas in equal times.

These newly formulated laws are of course derivable from the inverse square law together with the aid of the calculus and some rules of mechanics. It is important to point out that in some cases the new law, as reinterpreted by the inverse square law, actually contradicts its original counterpart. For example, if we think of Galileo's law of falling bodies as claiming that bodies fall to the surface of the earth with constant acceleration, then even though Newton's theory explains and unifies this law with the laws of celestial motion it nonetheless predicts something completely different; namely, the fact that accelerations increase as bodies approach the center of the earth. What Newton's theory explains is why Galileo's empirical law seemed to be correct or as in the case of Kepler's laws why they were correct to a certain degree of approximation. As incorrect laws neither is, strictly speaking, explained within the context of universal gravitation.

The traditional realist approach, and one that was certainly favoured by Whewell, is to attribute 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. The unification of theories may, in some instances, constitute part of what is involved in consilience and the accumulation of independent evidence. A theory that unifies a group of diverse phenomena is seen as having a greater degree of confirmation because of its ability to explain/predict a number of phenomena that were once thought to be distinct. So, it is important to keep in mind that the explanatory and predictive power applicable in cases of consilience must encompass phenomena from different domains.

The fact that a consilient theory amasses a great deal of independent evidence is important as a measure of the inductive support enjoyed by the theory. A consilience is similar to the testimony of two witnesses on behalf of the hypotheses:

> ...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. [<u>PIS</u> Vol.II p.285]

When two different classes of facts lead to the same hypothesis we may assume that we have discovered a vera causa. Consilience then can be regarded as having little to do with the structural aspects of the theory or its explanatory power simpliciter; instead it is seen as having a special kind of explanatory power which can be characterized as a measure of inductive support. In a truly consilient theory the hypothesis is put forth <u>not</u> simply because it is explanatory but rather because of the variety of independent evidence that supports it. Hence the domains of explanatory power and consilience as theoretical virtues need not overlap completely. Explanatory power as a structural virtue of theories may have little, if anything, to do with the amount of independent evidence that the theory has amassed.

There is however a notion of simplicity that is particularly relevant to consilience in the sense defined above. In virtue of the fact that a theory is consilient it will by definition require fewer theoretical entities or structures to account for a group of phenomena than a number of theories taken separately or in conjunction. Whewell himself saw consilience of inductions as giving rise "to a constant convergence of our theory towards simplicity and unity" [Butts, 1968 p.159]. Although he considered consilience and simplicity to be "hardly different" and exemplified by the same cases they were not "identical in their essence". Consilience was a contributing factor in the search for a simple, unified theory.

As I noted above, simplicity and explanatory power need not guarantee an accompanying degree of consilience. For instance, we may prefer a theory on the basis of its mathematical simplicity or because of the simplicity of its assumptions. The latter corresponds most closely to what I previously referred to as a high measure of deductive content. In these kinds of cases a law like the gravitation law of general relativity is considered simpler than Newton's law of gravitation because it can account for Mercury's perihelion, the red shift phenomenon and the bending of starlight, without the addition of any auxiliary assumptions. Mathematical simplicity, on the other hand, concerns the ease with which one can obtain the implications of a particular law. Newton's law of gravitation is mathematically simpler than that of general relativity, enabling us to easily derive the solution to a two-body problem within the classical theory; something that is not the case in relativity. On Whewell's account simplicity and unity were the culmination of a convergence of a number of consilient inductions. Structural simplicity may not necessarily contribute to this convergence.

In his discussion of the relationship between simplicity and inductive support Rosenkrantz [1981][33] introduces the term "sample coverage" to denote the proportion of experimental outcomes that a theory fits; or, more generally, the probability that a theory will fit a specified (or unspecified) outcome when a suitable hypothesis of independence obtains. Consequently when a very simple theory agrees with observations the probability of agreement that good occurring by chance is extremely small. Notice that this notion of sample coverage is distinct from what is normally involved in the case of simple theories and their relationship to conjunction of hypotheses. Unlike the account of simplicity and unification I have discussed above, theory conjunction serves, in most cases, to complicate our theory by introducing additional hypotheses. By contrast a simple theory is one that had few adjustible parameters. Although conjunction increases the logical content and applicability of the theory it is achieved only at the cost of increasing the number of parameters (hence making the theory less probable). Unfortunately this notion of theory conjunction plays a central role in Friedman's programme. He repeatedly emphasizes the importance of confirmation that results from the ability to predict and explain a variety of phenomena; but this confirmation results, on his account, from the fact that our theories evolve by conjunction. Rather than simply seeing unifying power as a criterion for a realist interpretation of theoretical structure, Friedman's case involves a more complicated methodology. The persistence or stability of particular structures through time enables our theories to evolve and increase their confirmation value (p.245). This approach is at odds with most cases of consilience/unification where there are significant changes made to the laws and terms of the hypotheses involved. 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 escapes many of the traditional problems associated with theory and entity reduction because it involves the unification of two groups of phenomena simply on the basis of independent evidence. In contrast, many reductions involve the identification of different levels of phenomena and theory for the purposes of theoretical simplicity or explanatory efficacy. In recognizing the need for correction and alterations to the hypotheses and laws (instead of simply "reducing" one group to another without taking account of their differences) the unification/consilience process frequently involves a conceptual reshuffling of the phenomena. Consequently, much of Friedman's discussion of confirmation and conjunction together with his motivation for a literal construal of theoretical structure is rendered irrelevant in the context of a truly consilient theory.

In discussing consilient theories it is important to note the contextual factors that are associated with the process. A theory becomes consilient when it shows that phenomena originally thought to be of a different kind 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 terrestial 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 action of similar types of vortices. This kind of historical relativity and the idea that a consilient theory is somehow a stamp of truth is thus in need of some further clarification.

This issue of historical relativism is perhaps the most important difficulty for a realist account of confirmation (one that equates consilient theories with reasons for true belief). Because the very notion of independent evidence together with the claim that 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 (unless of course truth is construed relativistically). On Friedman's account specific ontological claims are legitimated on the basis of unifying power. Because these entities/structures perform a unifying role in some contexts and not in others, our beliefs become dictated solely on the basis of currently accepted theory and the historical contingencies involved in the unifying process.[34] 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. A strategy of this kind simply rejects what is right about realism; namely, the search for "theory neutral" facts that can act as arbitrers in the evaluation of our theories and the accompanying epistemic and ontological attitudes that are crucial for understanding scientific methodology. This notion of independence or neutrality seems sacrificed on an account that uses unification (a context dependent process) 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. Instead our committments and beliefs become doubly abstracted; they are not simply relativized to a particular epistemic community but to the domain of a particular theory. Because a realism based on criteria like explanatory unification provides no ontological stability over time it would seem that as a justification for belief it proves unsatisfactory even for many realists. It is interesting to note that one of Friedman's justifications for a literal construal of theoretical structure is that it allows for a persistence of that structure over time, something which his own account, due to its historical variability, cannot guarantee independently of a particular context.

Finally, can we incorporate the account of consilience sketched above into a viable confirmation theory? Over and above the preference for straightforwardly simple theories that explain, the accuracy of our consilient theories is harder to ascribe to mere coincidence; they are what can be called "improbably accurate" [Rosenkrantz 1981]. Rosencrantz ~~suggests that~~ these simple theories are to be valued because they are more confirmable by "conforming data"; they have what he terms "higher cognitive growth potential". Rosencrantz provides historical justification for his argument by an analysis of the superiority of Copernican astronomy over the Ptolemaic theory. Very briefly the thrust of his argument is that the simplicity of the heliocentric view allows for an overdetermination of the phenomena, a criterion that produces distinct evidential value. We can account for all planetary motions in terms of a common cause, the earth's motion. On the view Rosencrantz provides the simplicity of the theory involves restricting its form to the extent that when the theory accounts for some very precise aspect of the phenomena it enjoys a substantial increase in confirmation. In other words the actual verification of the phenomena, something that was highly improbable given the restrictions or simplicity of the theory, is seen as 'miraculous' as the verification of a genuinely novel prediction. There is however an important difference between this type of simplicity and the kind of simplicity that results from a convergence of consilient hypotheses. In the latter case we amass a large variety of evidence from a number of different sources. This provides independent evidence that a group of phenomena are either of the same kind or that they share a common causal source, a conclusion that serves to insure a degree of overdetermination. The Rosencrantz account achieves overdetermination by restricting the form of the theory, so rather than fitting a wide variety of data it may fit only a few. But, because the theory has specified its parameters so closely the accuracy with which it fits the data makes it highly improbable that the theory is wrong. Although both accounts result in some degree of overdetermination Rosenkrantz's version lacks the strength that a broad range of independent data can bring to bear on a theory. He sees his account as providing a legitimate reason to abandon a well-entrenched theoretical approach for one that, in its infancy, is a good deal less accurate albeit simpler.[36] In light of our discussion of consilience (simplicity) and independent evidence the motivation for abandoning well-entrenched theories for less accurate ones is not altogether clear. Our preference for simple theories isn't to be explained by the form of the theory but rather because it has received a large measure of inductive support from disparate realms.

Historically we have accepted and should continue to accept the theories that provide the best supporting evidence. In many cases these have been consilient theories; but, the important question, the one that is the core of the realism debate, is whether we are justified in believing that consilient theories are true. If we are to take the notion of truth seriously, as something over and above warrented assertability, instrumental progress, etc., then it seems obvious that accepting and working within the context of a particular theory need not and indeed should not be accompanied by the belief that the theory is true. If fallibilism is to be a crucial aspect of scientific practice (and philosophy) then our epistemology and methodology ought to reflect that attitude. Realism and its emphasis on true belief fails, I think, in that regard.

The independent evidence accumulated by a consilient theory serves to reinforce arguments from coincidence or, similarily, inference to the most probable cause or a particular unifying structure. All are forms of inference to the best explanation. But, we accept arguments to the best explanation, the most probable cause or unifying structure because they provide the best explanation as to why consilient (as well as unifying) theories have been successful. However, as anti-realists are fond of pointing out, success of these theories cannot be equated with nor is it indicative of their truth. The point is perhaps best exemplified in the different attitudes taken by Whewell and Descartes toward arguments from coincidence. Whewell thinks it impossible that a theory which results in the convergence of two trains of induction could be in error. To illustrate his point he appeals to the metaphor of deecyphering a code in a unknown language, a metaphor used by Descartes in Principia:

> If I copy a long series of letters of which the last half dozen are concealed and if I guess those aright, as is found to be the case when they are afterward uncovered, this must be because I have made out the import of the inscription....We amy compare such occurrences [i.e. consiliences] to a case of interpreting an unknown character, in which two different inscriptions, deciphered by different persons, had given the same alphabet. We should, in such a case believe with great confidence that the alphabet was the true one.[On the Philosophy of Discovery p.274]

Whewell considers this scenario to be indicative of "a stamp of truth beyond the power of ingenuity to counterfeit" [Ibid.]. However, mindful of historical considerations a perhaps more reasonable approach is the one taken by Descartes in Principia Art. 204:

> That it suffices if I have explained what perceptible things may be like, even if perhaps they are not so.

Our scientific theories should be considered, at best, to be a description of what the world <u>may</u> be like, but this concerns possibilities rather than actuality and truth.[30]

<u>6. Conclusions and Summary.</u> What I have tried to argue is not that belief in theoretical structure is incorrect but only that the reasons provided as justification for that belief are inadequate to the task.

Friedman's programme suggests a limited brand of realism, restricting belief to the aspects of theoretical structure that participate in the unification of our theories. The difficulties that beset his approach are numerous. He suggests that the phenomenological aspects of a theory should be literally identified with the theoretical structure that explains their behavior. This identification gives rise to a literal interpretation of theoretical structure that enables us to perform the conjunctive inferences that are necessary for theory evolution. Not only does this approach <u>presuppose</u> that theories evolve by conjunction (something that, in the case of truly consilient theories, is clearly not the case) but it requires for its motivation the kind of strong realism that it was designed to counteract. If, as Friedman's model suggests, theories evolve and become unified by a conjunctive process, then we must believe that our hypotheses are literally true in order to apply the conjunction rule. So, rather than believing only in that part of the theory that unifies various phenomena, we must believe in the truth of the prior assumptions that were employed in achieving this unification.

In addition, this kind of conjunctive approach cannot account for the substantial changes in laws and terms that are often involved in the process of unification/consilience. Friedman sees his reductivist method as one which provides a stability of theoretical structure over time; the kind of stability that allows theories to evolve. However, in the context of a truly consilient theory, one that unifies a variety of phenomena, this kind of stability is simply not an issue. Instead, a new conception of the phenomena is introduced leaving us with a radically different picture than the previous theory provided. Because of these considerations it becomes difficult, if not impossible, to motivate a need for the conjunction rule on methodological grounds. Hence, the claim that it is a logical principle necessary to scientific methodology presents a somewhat idealized and abstract view of scientific practice. If one of the demands of a philosophy of science is that it explain practice then the embedding approach outlined above seems preferable.

Contrary to Friedman's demand, the reductivist programme in general has been notoriously unsuccessful in enabling us to provide a literal account of theoretical structure. The idea that the phenomenological and theoretical levels can be related to each other as submodel to model imposes a set of logical constraints that are often too stringent to account for the looseness of fit that exists between these two levels of phenomena. It is frequently the case that several different and sometimes incompatible models are necessary to account for the phenomena encompassed by a theory. Different representations may be employed for different purposes; for instance, the billiard ball model is used for deriving the perfect gas law, the weakly attracting rigid sphere model for deriving van der Waals equation and the model representing molecules as point centers of inverse power repulsion for facilitating transport calculations. If phenomenological properties are literally identified with theoretical structure then in what sense does the identity map function if theoretical structure must be depicted in a variety of ways. A literal interpretation in this context becomes a belief in a "thing or entity whose properties cannot be determined in any strict sense; a kind of "thing we know not what". This form of realism proves unsatisfactory because it is the specific properties of theoretical structure that allows us to derive phenomenological laws and make predictions. Without these properties much of the motivation for realism becomes unnecessary. Because the embedding approach provides a mechanism for incorporating the requirement for different models of the same structure and for changes in the models of the theory over time it provides a representation that is closer to scientific practice.

The final and perhaps most serious difficulty with

Friedman's account of realism concerns the variability of our ontological commitments. Because the justification for belief in theoretical structure supposedly arises only in the contexts of unifying entities or principles, and because unification is always relativized to a particular historical context, our beliefs become dictated <u>solely</u> on the ability of certain structures to provide unified explanations in the context of a <u>particular</u> theory. In other words, ontology becomes completely contingent on a methodological criterion without the advantage of a theory-neutral basis from which to assess these beliefs. Entities and structure come and go the way unifying contexts come and go; as a result we give up the kind of stability or persistence over time that Friedman seems at pains to preserve. One is reminded in this case of the Cheshire cat:

> "I wish you wouldn't keep appearing and vanishing so suddenly" replied Alice, "you make one quite giddy". "All right" said the cat; and this time it vanished quite slowly beginning with the end of the tail and ending with a grin which remained some time after the rest of it had gone. "Well! I've often seen a cat without a grin" thought Alice; "but a grin without a cat! It's the most curious thing I ever saw in my life!".

Footnotes

1. Putnam's argument is on somewhat firmer ground than Friedman insofar as he does not limit belief to the unifying structure. Friedman's selective brand of realism seems to require strong realist assumptions in order to achieve the kind of unification he endorses, despite his claims that only the unifying structure is to be interpreted literally.

2. Although Friedman is correct in seeing constructive empiricism as a generalization of van Fraassen's earlier views on space and time he is mistaken in equating the position with traditional forms of instrumentalism. One obvious reason for Friedman's characterization is perhaps van Fraassen's use of the embedding rather than the submodel approach, a move consistent with various froms of anti-realism. Although van Fraassen's literal interpretation renders our theories capable of being true or false it doesn't entail the corresponding requirement that we believe them or that we can assign them a truth value. We present a theory by specifying its models and delineating certain parts of those models (the empirical substructures) as candidates for the direct representation of observable phenomena. Once our theory is accepted (minimally, what is actual and observable finds a place in some nodel of the theory) it guides our linguistic practice. As a result the language receives its interpretation through the model(s) of the theory. Modal locutions as well as statements about thelretical structure and unobservable entities reflect the fact that our models specify many possible courses of events. van Fraassen sees the literal interpretation as specifying the model as the locus of possibility, not a reality behing the model. [1980 p.220] Because language is interpreted through the model rather than by some mysterious hookup with reality, the constructive empiricist can advocate a literal interpretation of theories while remaining agnostic about metaphysical committment to theoretical structure.

3. The details of this case will be discussed at leangth in chapter 3.

4. See The Scientific Papers of James Clerk Maxwell. Vol. 1 p.500.

5. For an interesting discussion of these issues see the Nagel-Feyerabend exchange in the literature; especially Feyerabend's "Explanation Reductin and Empiricism" reptinted in <u>Problems of Empiricism</u> Vol. 2, Feyerabend (ed.) Cambridge [1982]. See also Ernest Nagel, <u>The Structure of Science</u> and "Issues in the Logic of Reductive Explanation" in Kiefer and Munitz (eds.) <u>Mind, Science and History</u> SUNY Press [1970].

6. Cf. Feyerabend, op.cit., Nagel, op. cit. and Thomas Nickles, "Two Concepts of Intertheoretic Reduction" <u>Journal of Philosophy</u> Vol. LXX April, 1973.

7. Feyerabend uses this argument to make a case for a strong form of the incommensurability thesis. That is not my intention here; rather, I use the example only as an illustration of the difficulties involved in literal identifications within the reductionist programme.

8. One of the central features of thermodynamics is that the entropy of an isolated system increases with time until it reaches a maximum at equilibrium. The empirical foundation of this fact is the impossibility of a perpetum mobile or the arrow of time (entropy grows with time in an irreversible way) as indicated by heat conduction. Mechanical phenomena, on the other hand, are all time reversible so it appeared that any mechanical model, like the kinetic theory would be incompatible with the second law of thermodynamics. The kinetic theory teaches us that a gas system will pass from the less probable to the most probable microstate, the state of statistical equilibrium. Hence, it follows from this that the state of maximum entropy is the most probable state.

In order to avoid absurdities the kinetic theory required the abandonment of the absolute validity of the law of entropy. Maxwell, Boltzman and Gibbs proposed the idea that the principle of entropy expressed probabilities and hence had only approximate validity. Occasionally it was possible that a less probable state may arise and even when statistical equilibrium is reached there exists the possibility that some of the subsequent states will be among the highly improbable ones. A specific fromulation of this condition became known as the recurrence paradox (Poincare 1890). It stated that any mechanical system constrained to move in a finite volume with fixed total energy must eventually return to the neighbourhood of any specified initial condition. If a certain value of entropy is associated with every configuartion of the system (a disputable assumption) then instead of continually increasing with time the entropy must eventually decrease in order to return to its initial state. In order to aviod the apparent incompatibility with the H theorem (H being a magnitude associated with collisions of molecules, a magnitude that would decrease until it attained a minimum H_m with $H-H_m$ constituting the departure of the macroscopic state of a gas from statistical equilibrium) Boltzman maintained that the equilibrium state is not a single configuation but a collection of the majority of possible configurations. If one waits long enough some particular initial state (a fluxuation) is almost certain to occur; however, the probability of such a fluxuation is so small that one would have to wait an immensely long time before observing a recurrence of the initial state. Because the probability is so slight we can apply the thermodynamic principle of entropy without fear of error. The possibility of improbable states arising becomes much more favourable with small numbers of molecules are considered (as in the case of Brownian motion and fluxuations). Due to the nature of H it can be seen as proportional to (-) the thermodynamic entropy.

We see then that the kinetic theory leads to an extension of the thermodynamical concept of entropy. In the latter context differences in entropy were determined in connection with reversible transformations and the concept acquired meaning only when the gas was in the state of statistical equilibrium. When interpreted within the boundries of the kinetic theory entropy serves as a measure of the randomness with which the molecules are moving - the more orderly the motion the smaller the entropy. Entropy S was was defined as proportional to the probability of a system, i.e. $S = k \log W$ where W is the number of molecular microstates corresponding to a state of a macroscopic system as defined by thermodynamic variables like temperature and pressure. From a random sample of microstates the probability of finding a given microstate is proportional to its W. If W is large the system is disordered.

9. For a more specific account of this case see I. Khinchin, The Mathematical Foundations of Statistical Mechanics.

10. In statistical mechanics we describe the state of the 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:

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

where H is the so-called Hamiltonian function of the 2s variables $q_1, \ldots p_s$.

11. See Khinchin, Ch. 2, op. cit. for details.

12. In other words, the function $H(q_i; p_k)$ is an integral of the system described by the equations of motion.

13. For a description of the derivation taking account of the sizes and forces of the molecules see D. Tabor, Gases, Liquids and Solids, Chapter 5.

14. For Clausius' original paper on this topic see Philosophical Magazine, 1870. pp.122-27. Additional discussion of the historical origins of van der Waals' law can be found in Elizabeth Garber, "Molecular Science in Britian" HSPS Vol.1; M.J. Klein, "Historical Origins of van der Waals Equation" Physica, 73 pp.28-47 and Peter Clarke, "Thermodynamics vs. Atomism" in C. Howson (ed.) Method and Appraisal in the Physical Sciences Cambridge, 1976.

15. From equation (5) it was possible to see that the pressure of a gas could arise from the motion of particles or from possible repulsive forces between them. Because experimental evidence (Laplace's theory of capilliarity and surface tension of liquids) had already indicated that forces acted between the molecules of a gas, van der Waals assumed that these forces would affect the pressure term in the ideal gas law. The size of the correction could be computed using the viral theorem. Since the form of $\Phi(r_{ij})$ (the viral of the intermolecular forces) was unknown $\langle 1/2 \sum r_{ij} \Phi(r_{ij}) \rangle$ could be replaced by an effective intermolecular pressure term P' giving us

$$\sum 1/2\ mu^2 = 3/2\ (P + P')V.$$

For gases the molecular pressure P' is expected to be only a correction to the external pressure P whereas for liquids the external pressure (in normal circumstances) is likely to be negligible compared to P'.

16. See Klein, op.cit. pp.39ff. for details of van der Waals method.

17. Since the mean free path is supposedly reduced for spherical molecules (by a factor of λ/λ_0 where λ_0 is the mean free path for point molecules) the pressure must be increased by reciprocal of this ratio (pressure being proportional to number of collisions per unit time). The equation of state could now be written in the form

$$\langle \sum 1/2\ mu^2 \rangle = 3/2(P+P')(v-b)$$

where $b = 4N_0v_m$, the force times the actual volume of the N_0 molecules in a mole. In short (V - b) is simply the volume of the container minus the sum of molecular volumes.

18. The number of interacting pairs was proportional to the square of the density of the fluid or inversely proportional to v^2 where v is the volume per mole. By introducing a proportionality factor a the internal pressure a/v^2 giving us

$$\langle \sum 1/2\ mu^2 \rangle = 3/2\ (P + a/v^2)\ (v - b).$$

19. Andrews original paper reporting the results of his experiments on CO_2 was published in the Philosophical Transactions of the Royal Society, 159, pp.575-90. See also Clarke, op. cit., Tabor, op. cit. and S. Brush, The Kind of Motion We Call Heat New York: North Holland [1976] Bk. 1 Ch. VII.

20. For a brief discussion of Maxwell's dissatisfaction with the relationship of van der Waals equation to the kinetic theory see his Collected Scientific Papers, op. cit. Vol. II, pp.407-08. For a detailed quantivative treatment of the relationship between the virial and the van der Waals equation see E. Segre, From Falling Bodies to Radio Waves, Appendix 14, pp.281-83.

21. See S. Brush [1983] Statistical Physics and the Atomic Theory of Matter Princeton: Princeton U. Press

22. Cf. Tabor, op. cit. and Sir J.Jeans The Kinetic Theory of Gases, especially Chapter 3.

23. Ibid.

24. The critical temperature T_c specifies the point above which liquification is impossible while P_c and V_c are the pressure and volume at which liquification first begins, the point when the substance is at a temperature just below T_c. So long as a gas is kept above critical temperature no pressure, however great, can liquify it. A gas below critical temperature is usually

described as a vapour. Cf. Jeans as well as Tabor.

25. Cf. Jeans p.94 for a table of comparisons.

26. It is interesting to note that, the van der Waals explanation of critical temperature is not tied to the particular approximations used in deriving his equation. The existence of a critical temperature is usually assumed to depend upon the minimum in the potential energy curve but as yet there is no satisfactory detailed theory based on the molecular forces. For details see C.H. Collie Kinetic Theory and Entropy Chapter 2.

27. That this happens can be seen from the "reduced equation of state. If a and b are eliminated from the equation in terms of critical constants one obtains this "reduced" equation which gives reduced values for pressure (where reduced p $=p/p_c$), temperature and volume. The equation is supposedly the same for all gases since the quatitiies a and b, which vary from gas to gas have disappeared. If this equation could be regarded as absolutely true then whenever any two "reduced" quantities are known the third could also be given and, similarily, when any two quantities are the same fo tow gases then the third will also be the same. (The Law of Corresponding States) This law is true only if the nature of the gas can be specified by two physical constants. But, as in the case of van der Waals law, the law of corresponding states is true only as a first approximation. For details see Jeans, op. cit., Collie, op. cit., and S. Brush, The Kind of Motion we call Heat, Book 1, Chapter 7.

28. This is the issue raised by Van Fraassen in the Scientific Image.

29. In an attempt to counter the Putnam-Boyd objections to anti-realism van Fraassen argued that theory evolution could not proceed by simple conjunction; there was usually if not always substantial corrections made to the individual theories that were brought together.

30. Richard Boyd suggests this is his "Lex Orandi est Lex Credendi" in Churchland and Hooker, Images of Science, 1985.

31. Friedman cites his own model as an instance of Whewell's consilience of inductions. However, it is important to note that although consilience involves a reductive process it does not proceed by conjunction in the way Friedman's model does.

32. See Laudan [1971] "William Whewell's Consilience of Inductions" The Monist Vol. 55; Butts [1968] William Whewell's Theory of Sicentific Method Pittsburgh: U. of Pittsburgh Press, [1973] "Whewell's Logic of Induction" in Giere and Westfall (eds.) Foundations of Scientific Method: The Nineteenth Century Bloomington: Indiana U. Press, [1977] "Consilience of Inductions and the Problem of Conceptual Change" in Colodny (ed.) Logic, Laws and Life Pittsburgh: U. of Pittsburgh Press; Hesse [1968],

"Consilience of Inductions" in Lakatos (ed.) The Problem of Inductive Logic Amsterdam: North Holland; M. Fisch [1985] "Whewell's Consilience of Inductions" Philosophy of Science Vol. 52, No.2.

33. The variety of evidence that increases the confirmation relations on Friedman's account seems to result from a particular hypothesis being conjoined with another hypothesis or bit of theoretical structure. It has little, if anything, to do with the fact that the hypothesis itself is supported by a large amount of independent evidence as is characteristic of consilient theories.

34. Friedman himself points out (p.249) 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 have been interpreted literally. We now know that classical electrodynamics is false and absolute rest is not needed for formulating an accurate law of motion, hence we can assign it a purely representative status

35. See his discussion of Copernicus in his Foundations and Applications of Inductive Probability [1981]..

36. It is interesting to note the way in which Whewell interpreted the argument from coincidence as evidence for truth. Together with his famous claim that the history of science offers no examples of a consilient theory that proved to be false is his somewhat Humean psychological account of belief and necessary truth. As an hypothesis is subjected to more and more increasingly difficult trials we form the conviction that "no other law than those proposed can account for the known facts" This conviction "finds its place in the mind gradually, as the contemplation of the consequences of the law and the various relations of the facts become steady and familiar. Just as in Hume's case, where the repeated impression of a constant conjunction leads the mind to form an idea of a necessary connection between cause and effect, so too is it in Whewell's account. The accretion of independent evidence or different inductive proofs indicating the same conclusion persuades us that we have discovered the true cause or a necessary truth. Our belief becomes so strong that we cannot conceive it possible to doubt the truth of our hypothesis. See The Philosophy of the Inductive Sciences, Vol. II, p. 268 & 622.

Chapter II

A Study in Theory Unification: The Case of Maxwell's Electromagnetic Theory

1. Introduction. In chapter 1 I argued that Friedman's reductivist programme and reductivist accounts in general fail to capture the mechanisms required to model most cases of theory unification and consilience. The semantical changes together with the corrections to the laws of the theories involved cannot be easily incorporated into an account that emphasizes the importance of strict logical derivations and relationships between different levels of phenomena. In fact, as I tried to show, a reductivist approach often precludes interpreting theoretical structure in a literal fashion. This problem was apparent in the discussion of theory reduction (thermodynamics to statistical mechanics) and entity reduction (the need for many different molecular models in the explanation of the behavior of real gases). In addition the changes made to Kepler's and Galileo's laws on the basis of the inverse square law illustrate just how reduction and unification involve processes much more complicated than the model-submodel account can accommodate.

Despite these difficulties there have been cases of reduction and unification that were relatively straightforward; instances that provide what might seem to be evidence for a view like Friedman's. One such example is

Maxwell's electromagnetic theory of light. Unlike the case of universal gravitation there were relatively few corrections to the laws of either electrical or optical theory. It was possible to show that the basic wave equation for light was deducible from Maxwell's equations describing the electric and magnetic field. There were some difficulties in the relationship between the preceeding theory of physical optics and the electromagnetic theory of light. Although some aspects of these theories were not logically compatible they nevertheless did bear a very close relationship to each other; a relationship that would appear to at least approximate Friedman's model of theory reduction and unification. Regardless of the difficulties presented by physical optics in deriving suitable laws of reflection and refraction it is important to stress the relative ease with which light could be identified with electromagnetic waves. It was this unification that resulted in the identification of the luminiferous and electromagnetic aether. Not only did it take place in a relatively straightforward manner but, in addition, the unification provided by the electromagnetic theory satisfied the requirements of a Whewellian consilience by displaying the appropriate degree of independence with respect to the phenomena that were unified. So, although the luminiferous aether was "seen in a new light" its identification with the electromagnetic aether involved no substantial changes to either theory and thus seems to provide a good example of the kind of reduction and unifying power that Friedman describes in his account of theoretical inference.

Despite this straightforward reduction the issue becomes substantially more complicated when we consider unification as an argument for realism about theoretical structures (in this case the aether and the displacement current). Not only was there considerable difficulty in specifying a workable model of the aether but the introduction of a "displacement current" necessary for the derivation of the electromagnitic theory of light was highly contentious due to its seemingly ad hoc postulation. Although it was intended as a counterpart to Faraday's electrotonic state, the displacement current took on a substantially different form within the domain of Maxwell's theory. Maxwell grew increasingly dissatisfied with the attempts to devise an appropriate mechanical model for his theory and resorted to formulating his account of elecgromagnetism strictly within the context of the Lagrangian formalism. Thus the electromagnetic theory could be considered a dynamical mathematical theory that was independent of any specific mechanical model of the aether or qualitative account of displacement. In light of these difficulties I argue that, contrary to Friedman, a successful unification need not support or be an ingredient in a realistic construal of the unifying physical structure. In other words, unification is not sufficient to sanction a realism about the physical interpretation of the mathematical model of a theory.[1]

In discussing the development of the electromagnetic theory it is particularly interesting to note the role played by models and analogies in Maxwell's thought. In chapter one discussed the advantages of the embedding approach and an accompanying anti-realism about theoretical models. Because the articulation of a particular theory often requires a variety of models the idea that we can specify one as a literally true description of reality presents an overly abstract view of the structure of scientific theories. This approach encorporating many different models is especially prominent throughout Maxwell's work. As I shall try to show the electromagnetic theory provides a nice historical account of the way in which the embedding approach meshes nicely with scientific practice. The mathematical unification acheived in the electromagnetic theory carried no accompanying commitment to any particular physical model. Although Maxwell relied heavily on dynamical models and analogies to enhance the understanding of the theory he remained agnostic as to the literal truth of these models as actual representations of the material world. Models for Maxwell, as for his predecessors Faraday and Thomsom (Lord Kelvin), were not depictions of how nature <u>actually</u> behaved but rather, like analogies, they provided us with <u>various different</u> conceptions of the phenomena. Analogies were seen as an important heuristic tool for theory construction and as such figured importantly in the discovery process rather than as a method of justification. Although they functioned in much the same way as models, providing possible interpretations of the phenomena, their status was sometimes limited to purely fictional representations. What Maxwell calls the "method of physical analogy" is simply a mechanism for developing models of the phenomena that will serve to illustrate physical ideas that may be regarded as possible interpretations of the mathematical formalism. Although Maxwell was severely criticized (especially by Duhem) for putting forth inconsistent models, given his views on their role in theory construction this can hardly be seen as a legitimate criticism of his method.[a]

I begin the historical discussion with an account of Maxwell's debt to Faraday and Thomson; the two most important figures from whom Maxwell's views on models, analogies and the electromagnetic theory itself can be traced. From there I examine the specific details of the unification of electromagnetics and optics as well as the problems that arose in the attempt to establish a realistic view of theoretical structure. I conclude with a discussion of the relationship between mathematical and physical aspects of theories and models and the implications of this distinction for a view like Friedman's.

2. Debt to Faraday and Thomson.

2.1 The Origin of "Lines of Force". In 1831 Faraday discovered that a magnet could induce an electrical current in a wire when the wire was made to cut the magnetic lines of force. This discovery of electromagnetic induction (the induction of an electric current between two coils of wire wound around an iron ring) was explained by a "peculiar condition" in the magnet or ring.[3] This "electrical condition of matter" was termed the "electrotonic state" and was said to be the result of a state of tension in the particles of the ring. It was the creation and dissolution of this electrotonic state that caused the induction of the current.[4] Faraday defined the state of tension, or polarity, as a state in which a molecule acquires opposite powers on different parts. Hence the electrotonic state was thought to involve the polarization of molecules that resulted from an electric force, where polarization was construed as opposite electrical states on different parts.

Electrochemical decomposition was also explained by the transmission of forces by the particles of matter. In his first paper on electrostatic induction in 1837[5] Faraday showed that the force between two charged bodies depends on the insulating medium that surrounds them and not merely on their shape and position. (Cavendish had proven the same result long prior to Faraday but his work remained unpublished.) According to Faraday's theory, induction was an action caused by contiguous particles and took place along curved lines which he called "lines of force". Faraday showed experimentally that different substances had different capacities for the mediation of electrostatic forces.[5] These lines of force were said to denote the disposition of individual particles to transmit an electric force from particle to particle.[6] The particles of the medium that transmitted this force (the dielectric) were subject to a state of electrical tension that led to the propogation of electrostatic forces. Thus, the electrotonic state came to be identified with the action of the particles and the lines of force.

Faraday used the term "lines of force" as a "temporary conventional mode" of expressing the direction of the power[7]. The idea behind these lines was that they were a kind of geometrical representation of the lines of polarized particles of the dielectric that were subjected to electric tension. The polarization was not considered to be the result of mutual contact; instead the particles acted by their associated polar forces. Although Faraday denied that electrostatic induction took place by action at a distance he was at a loss as to how to explain the mode of transmission of the tension from one polarized particle to another.[8] Faraday offered a preliminary solution to this difficulty by maintaining that the intermediate spaces between atoms could account for the communication of electric action. However, if this space was considered as an ingredient in the continuous transmission of force then one is led to the absurd conclusion that space may be a non-conductor in non-conducting bodies and a conductor in conducting bodies. Faraday saw only one way to overcome this difficulty. By arguing that there exists a system of powers and forces around the atomic centers of matter he concluded that substance itself should be seen to consist of these powers; as a result matter was thought to be continuous throughout all of space. Atoms could thus be thought to penetrate to the very centers of force; they were highly elastic and mutually penetrable and perhaps most importantly, electric action could be explained without presupposing that particles act across insensible distances.

In this new theory of matter, which was in stark contrast to Faraday's original particulate theory, the lines of force played an extremely important role (See "Thoughts on X-Rays Vibrations" [1846]). Electrical action could now be <u>explained</u> solely by lines of force as opposed to the polarization of particles by the interaction of forces. Faraday even went so far as to suggest that light could be understood as the result of vibrations of the lines of force instead of regarding it as being carried by a quasi-material aether. Although he recognized this hypothesis as mere speculation the germ of the idea was retained, albeit in a different form, in Maxwell's development of an electromagnetic theory of light.

The notion of lines of force received a further strengthening in its application to magnetic phenomena. The fundamental importance of lines of force as entities distinct from the particles of matter can be seen in Faraday's explanation of paramagnetic and diamagnetic phenomena in terms of the relative magnetic conductivity of the bodies and the surrounding medium. Prior to the investigations of diamagnetism the direction of the magnetic lines of force was invoked to explain the rotational effects of magnets on light and the effect of magnetism on the alignment of crystals. In the former case a magnetic field rotates the plane of polarization of light passing through certain transparent substances. However, the direction of rotation depends only on the direction of the magnetic lines of force and not on the direction of the light ray; the power of the rotating polarized light being directly proportional to the intensity of the magnetic force. Faraday recognized that the polarity was associated with the entire line of magnetic force and the light ray rather than the particles of the dielectric. Because he had not dissociated lines of force from matter he called this new magnetic force or mode of action of matter "diamagnetism".[10] However, in a series of experiments on diamagnetism Faraday failed to detect any polarity in these so-called diamagnetic substances[11] causing him to abandon the idea that molecules of diamagnetics were polarized. Instead he concluded that the diamagnetics merely interacted with lines of force; the latter now conceived as entities distinct from particles of matter. The relative conductability of the paramagnetic and diamagnetic bodies was then explained in terms of the propensity of lines of force to pass through the bodies.[12] In fact, magnetic conductivity - the differences in the magnetic character of all substances - was explained by the propensity of lines of force to pass through each substance. Polarity no longer existed as a state of matter, instead it represented the direction of the lines of force in the force fields.[13]

In a series of experiments using magnets, copper discs and wire Faraday showed that lines of force were general strains that did not move with the magnet and were produced by the magnet as a whole rather than by specific parts; that is, they did not end on the poles of the magnet.[14] In addition, he concluded that the lines of force within the magnet were exactly equal in amount to the ones outside the magnet. Unlike his original account there was no polarization of molecules conceived as centers of action; instead the lines of force were seen as continuous curves. They were thought to be in the medium, because the magnet itself could not exist without a surrounding medium or space and, moreover, forces could only be related to each other by these curved lines through the surrounding medium. Although lines of force seemed to be given a kind of physical meaning in Faraday's 1852 "On the Physical Character of the Lines of Force", their exact nature remained a mystery:

> The term line of magnetic force is intended to express simply the direction of the force in any given place, and not any physical idea or notion of the manner in which the force may be there exerted; as by actions at a distance, or pulsations, or waves or a current or what not. [ER,3, pp. 328ff,; 368ff.; 402]

In addition the nature of the medium or space, as well as the essence of the magnet itself, remained a puzzle.

We have seen how the properties of electric induction led Faraday to the view that this is not an action at sensible or insensible distances but one that requires the intervention of a material medium. Similarily for the case of magnetic lines of force; indeed they had many properties in common with electric induction such as their curvature and the fact that current is induced in a circuit by mere motion in a magnetic field.[15] But, here again the existence of these lines of force was highly questionable. As Faraday noted toward the end of his Experimental Researches:

> If they [magnetic lines of force] exist, it is not by a succession of particles, as in the case of static electric induction...but by the condition of space free from such material particles. A magnet placed in the middle of the best vacuum we can produce... acts as well upon a needle as if were surrounded by air, water or glass; and therefore these lines exist in such a vaccum as well as where there is matter [3,para.3258].

What he had characterized was a model for the lines of force: a set of possible constraints that he saw as providing a plausible account of their nature. But, even as a model Faraday's account was relatively ambiguous, perhaps because he wished to emphasize the fact that his ideas about lines of force should not be understood as a true physical theory. Nevertheless, it was often difficult to unravel Faraday's conception of these lines. If they were strains, but not strains of material particles, then how were they to be explained? Faraday rejected the then current model of the aether as a possible explanation on the grounds that one could not detect a polarity of that aether. (Because the aether was assumed to be particulate the lines of magnetic force would have to be seen as strains of these particles. It followed from this that there ought to be an aetheral polarity analagous to the electrostatic lines of force.)[10] Despite the denial of the particulate aether Faraday claimed that in order for the magnet to exist it required a surrounding medium. However, he remained content simply to maintain that these lines of force were lines of strain. Although Faraday did remark that "...the surrounding magnetic medium, deprived of all material substance may be, I cannot tell, perhaps the aether" [3, p.425], it is important to note that any aether postulated as an explanation of magnetic phenomena would of course have to be in keeping with Faraday's theory of a continuous force plenum. As an answer to Hare's objections [cf., note 8] about action at a distance Faraday had replaced the molecular aether with a prevading medium of forces with continuous lines of force replacing contiguous particles as the cause of the activities in the medium. The particulate aether had assumed successive centers of action while the new theory presupposed lines of action. It was these lines of force between centers of particles that carried the vibrations required by magnetic phenomena [ER, pp.451-52]. Hence, the dualism of matter and force necessitated by traditional aether models could be replaced by an all-pervading but continuously differentiated force.

Although this theory of force provided the continuous medium required for the explanation and existence of magnetism the nature of these lines of force and hence the nature of the aether itself remained a mystery. Faraday continued to equate the lines with the idea of a strain, analagous to a bundle of elastic strings stretched under tension. Faraday also used the term "conducting power" but it was considered no more than a convenient way of saying that iron concentrates lines of force and it did not imply any physical hypothesis (or real analogy) about magnetism travelling along lines of force in the way charges travelled along lines of current flow [ER, 3, p.397]. The relation between material bodies and these spatial strains was illustrated by using an analogy that compared the magnet to a voltaic battery immersed in water, or any other electrolyte.[17] Upon removal of the elctrolyte the voltaic cell became an inert container filled with chemicals. Only when the external medium permitted passage of electricity did the cell become a center of electrical force. The same was true of the magnet:

> I incline to consider this outer medium as essential to the magnet; that it is that which relates the external polarities to each other by curved lines of power, and that these must be so related as a matter of necessity. Just as in the case of the battery...there is mo line of force either in or out of the battery, if this relation be cut off by removing or intercepting the conducting medium. [ER, 3, para. 3276]

This idea of a force-field pervading empty space constituted what came to be known as Faraday's field theory[18] and although he remained agnostic as to the precise nature of the magnetic lines of force and the medium they comprised, both were fundamental in the formulation of an alternative conception to action at a distance. It was this notion of an alternative conception of the phenomena that became an important consideration in Maxwell's work, especially with respect to the role of models and analogies in the development and explication of scientific theory. As will be evident in the work of Thomson and Maxwell, ideas like that of "lines of force" can assist in the development of theoretical models and analogies regardless of whether they can be verified to have actual counterparts in reality.

2. Thomson's Contributions. In 1841 Thomson published a mathematical demonstration that the formulas of electricity deduced from laws of action at a distance are identical with the formulas of heat distribution deduced from the ideas of action beween contiguous particles. (An analogy drawn from Fourier's theory of heat.)[19] Again in Janurary 1845 Thomson, using Green's notion of a potential function, demonstrated that Coulomb's electrical action at a distance and Faraday's notion of action by contiguous particles in the medium led to the same mathematical theory.[20] Having undertaken a study of Faraday's electrostatics Thomson extended the mathematical analogy between thermal and electrostatic phenomena to the suggestion of a corresponding physical analogy. The physical model of propagation of heat from particle to particle suggested an analagous propagation of electrical forces by the action of contiguous particles (molecular vorticies) of an intervening medium. Although Faraday would later transform this analogy into his field conception of lines of force Thomson denied any justification, based on the analogy, for its introduction as a true physical hypothesis. Instead this mathematical correspondence merely indicated that a possible model of propagation of electrostatic forces could be based on a theory of action between contiguous particles.[21]

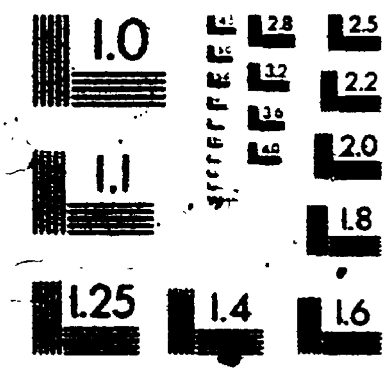
Perhaps the most important mathematical analogy developed by Thomson, at least in so far as Faraday's final version of field theory is concerned is the one outlined in his 1847 paper "Mechanical Representation of Electric, Magnetic and Galvanic Forces".** As we saw above the foremost difficulty with Faraday's early version of the field theory was its inability to explain the mechanism of molecular interactions between the contiguous particles of the dielectric medium. Thomson's attempt at a solution was contained in what he termed a "sketch of a mathematical analogy" that dealt with the propagation of electric and magnetic forces in terms of the linear and rotational strain of an elastic solid. Thomson used the mathematical methods developed by Stokes to treat rotations and strains in a continuous aether. He showed that the distributions of the linear and rotational strains in an elastic solid in space would be analogous to the distributions of electric force (for point charges) and magnetic force (for magnetic dipoles and electric currents). Although this mechanical model had physical implications (Faraday's discovery of the rotational effect of magnetism on light was consistent with the representation of magnetic force by the rotational strain of an elastic solid)** the <u>relationship</u> between mechanical strain and electrical and magnetic phenomena remained physically unclear. However, as Thomson noted, if a physical theory was to be discovered on the basis of this speculative anaolgy it would, when taken in connection with the undulatory theory of light, most probably explain the effect of magnetism on polarized light and hence exhibit a connection with the luminiferous aether.[84]

In 1849 Thomson turned his attention to the mathematical representation of Faraday's theory of the magnetic field as a structure comprised of lines of force. In a paper entitled "A Mathematical Theory of Magnetism [1849] & [1850][85] Thomson observed that it was "convenient" to conceive of magnetic force as due to a continuous distribution of attractions and repulsions mutually exerted between portions of an "imaginary matter". This analogy enabled him to once again use the potential function[86] in defining the resultant magnetic force at any point in space, whether or not the point was occupied by magnetized matter. The imaginary matter of this analogy possessed none of the qualities of ordinary matter and hence it could not be identified with a solid or the magnetic fluid. The physical picture suggested by this analogy resulted in Thomson's having little difficulty in accepting the notion of a magnetic force in empty space.[87] This magnetic matter was considered an embodiment of Faraday's ideas about the primacy of lines of force, a plenum that represented the spatial distribution of the field of force. In fact, Thomson concluded that it was this magnetic substance that determined the total force in the magnet, a claim that suggested that the magnetic action lay in the imaginary matter and not in ordinary matter. Indeed it was these views of Thomson's that strongly influenced Faraday's concept of the field and the accompanying notion that lines of force could exist in what seemed to be an otherwise empty space.

Thomson's 1849-50 paper is particularly interesting for its emphasis on the gap between mathematical and physical representations of phenomena. In the opening paragraph of the paper Thomson notes that we must consider the investigation of the mathematical and not the physical laws regulating magnetic action as primary.[88] He was successful in showing that the hypothesis of two magnetic fluids and of action-at a distance is not necessary for a mathematical theory of magnetism. The representations of the lines of magnetic force by a magnetic plenum were not intended as models of the physical constitution of the field; instead they were merely meant to provide physical illustrations or analogies for the corresponding mathematical accounts.

In 1856 Thomson attempted to provide what he called a "dynamical illustration" of the physical field.[89] Although his discussion of the field involved the notion of molecular motions he carefully avoided commitment to any specific molecular model. The dynamical theory of heat provided the model for the physical field but no specific details accompanied Thomson's account over and above the very general hypothesis that heat was a form of molecular motion.[90] The physical structure of the aether whose motions constituted the field was not specified; instead several different possible accounts were offered. The aether could be seen to

2



be a continous fluid permeating the spaces between molecules of gross bodies or it could be constituted of discrete molecules, or alternatively, it was possible that all matter was ultimately continuous with its apparent molecular structure produced by the vortical motions of a continuous aether.

While Thomson went on to develop a theory of a continuous aether based on a physical embodiment of Faraday's force plenum Maxwell, in 1856, embarked on an exploration of the geometrical implication of Faraday's notion of lines of force. Maxwell saw the concept of "lines of force" as providing a purely geometrical representation of the structure of the field. His interest had been sparked by Thomson's early paper on the geometrical analogy between thermal and electrical phenomena. And in keeping with the ideas of Faraday and Thomson, Maxwell intended to emphasize the gap between mathematical and physical representations by providing a framework of "geometrical truths" that could be applied to both thermal and electrical phenomena. In what follows I discuss how this mathematical approach to physical theory influenced Maxwell's work, especially his ideas about the aether and its role in the development of a unified electromagnetic theory. In keeping with some of the general themes addressed in chapter one, especially the motivation for a realism about unifying theoretical structure, particular emphasis will be placed on the status and role of the physical implication of a mathematical representation of nature; as well as the differences between mathematical and physical unification. Maxwell's work, seen as an extension and development of the views of Faraday and Thomson, raises a number of interesting issues that provide insight into contemporary debates about theory unification and realism.

3. Development of the Electromagnetic Theory - The Early Stages.

3.1 Maxwell circa 1856 - "On Faraday's Lines of Force". The goal of Maxwell's 1856 paper was to present Faraday's account of the electromagnetic theory in mathematically precise yet pictorial form. Having been strongly influenced by Thomson's presentation of his mathematical and physical ideas in an analogical way, particularly his use of molecular vortices as a representation of the field, Maxwell employed the method of models and analogies as a means of illustrating Faraday's conception of lines of force. He felt that Faraday's "lines" could be employed as a purely geometrical representation of the structure of the field with the direction of forces acting in the field represented by lines of force filling space. In order to employ these lines of force as a quantitative expression of the forces themselves Maxwell conceived of them as tubes carrying incompressible fluid made up of a "a collection of imaginary properties".[11] This geometrical model defined the motion of the fluid by dividing the space it occupied into tubes with forces being represented by the motion of the fluid. Seizing on the formal equivalence of the equations of heat flow and action at a distance Maxwell sought to substitute the flow of the ideal fluid for the notion of action at a distance. The pressure in the tubes containing the fluid varied inversely as the distance from the source of the fluid but the crucial feature of the model was that the action was not at a distance; the energy of the system was in the tubes of force. Its origin was Thomson's representation of the analogy between heat flow and electrostatic action.[32] In keeping with Thomson's mitigated skepticism about the status of models Maxwell remained cautious, warning that "the two subjects will assume very different aspects" if their resemblance is pushed too far.[33] Nevertheless, the mathematical resemblance of the laws would remain something that could be considered useful in the development of further mathematical ideas.[34] The importance of Maxwell's analogy was that it had shown that one could look at electrical and magnetic action from two different yet mathematically equivalent points of view.[35]

Maxwell went on to provide an illustration of the phenomena of electrostatics, current electricity and magnetism by drawing analogies between them, and the motions of the incompressible fluid. At no time however did Maxwell intend his account to be anything more than an analogy to be used for heuristic purposes. In fact the incompressible fluid was not even considered as a hypothetical entity. Consequently the geometrical model was not put forth as a physical hypothesis, instead the emphasis was on mathematical rather than physical similarity; a resemblance between mathematical relations rather than phenomena or things related.[36] The advantage of the method of analogy over a purely analytical formalism was the visual representation provided by the lines, surfaces and tubes. Although Maxwell referred to his method as one that presents <u>physical</u> analogies he did so because he considered it a method of obtaining physical <u>ideas</u> without adopting a physical <u>theory</u>.

> By a physical analogy I mean that partial similarity between the laws of one science and those of another which makes each of them illustrate the other. Thus all the mathematical sciences are founded on relations between physical laws and laws of numbers, so that the aim of exact science is to reduce the problem of nature to the determination of quantities by operations with numbers [<u>Papers</u>, 1, p.156].

The kind of emphasis Maxwell places on analogy is indicative of his committment to a strong distinction between mathematical and physical theory. In an early draft on the 1856 paper[37] Maxwell remarked that he had assumed a purely imaginary fluid because "While the mathematical laws of the conduction of heat derived from the idea of heat as a substance are admitted to be true, the theory of heat has been so modified that we can no longer apply to it the idea of a substance".[38] That the purpose of the analogies was purely heuristic is evident from other more specific remarks.

> By referring everything to the purely geometrical idea of the motion of an imaginary fluid, I hope to attain generality and precision, and to avoid the dangers arising from a premature theory professing to explain the cause of the phenomena. If the results of mere speculation which I have collected are found to be of any use to experimental philosophers in <u>arranging</u> and <u>interpreting</u> their results, they will have served their purpose, and a mature theory, in which physical facts will be physically explained, will be formed by those who, by interrogating nature herself, can obtain the only true solution of the questions which the mathematical theory suggests [Papers, 1, p.159].[39] [emphasis added]

In essence then the analogy served as a means for applying specific techniques to a variety of fields while at the same time carving out a middle ground between mere mathematical abstraction and full blown committment to a physical hypothesis. The dangers of adopting a particular hypothesis consist of seeing the phenomena only through a medium; we become "liable to that blindness to facts and rashness in assumptions which a partial explanation encourages".[40] The method of analogy provides a means of investigation that facilitates a clear physical conception of the phenomena without being carried beyond the truth due to committment to a particular hypothesis.[41] Instead of providing a literal interpretation of possible states of affairs analogies furnish guidelines for constructing and developing theories as well as suggesting possible approaches for experimentation. They supply the mechanisms for what has been traditionally termed the "context of discovery". In contrast to the analogical method a purely mathematical approach causes us to lose sight of the phenomena to be explained and its lack of heuristic power prevents possible extensions of the "views on the connections of the subject".[42]

The second part of "On Faraday's Lines of Force" consisted of a mathematical formulation of Faraday's concept of the "electrotonic state". If we recall from the discussion above this concept was used to represent the electrical tension of matter that was associated with Faraday's early theory of particulate polarization. It was the electrotonic state that eventually gave way to the idea of lines of force. Maxwell introduced a mathematical expression (later called vector potential) that enabled one to conceive of the electrotonic state at any point in space as a quantity determinate in magnitude and direction and represented by any mechanical system that has at every point a quantity like velocity, displacement or force whose direction and magnitude correspond to "those of the supposed electrotonic state".[43] Again Maxwell was quick to point out that his representation involved no physical theory, it was what he called a kind of artificial notation. However, he used the electrotonic state because it reduced the attraction of currents and electrified bodies to one principle without introducing any new assumptions.[44] In addition, the electrotonic state enabled Maxwell to avoid considering the quantity representing magnetic induction (the number of lines of force) passing through the circuit; instead he could use the more natural method of considering the current with reference to quantities existing in the same space as the current.[45] Although the nature of the electrotonic state was sometimes represented by a variety of physical entities its essence was its mathematical form. As Maxwell pointed out:

> The idea of the electrotonic state has not yet presented itself to my mind in such a form that its nature and properties can be clearly explained without reference to mere symbols:...By a careful study of the laws of elastic solids and the motions of viscous fluids I hope to discover a method of forming a mathematical conception of this electrotonic state adapted to general reasoning [Papers, 1, p.188].

In keeping with his cautious attitude toward physical hypotheses Maxwell saw the emphasis on mathematical formulations as providing a distinct advantage over accounts that presuppose the truth of specific background theoretical assumptions. For example, in the development of Weber's electromagnetic theory it was assumed that the attraction or repulsion of moving electrical particles upon one another was dependent on their velocities; an assumption that was itself rather problematic. As Maxwell pointed out, "if the forces in nature are to be reduced to forces acting between particles, the principle of conservation of force requires that these forces should be in the lines joining the particles and functions of the distance only" [Ibid. p.208]. This suggested that Ampere's theory of two separate electrical fluids, upon which Weber's account had been based, needed to be abandoned.[46] The mathematical account provided by Maxwell suggested an alternative conception insofar as it provided the formal apparatus necessary for the beginnings of a different and promising approach.

Unfortunately people were reluctant to pursue Maxwell's programme for several reasons. Not only was there no physical/mechanical basis for the analogy but the flow analogy itself failed to provide an image of the coexistence and interaction of electrical fields, magnetic fields, and electric currents. The flow lines of the imcompressible fluid were taken to correspond to electric or magnetic lines of force, or lines of electric current, depending on the context of the problem at hand. As a result the analogy provided a fragmented understanding of each of the three electrogmanetic phenomenon by considering each one in isolation from the others.[47] Thomson's theory of molecular vortices offered a possible solution to this problem by promising to provide the kind of comprehensive explanatory power that the flow analogy lacked. And, unlike the imaginary fluid of the flow analogy, which was simply an illustrative device, an account of the connections and interactions of the various phenomena was to be considered as a genuine physical <u>possibility</u>. It was an approach that could at least be <u>understood</u> as having a physical basis rather than simply instrumental value. But, although the explanatory power of the vortex hypothesis was an important consideration it was <u>not</u> taken to provide any evidence for the claim that electromagnetic phenomena were <u>actually caused</u> by vortical motions in the medium. As Maxwell remarked:

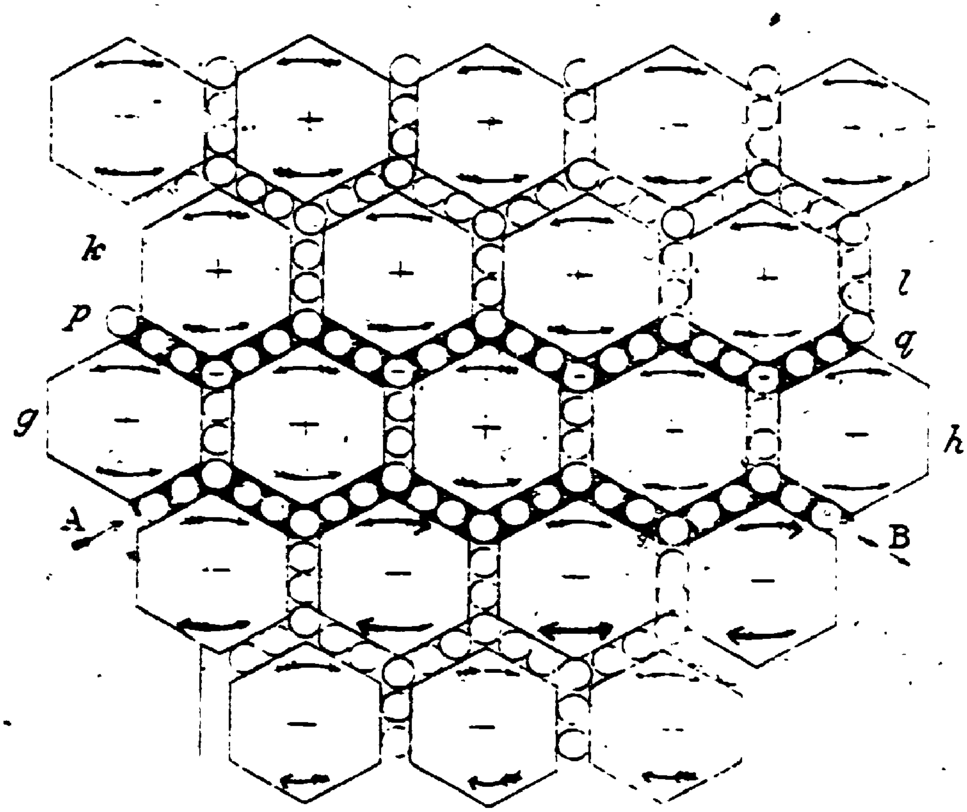
> If, by the molecular-vortex hypothesis we can connect the phenomena of magnetic attraction with electromagnetic phenomena, and with those of induced currents, we shall have found a theory which, if not true, can only be proved to be erroneous by experiments which will greatly enlarge our knowledge of this part of physics.[41]

Using Thomson's molecular vortex model Maxwell went on to develop an account that took him beyond the flow analogy to a mechanical model of electromagnetic phenomena.

3.2 "On Physical Lines of Force". Faraday's distinction between the geometrical treatment of lines of force (a descriptive account of their distribution in space) and their physical treatment, (dealing with their dynamical tendancies) was carried over quite distinctively into Maxwell's work. In "Faraday's Lines" (hereafter FL) there was no mechanism for understanding the forces of attraction and repulsion between magnetic poles. In keeping with the method of Thomson's 1847 paper Maxwell had used mechanical analogies or <u>illustrations</u> to assist the imagination, rather than to account for the phenomena. In contrast the "Physical Lines" (PL) paper was an attempt to "examine" the phenomena from a mechanical point of view and to determine what tensions in, or motions of, a medium were <u>capable</u> of producing the mechanical phenomena observed [PL, p.467]. The physical behavior of the magnetic lines was what was needed in order to account for magnetic forces.[49]

In Thomson's 1856 paper Faraday's discovery of the rotation of the plane of polarization of light by magnets was explained by a theory that construed magnetism as the rotation of molecular vortices in a fluid aether. This effect led Maxwell to the view that in a magnetic field the medium (or aether) was in rotation around the lines of force; the rotation being performed by molecular vortices having their axes parallel to the lines of force. The fact that magnetic poles attracted or repelled each other led Maxwell (as well as Faraday) to the view that the lines of force extending from these poles represented a state of tension in the medium. In addition the concept of polarity was modified to include the vortices, with the polarity of the lines of force being represented by the polarity of the vorticies that constituted them. So, by quantitatively evaluating the stress tensor in the medium Maxwell could account for all classes of magnetic forces. He then went on to derive a general expression for the force F on a unit of volume of the electromagnetic medium subject to the vortex motion.

Part II of PL offered an attempt to specify the forces that caused the medium to move in the way described by the model and to account for the occurence of electric currents. In order to do this Maxwell needed to provide an explanation of the transmission of rotation in the same direction from vortex to vortex. Because of the difficulties involved in specifying a mechanical model of vortex rotation Maxwell was forced to introduce a layer of spherical particles residing on the surfaces of the vortices and separating them from each other. The problem was roughly the following: because neighbouring lines of force in a magnetic field point in approximately the same direction, adjacent vortices in the electromagnetic medium must rotate in the same way. Hence, as Maxwell pointed out, the contiguous portions of consecutive vortices must be moving in opposite directions. By introducing a layer of particles dividing the medium into cells one could account for the motion of the vortices required by Maxwell's model by means of the rotation of the material cells. The particles separating the cells could then be seen to rotate in an opposite direction to the vorticies they separated. Each particle revolved on its own axis and in the opposite direction to the neighbouring vorticies. So, by postulating this layer of particles one could easily account for the rotation of vortices about parallel axes in the same direction. By placing the particles <u>between</u> contiguous vortices the rotation of each vortex would cause the neighbouring vorticies to revolve in the same direction.



These particles came to be known as idle wheels that rolled without slipping on the surfaces of the vortices. If the adjacent vorticies were not revolving at the same rate the idle wheel particles would acquire a translatory motion, with electric current being represented as the transferrence or flow of the moveable particles. In other words, the motion of the particles from vortex to vortex through the medium constituted a current. The tangential action (resulting from their motion) of the particles on the cells in the electromagnetic medium caused the cells to rotate and so accounted for magnetic fields due to currents and electromotive force. If the angular velocity of each vortex remained constant the particles separating the vortices experienced no net force pushing them in any speicifc direction. However, when a change in the state of rotation of the vortices was transmitted through the medium, a force did result, thereby enabling the model to account for electromagnetic induction.

The basic difference between the position taken by Maxwell in FL and the one articulated in PL was that the electrotonic state was now defined in terms of the motion of the vortices, which in turn determined the nature of the lines of force. By explaining the electromotive force in terms of the forces exerted by the vortices on the particles between them Maxwell was able to define it as the time rate of change of the electrotonic state.

Although Maxwell was successful in developing the mathematics required for his mechanical model he was insistent that it be considered provisional and temporary in nature. In fact he explicitly states that:

> The conception of a particle having its motion connected with that of a vortex by perfect rolling contact may appear somewhat awkward. I do not bring it forward as a mode of connexion existing in nature, or even as that which I would willingly assent to as an electrical hypothesis. It is, however, a model of connexion which is mechanically conceivable, ~~and~~ easily investigated, and it serves to bring out the actual mechanical connexions between the known electromagnetic phenomena; so that I venture to say that anyone who understands the provisional and temporary character of this hypothesis, will find himself rather helped than hindered by it

> in his search after the true interpretation of the phenomena [PL, P.486].

It would seem then on the basis of this passage that the idle wheel hypothesis had much the same status as the analogies Maxwell used in FL. Although it solved the mechanical problem raised in part one of PL its chief value seemed to be heuristic, serving as a way of <u>conceiving</u> the phenomena that would perhaps suggest further development of the model. It is interesting to note however that Maxwell's reluctance to see the idle wheel hypothesis as any thing more than heuristic at this point is not necessarily extended to the more straightforward account of molecular vortices. In the introduction to part two of PL Maxwell remarked that the hypothesis of molecular vortices gave a "probable" answer to the question concerning the mechanical cause of the difference in the pressure of the medium in different directions. This answer was to be distinguished from answers to the questions of how the vortices were set in rotation and why they were arranged according to the known laws of lines of force about magnets and currents.

> These questions are certainly of a higher order difficulty than ... the former; and I wish to separate the suggestions I may offer by way of a provisional answer to them from the ... hypothesis of vorticies [PL, p. 468].

It appears then that even if Maxwell considered the vorticies themselves as more than simply heuristic devices he was severely constrained as to how their <u>properties</u>

(specifically rotation) were to be accounted for in a legitimate and coherent model. This difficulty is particularly evident from Maxwell's remark at the end of part two of PL:

> We have now shown in what way the electromagnetic phenomena may be imitated by an imaginary system of molecular vorticies... [We] find here the conditions which must be fulfilled in order to give... [the hypothesis] mathematical coherence, and a comparison, so far satisfactory, between its necessary results and known facts [p. 488].

At this point Maxwell had little faith in his theory as a serious alternative to Weber's action at a distance programme. He had been unable to extend his model to electroststics and consequently it lacked the comprehensiveness of Weber's model.[30] However, this difficulty was quickly overcome when, after a period of nine months, Maxwell published parts III and IV of PL which contained an account of electrostatics that proposed a new model of the aether as well as the first derivation of his theory of light.

In part II the magneto-electric medium was seen to be cellular with each cell consisting of a molecular vortex - a rotating parcel of fluid - surrounded by a cell wall consisting of a layer of small spherical particles that functioned as idle wheels. However, in order to explain what was meant by a charged body (the condition of a body when it was charged with respect to the surrounding medium) and to derive the law of attraction between charged bodies he proposed a rather different model of the aether. Instead of the hydrodynamic model he used an elastic solid model in which the aetherial substance formed spherical cells endowed with elasticity. The cells were separated by electric particles whose action on the cells would result in a kind of distortion. Hence, the effect of an electromotive force was to distort the cells by a change in the position of the electric particles. This called into play an elastic force which set off a kind of chain reaction. Maxwell saw the distortion of the cells as a displacement of electricity within each molecule with the total effect over the entire medium producing a "general displacement of electricity in a given direction" [PL, p.491]. Understood literally the notion of displacement meant that the elements of the dielectric had changed position.

Because changes in displacement involved a motion of electricity Maxwell argued that they should be "treated as" [Ibid.] currents [italics added]. Strictly speaking displacement could not be equated with a current (because when it attained a certain value it remained constant) but it could be seen as the "commencement of a current" [PL, p.491] with its variations constituting currents in the positive or negative direction according to whether the displacement is increasing or diminishing. Displacement also served as a model for dielectric polarization, a state in which opposite ends of a dielectric molecule become equally and oppositely charged (something Maxwell pictured as due to a displacement of electricity). Electromotive force was responsible for distorting the cells and its action on the dielectric produced a state of polarization.[71] As Joan Bromberg has pointed out, whenever Maxwell introduces the displacement current he derives it from the idea of displacement simpliciter with no attempt to justify it further.[72] Because displacement satisfied the phenomenological law $R = -4\pi E^2 h$ (cf. n.51) which expressed the relation between polarization and force Maxwell was able to use it to calculate the ether's elasticity (the coefficient of rigidity); the crucial step that led him to identify the electromagnetic and luminiferous aethers. So, given the assumption of elasticity of the electromagnetic medium Maxwell's model had a great deal of explanatory power.

It is interesting to note that until this time no mention of the optical aether had been made. However, once the electromagnetic medium was endowed with elasticity Maxwell relied on the optical aether in support of his assumption about elasticity:

> The undulatory theory of light requires us to admit this kind of elasticity in the lumininferous medium in order to account for transverse vibrations. We need not then be surprised if the magneto-electric medium possesses the same property [PL. p.489].

Maxwell had introduced the equation for displacement as an empirical relation; one that was dependent on observations on the properties of dielectrics subjected to electric forces. Given this relation he went on to point out that it was independent of any theory about the internal mechanisms of dielectrics and:

> ...when we find electromotive force producing electric displacement in a dielectric; and when we find the dielectric recovering from its state of electric displacement with an equal electromotive force, we cannot help regarding the phenomena as those of an elastic body, yielding to pressure and recovering its form when the pressure is removed [PL, p.492].

After a series of mathematical steps[53] Maxwell had the necessary tools for calculating the velocity with which transverse waves were propogated through the electromagnetic aether. Using the formula $V = \sqrt{m/Q}$ where m is the coefficient of rigidity and Q is the aethereal mass density Maxwell arrived at a value for V that agreed with the value calculated for the velocity of light (V = 310,740,000,000 mm/sec.).[54]

> The velocity of transverse undulations in our <u>hypothetical</u> [italics added] medium, calculated from the electro-magnetic experiments of Kohlrausch and Weber, agrees so exactly with the velocity of light calculated from the optical experiments of M. Fizeau that we can scarcely avoid the inference that light consists in the transverse undulations of the same medium which is the cause of electric and magnetic phenomena [PL, p.500].

That this unification of optics and electromagnetism was a complete surprise to Maxwell is evident from his letter to

Faraday, Oct. 19/1861 wherein he claims that he "worked out the formulae in the country before seeing Weber's number ...and I think we have now strong reason to believe, whether my theory is a fact or not that the luminiferous and electromagnetic medium are one".[35]

In addition to the fact that Maxwell was supposedly unaware of Weber's result and hence of the similarity between the velocity of propogation of magnetic effects and light, further evidence for the unexpectedness of this unification stems from the way in which the wave equation was derived from the model. An examination of Maxwell's calculations reveals that the value for V was not in fact a direct consequence of the model. Because of the way in which Maxwell defines the value of m (the coefficient of rigidity) the correct expression for V is $\sqrt{m/2Q}$. Hence, Maxwell's model yields a velocity of propogation equal to the ratio of the electrical units divided by $\sqrt{2}$.[36]

Because Maxwell had set out to calculate only the velocity of propagation of electromagnetic waves and given the way in which he arrived at the result it would seem that his unification of the electromagnetic and luminiferous aethers displayed all the characteristics of a truly consilient theory. Hence, given the kinds of arguments put forth by Friedman the unification that Maxwell achieved ought to be taken as evidence for a realistic interpretation of the theoretical structure and entities (specifically the vortex model of the aether and the displacement current) that facilitated the consilience. This is especially important in light of the role these entities played in formulating the theory of light. The necessary calculations were the result of using electromagnetic laws together with the laws of mechanics that governed the aether model. What Maxwell had in fact shown was that given the specific assumptions employed in developing the mechanical details of his model, the elastic properties of the electromagnetic medium were just those required of the luminiferous aether by the wave theory of light.[27] Hence, what was effected was the reduction of electromagnetism and optice to the mechanics of one aether rather than a reduction of optics to electromagnetism simpliciter.

However, as evidence for a realistic interpretation of Maxwell's model the unification provided little in the way of support. Despite the novelty of the result and the rather straightforward way the reduction occurred both Maxwell and his contemporaries sought evidence independent of the theory that would give credence to the model and the background assumptions employed in the derivation. In what follows I will examine some of the specific difficulties raised by Maxwell's model and how they relate to the more general issue of unification as a criterion for realism. It seems more than slightly ironic that the very structure that enabled Maxwell to achieve his spectacular unification was soon abandoned for a purely formal account devoid of commitment to any particular mechanical model.

4. Unification and Realism - Some Problems for the Electromagnetic Theory. At the beginning of PL Maxwell refers to his vortex aether model as providing a mechanical account that would direct experiment and indicate features that a true theory would have to incorporate:

> My object in this paper is to clear the way for speculation in this direction, by investigating the mechanical results of certain strains of tension and motion in a medium, and comparing these with the observed phenomena of magnetism and electricity. By pointing out the mechanical consequences of such hypotheses, I hope to be of some use to those who consider the phenomena as due to the action of a medium [PL, p.452].

Although Maxwell distinguished between the idle wheel hypothesis and his commitment to the molecular vorticies that supposedly constituted the electromagnetic medium, the reality of these vorticies was by no means an unproblematic issue. Maxwell does not claim (contrary to many commentators, Cf. Siegel [1985]) that the vortex hypothesis is a probable one but only that it gives a probable answer to the question of what causes differences in pressure in the medium. That is, there is no independent reason to think that the vortex hypothesis is probable over and above its explanatory role. Part IV of PL was intended as a proof of the rotary character of magnetism; an argument which, if successful, would support a realistic interpretation of the vorticies and their properties as described by the model. But as became evident in Maxwell's later work the general force of the argument remained independent from any <u>specific</u> mechanical model of the medium. In other words, the details of the theory of molecular vorticies <u>were not</u> supported by the argument for the rotational character of magnetism; it could be given independently of the specifics of the theory of molecular vorticies.[38]

But molecular vorticies were not the only problem to beset Maxwell's theory, the notion of a displacement raised further difficulties. The idea that the electrostatic state was a displacement of something from equilibrium was not new. Thomson had compared electric force to the displacement in an elastic solid while Faraday hypothesized that when the dielectric (which is made up of small conductors) is subjected to an electrostatic field there is a displacement of electric charge on each of the small conductors. An electric current can then be defined as the motion of these charges when the field is varied. It was from this account that Maxwell arrived at the idea that variations of displacement are to be counted as currents. However, in using Faraday's idea Maxwell completely transformed it. According to Faraday's notion displacement was applicable only to ponderable dielectrics and was introduced specifically to explain why the inductive capacity of these dielectrics was different from that of free aether. On

Maxwell's account displacement occurs wherever there is electric force regardless of the presence of material bodies.

In addition to this reinterpretation, the term specified as the displacement term was not a natural mechanical consequence of the model. As a result Maxwell's analogy between electric and elastic displacement could not adequately account for the electrostatic phenomena without additional assumptions. According to Maxwell's model electricity is constituted by small particles and given this view an obvious interpretation of electric charge is to link it to an accumulation of particles.[34] However, Maxwell's model also implies that all currents flow in closed circuits thereby ruling out an accumulation of particles at any place in the circuit. As a result a component must be added to the current so that it is no longer circuital. The dielectric between the coatings of a condenser fulfilled this need and was seen as the origin of the process known as the displacement current. The force of this current was proportional to the rate of increase of the electric force in the dielectric and therefore produced the same magnetic effects as a true current. Hence the charged current could be regarded as flowing in a closed circuit. The major difficulty with this account was that there was no experimental reasons for introducing the displacement current. Because the term had the value zero for the case of steady currents flowing in closed circuits Ampere's law, which gives the relationship between an electric current and the corresponding magnetic field, remained sufficient for dealing with the available data. However, for the case of open circuits the term took non-zero values and thus gave definite predictions for the magnetic effects of open circuits through a modification of the mathematical structure of Ampere's law.[40] It was the electrical rather than the mechanical interpretation of displacement that allowed Maxwell to introduce it into Ampere's law; the same move that enabled him to get a numerical value for the elasticity of the electromagnetic aether. But again the problem was that there was no available data against which to test the predictions provided by the alteration to Ampere's law. In this sense the introduction of displacement was a purely theoretical move devoid of an experimental justification. So, it was the difficulties that surrounded the postulation of this displacement current in space free from matter and electric charge, together with the idea that all currents could be regarded as closed that proved to be a serious hindrance to the acceptance of Maxwell's theory by his contemporaries.[41]

However, it was not merely the <u>postulation</u> of a displacement current and its physical presuppositions that proved to be a source of difficulty.[42] In his derivation of the wave equation the relation between the electromotive force and displacement ($R = -4\pi E^2 h$) played a dual role.

functioning as both an equation of elasticity and as an electrical equation; a situation that resulted in R being interpreted as an electromotive force in the direction of displacement as well as an elastic restoring force opposite to it.[63] Similarily, E was considered an electric constant as well as an elastic coefficient while h was interpreted as a charge per unit area and a linear displacement. In virtue of its double interpretation the equation served as a bridge between the electrical and mechanical parts of the model. The electrical part included the introduction of the displacement current, the calculation of a numerical value for E (the dielectric constant) and the creation of equations ([112] in PL) that describe the quantity of an electric current per unit area[64] as well as the displacement. These equations were also given a rather ambiguous, if not dual, interpretation by Maxwell.[65] It was not until the mechanical part of the model (the derivation of $R = 4\pi m h \left(3/1 + 5/3\ m/\mu\right) = -4\pi(\pi m)h$ and the calculation of the wave velocity from $V = \sqrt{m/\rho}$) was expressed in purely dynamical terms (in "A Dynamical Theory of the Electromagnetic Field" [DT] and Treatise on Electricity and Magnetism [TEM]), thereby eliminating the need for a bridge between the mechanical and electrical parts of the model, that these ambiguities were resolved.

The difficulties Maxwell encountered led him not only to abandon his mechanical aether model but also to transform his view of the displacement current. In doing so Maxwell revised his electro-mechanical theory of light into a purely electromagnetic account. His new approach utilized the mechanical properties of an all-prevading medium but only in a very restricted fashion. Instead the theory was now based on general laws of dynamics, structural principles to which the mechanical properties are subject. Hence the unification of the two aethers was transformed from a physical/mechanical theory to a formal dynamical account devoid of any specific theoretical hypotheses.

5. The Electromagnetic Theory - Later Developments.

5.1 The Dynamical Theory. The problems Maxwell encountered with his aether model led to the notion of a medium being used in a somewhat indirect sense in both DT and TEM. Not only was the idle wheel hypothesis considered to be awkward and provisional but according to the correspondences established by the theory, the particles were assumed to be electrical in nature. This raised an important difficulty for the hypothesis regarding the nature of the medium since the idea of electrical particles or substances in field theory was not looked favourably upon by either Maxwell or by Faraday and others.[44] In correspondence he remarked that he "wanted to clear the electromagnetic theory of light from any unwarrentable assumptions, so that we may safely determine the velocity of light by measuring the attraction between bodies kept at a given difference of potential, the value of which is known in electromagnetic measure".[57]

Maxwell intended to connect the experiments on the ratio of units with measurements of the velocity of light in a direct way that would avoid unnecessary hypothetical elements. By 1864 he wrote to Stokes claiming that he had:

> ...materials for calculating the velocity of transmission of a magnetic disturbance through air founded on experimental evidence, without any hypothesis about the structure of the medium or any mechanical explanation of electricity or magnetism.[58]

In his "Dynamical Theory", published in 1865, a new theoretical framework based on experiments and a few general dynamical principles furnished an account of the relationship between electromagnetic quantities and the velocity of light. Maxwell was able to derive his basic wave equations of electromagnetism from the general dynamical equations of mechanical systems, from which the propagation of electromagnetic waves through space followed without any special assumptions about molecular vorticies or forces between electrical particles.

Although Maxwell had abandoned his vortex aether model as a means of deriving his theory of light he remained committed in some sense to the __idea__ of a medium filling space. In fact, he explicitly remarks that the Faraday effect (the rotation of the plane of polarization of polarized light), the polarization of dielectrics, and the phenomena of optics all led him to the same conclusion; namely, that there is an is an aetheral medium pervading all bodies [DT, p.532]. What is particularly important to note however is the way in which the medium figures or doesn't figure in the account presented in DT. Maxwell claims that the theory he proposes may be called a theory of the electromagnetic field because it concerns the space in the neighbourhood of the electric or magnetic bodies. Moreover, it may be called a <u>dynamical</u> theory because it assumes that in that space there is matter in motion which produces the observed electromagnetic phenomena. The additional reference of the term "dynamical" was to the abstract dynamics developed by Lagrange and elaborated by Hamilton.[69] Lagrangian methods had been used by both MacCullagh and Green[70] in their studies on the luminiferous aether. The approach consisted of the postulation of an appropriate potential energy function for the aether without specifying the details of its mechanical structure. Once this was done a wave equation was derived, using the Lagrangian variational principle[71], that allowed the aether to be treated as a mechanical system without any specification of the machinery that gives rise to the characteristics exhibited by the potential energy function. In contrast to the method of PL where Maxwell attempted to describe "a particular kind of motion and a particular kind of strain so arranged as to account for the phenomena", in DT he "avoid[ed] any hypothesis of this kind". He goes on to point out that

...in using such words as electric momentum and electric elasticity in reference to known phenomena of the induction of currents and the polarization of dielectrics, I wish to merely direct the mind of the reader to mechanical phenomena, which will assist him in understanding the electrical ones. All such phrases in the present paper are to be considered as illustrative and not as explanatory [p.526].

Once again we see Maxwell invoking the distinction between theoretical hypotheses that are to be understood as real physical possibilities and those that are merely instrumental, as in the case of the analogies introduced in the earlier papers. Because Maxwell is agnostic as to the theoretical structure of the medium and the mechanisms involved in displacement, his intention was to use the descriptive terminology as a way of providing some physical understanding of how one might possibly account for the phenomena. However, since no true hypothesis was being put forth Maxwell intended that terms be used in a way similar to that of analogies. But, unlike the case of idle wheels and the imaginary fluid Maxwell, seems to regard the aether and displacement as physical possibilities despite their rather precarious ontological status. Maxwell does remark that when speaking about the energy of the field he wished to be understood literally. All energy is to be considered mechanical energy whether it exists in the form of motion, elasticity or any other form. So, in other words, the energy in electromagnetic phenomena is to be understood as mechanical energy with the only, and perhaps most crucial, question being, Where does it reside? According to action at a distance theorists the energy resides in electrical bodies, conducting circuits and magnets in the form of potential energy (the power to produce certain effects at a distance). On Maxwell's electromagnetic theory the energy resides in the electromagnetic field, that is, in the space surrounding the electrified and magnetic bodies as well as in the bodies themselves. The energy consists of two different forms, magnetic and electric polarization which, according to what Maxwell terms a "very probable hypothesis"[DT p.564], can be described as the motion and strain of one and the same medium.

The fact that there was some reason to believe that the medium was capable of motion stemmed from the phenomena of heat and light (DT, p.525), especially the Faraday effect (magneto-optic rotation). In the latter case Maxwell claims that

> We have <u>warrantable grounds for inquiring</u> whether there may not be a motion of the aethereal medium going on wherever magnetic effects are observed, and we have some reason to suppose that this motion is one of rotation [DT, p.529] [emphasis added].

Maxwell thought it possible that the motion of the medium (caused by electric currents and magnets) could be communicated from one part to another by forces arising from the connections of those parts. Due to the action of the forces there is a certain yielding that depends on the elasticity of the connections so that the energy of the medium is composed of the actual energy of the motion of the parts and the potential energy is stored in the connections as elasticity (displacement). However, aside from the very general claim about matter in motion Maxwell offered no hypothesis about the nature of the medium, but instead claimed that we are "led to the conception of a complicated mechanism capable of a vast variety of motion"[DT p.533]. Such a mechanism, he went on to remark, must be subject to the general laws of dynamics and one ought to be able to work out all the consequences of its motion provided that one knows the form of the relation between the motions of the parts [DT, p.532]. Because <u>all</u> the consequences of the motion of the medium could be deduced, Maxwell was thereby able to account for mechanical and electrical actions that take place in accordance with the fundamental principles of dynamics.

Although Maxwell could deduce certain consequences from structural considerations that apply to the medium, the method is different from the one presented in PL because the medium is introduced and presented in a significantly different way. In the latter case the mechanical properties of the model provided the mechanisms for deducing electromagnetic effects whereas in DT the model plays virtually no role; all consequences are derived from the Lagrangian formalism. In addition to this the very idea of an electromagnetic medium was introduced in rather different manner using a kind of analogical method whose origins could, at least in this context, be traced to the beginning of DT. Based on some considerations published by Thomson[72] concerning the density of the luminiferous medium, Maxwell claimed that we can take the data from a branch of science independent of the one in which we are dealing and, given certain properties of an optical medium (especially elasticity and the velocity computation), we can draw conclusions about its identity with the electromagnetic medium. In PL he postulated a mechanical electromagnetic medium and subsequently discovered that it could be identified with the optical aether whereas in DT he introduced the optical aether and discussed the phenomena of light that led to its postualtion and the delineation of its properties. This was followed by a discussion of the Faraday effect, its implication that magnetism consists in a motion of the optical aether and, finally, a discussion of the facts about electricity amd magnetism that independently suggests the existence of an electromagnetic aether. These three independent domains of evidence lead to the assumption that there is an aethereal medium prevading all space. The physical analogy that lends plausibility to the assumption of an elastic electromagnetic medium in PL (cf. p.489) is reintroduced in DT as a kind of consilience of inductions. Although it is the elasticity equation in PL that suggests the identity of the two aethers, its initial formulation stems from a physical analogy. Maxwell postulated the existence of an electromagnetic medium endowed with elasticity and then brought in the optical aether to lend credence to the idea:

> The undulatory theory of light requires us to admit this kind of elasticity in the luminiferous medium in order to account for the transverse vibrations. We need not then be surprised if the magneto-electric medium possesses the same property [PL, p.489].

This physical analogy is not used in DT and instead Maxwell appeals to a kind of mathematical analogy based on a formal account of the motions of the parts of the medium which can be used to determine *all* the consequences of the motion of the medium. This, together with a kind of consilience based on what Maxwell terms "experimental facts of three kinds"[DT p.564], justifies the introduction of laws of mechanical forces laws that would be *applicable* to the medium but do not require the *hypothesis* of the medium for their introduction. As I mentioned above, although there is a sense in which Maxwell accepts the existence of the medium no *specific* model endowed with particular properties can be given.[73] He makes a point of specifying that unlike the case of PL where his conclusions depended completely on his mechanical model of the aether and the displacement current, in DT the conclusions depend only on experimental facts from which certain dynamical laws can be deduced: laws that are independent of any hypothesis concerning the existence of a medium. Although Maxwell retained the idea that electric and magnetic energy are diseminated he avoided commitment to any hypothesis about their mechanical form in space.

Despite Maxwell's claims to provide deductions from experimental facts it is important to point out that his account still requires the postulation of a displacement current; something that had not been verified by experiment. Using a mechanical analogy between two closed currents interacting by means of the aether and the interaction of two driving points of some mechanism, Maxwell developed the theory of interacting closed currents.[74] Because all physical concepts in DT, except matter and motion, were to be understood as merely illustrative, Maxwell was able to develop an account that was consistent with mechanical principles but did not involve any specific details of the properties or structure of the aether.[75] This application to closed currents was extended in part III of DT to open currents by the introduction of the displacement current. It was also in this section that Maxwell formulated the general equations of the theory (now commonly referred to as Maxwell's equations); which, in contrast to those presented in PL, could be interpreted without ambiguity.[76]

Because the mechanical model had been abandoned, displacement was no longer seen as an elastic restoring force for the mechanical aether. Hence the displacement current was freed from its connections with mechanism and the negative value in $R = -4E^2h$, disappeared[77]; and as I mentioned above, the part of the theory that concerned the electrostatic equations was established without the use of any of the mechanical properties of the aether.[78]

Maxwell did retain a basic feature of his theory of charge as formulated in PL; that is, the identification of a charge and a discontinuity in displacement in a dielectric medium. The main difference between the accounts in PL and DT was that in the former case an equation connecting displacement with charge was not explicity given, whereas in DT the relation appeared as one of the basic equations.[79] Given Maxwell's claim about open currents the notion of a displacement that gives rise to transient currents was essential; without it there would be no basis for the electromagnetic theory of light. So, although displacement retained a prominent position in DT its role was somewhat different from the original one it played in PL. It was no longer associated with the change in position of rolling particles, rather Maxwell defined it as the motion of electricity; that is, in terms of a quantity of charge crossing a designated area.[80] Finally, in addition to the problems with the displacement current discussed above a further source of difficulty for Maxwell's account was that the postulation of the displacement current was not necessary; action at a distance theories could explain electromagnetic phenomena perfectly well without it.

An important dimension of DT was the revival of various of Maxwell's earlier ideas about the geometry of lines of force. The intention was to extend the electromagnetic theory to explain gravitation but unfortunately he was unsuccessful in this endeavour due to an inability to represent the physical nature of lines of gravitational force. In DT he had pointed out that gravitation, being an attractive force, had the consequence that the energy of any gravitational field of a material constitution was less wherever there was a resultant gravitational force. Hence, those parts of space in which there was no resultant force would possess an enormous energy. However, as Maxwell himself admitted, "I am unable to understand in what way a medium can possess such properties" [DT, p.571].

In summary then, Maxwell's goal in DT was to use the equations governing electrogmanetic phenomena to present an account of the possible conditions that existed in a mechanical medium responsible for the exchange and transmission of kinetic and potential energy. The method allowed Maxwell to treat field variables as generalized mechanical variables interpreted within the context of the Lagrangian formalism. What this meant with regard to the physical model of the theory is discussed below. for now let me stress that although the dynamical theory was applicable to mechanical phenomena the equations were not connected to any specific mechanical model. Nor was the __motivation__ for the theory based on mechanical hypotheses. Maxwell was anxious to give some physical interpretation to the formalism but he is quick to remind us that the concepts he uses are to be understood as merely illustrative. The core of the theory is the Lagrangian formalism, a highly abstract and general method applicable to a __wide__ range of __different__ models of the data. In contrast the theory expounded in PL was concrete and intimately tied to the specific mechanical model Maxwell used.

5.2 The Treatise on Electricity and Magnetism. The methodology used in DT was extended and more fully developed in __TEM__. Once Maxwell had formally expressed the fundamental laws of electromagnetism and the relations between the various quantities involved he had to show how these laws could be deduced from general dynamical laws applicable to any system of moving bodies. Although the connections of the motions of the medium with specific variables were eliminated from the equations (the variables being independent of any particular form of these connections) Maxwell did make use of mechanical principles in the development of the theory. For example, in the chapter devoted to the Faraday effect Maxwell once again appealed to the hypothesis of molecular vortices. But in applying the Lagrangian formalism to magneto-optic rotation Maxwell remained agnostic about their nature and remarked that:

> the consideration of the action of magnetism on polarized light leads... to the conclusion that in a medium under the action of magnetic force <u>something</u> belonging to the same mathematical class as angular velocity, whose axis is in the direction of the magnetic force, forms part of the phenomena [<u>TEM</u> 2:399-417, esp.406,408] [emphasis added].

Given a mechanical ontology and a universe of matter in motion Maxwell assumed that that which belongs to the same mathematical class as an angular velocity must be the angular velocity of some rotating portion(s) of a medium filling space. On the basis of experiment it was shown that there were no sizable angular momenta associated with these rotations. So, because the rotating parts of the medium had to be small, it seemed reasonable to suppose that the rotation was caused by molecular vorticies each rotating on its individual axis.

In light of Maxwell's remarks in DT on the status of mechanical and electrical concepts it seems reasonable to assume that the hypothesis of a vortex aether model (even one devoid of idle wheels) represented, at the <u>very most</u>, a physical possibility and nothing more. In the treatise he stresses that his aim is to formulate a dynamical theory of the field that would emphasize the primacy of concepts like energy, velocity and momentum. In doing so Maxwell attempted to translate the results achieved through the use of the Lagrangian method from the language of the calculus into the language of dynamics. The outcome of this process would be the depiction through language of a mental image of a property of moving bodies rather than an algebraic process [TEM 2:210].[81] This method kept "out of view the mechanism by which the parts of the system are connected" [TEM 2:209] allowing Maxwell to avoid the explicit formulation of a mechanical aether model while keeping in mind the ideas "appropriate to the fundamental science of dynamics" [TEM 2:210].

When using the abstract principles of dynamics Maxwell remarked that in the study of any complex objects we must "fix our attention on those elements of it which we are able to observe and to cause to vary, and ignore those which we can neither observe not cause to vary".[82] He uses the allegory of the belfry to illustrate this principle of inquiry.

> In an ordinary belfry, each bell has one rope which comes down through a hole in the floor to the bell-ringer's room. But suppose that each rope, instead of acting on one bell contributes to the motion of many pieces of machinery, and that the motion of each piece is determined not by the motion of one rope alone, but by that of several and suppose, further, that all this machinery is silent and utterly unknown to the men at the ropes, who can only see as far as the holes in the floor above them. Suppose all this, what is the scientific duty of the men below? They have full command of the ropes, but of nothing else. They can give each rope any position and any velocity, and they can estimate its momentum by stopping all the ropes at once, and feeling what sort of tug each rope gives. If they take the trouble to ascertain how much work they have to do in order to drag the ropes down to a given

> set of positions, and to have found the potential energy of the system in terms of the known coordinates. If they then find the tug on any one rope arising from a velocity equal to unity communicated to itself or to any other rope, they can express the kinetic energy in terms of the coordinates and velocities. These data are sufficient to determine the motion of every one of the ropes when it and all the others are acted on by any given forces. This is all that the men at the rope can ever know. If the machinery above has more degrees of freedom than there are ropes the coordinates which express these degrees of freedom must be ignored. There is no help for it [*Papers*, pp.783-84].

Although Maxwell claimed that there was "good evidence for the opinion that some phenomenon of rotation is going on in the magnetic field"[*TEM* 2:470 Sec.31], he nevertheless claimed that the attempt made to imagine a working model of the mechanism connecting the different parts of the field ought to be seen as nothing more than a demonstration that it is *possible* to *imagine* a mechanism capable of producing a connection mechanically equivalent to the actual connection of the parts of the field what ever it might be. The problem of determining the mechanism required to establish a given species of these connections was that there existed an infinite number of solutions. This too was a consequence of the Lagrangian method; it enabled one to express all that could be deduced, within reasonable limits, from a given set of observational data. That Maxwell focused on observable consequences rather than specifying a particular mechanical model as a candidate for truth is further evident from his remarks in chapter IV of TEM.

> What I propose to do now is to examine the consequences of the assumption that the phenomena of the electric current are those of a moving system, the motion being communicated from one part of the system to another by forces, the nature and laws of which we do not even attempt to define, because we can eliminate these forces from the equations of motion by the method given by Lagrange for any connected system [TEM, 198].

In the application of Lagrangian formalism[14] to electromagnetism the information derived from experiments on interacting electric currents was analagous to observations of the behavior of ropes in the case of the belfry. The hidden mechanism was composed of motions that were assumed to exist in the medium surrounding the currents. Once we identified the generalized coordinates and velocities for a system of interacting currents we can then form a general expression for the kinetic energy of the system and apply the Lagrangian equations of motion in order to derive expressions for the generalized forces. Operating within the mechanical constraints of matter in motion Maxwell's commitment to the aether seems to be a commitment to **something** functioning as the seat of energy for electromagnetic phenomena. Beyond that and beyond the fact that the electromagnetic effects could be described by equations Maxwell was unwilling to speculate.

Among the key features of the Treatise are Maxwell's representations of the energy of the field and the stress in the medium. The kinetic or electromagnetic energy of the field was expressed in terms of the magnetic force and magnetic induction, or alternatively, by the electrotonic state and electric current; both of which had the same observable consequences. Electrostatic or potential energy was expressed in terms of electric force and the displacement current. Although the notion of stress in the medium was represented by electromagnetic quantities and closely associated with the theory of polarization outlined in TEM, Maxwell was vague as to the nature of this stress or strain and the nature of polarization. The latter was defined in terms of a particulate theory but no mechanical model was suggested.[84] As a result the two questions that would naturally arise, namely, what is the state of stress in the aether which will enable it to produce the observed electrostatic attractions and repulsions between charged bodies, and, what is the mechanical structure of the aether that would give rise to this state of stress, were not sufficiently answered:

> It must be carefully borne in mind that we have made only one step in the theory of the action of the medium. We have supposed it to be in a state of stress, but have not in any way accounted for this stress, or explained how it is maintained... I have not been able to make the next step, namely, to account by mechanical considerations for these stresses in the dielectric. I therefore leave the theory at this point [TEM, V.I p. 132].

Indeed, in the Treatise #110-11 and #645 Maxwell emphasized that while he postulated the existence of the stress he was not making suggestions about its nature but was merely showing that a representation of such a stress is <u>possible</u>.

Displacement, magnetic induction and electric and magnetic forces were all defined in the treatise as vector quantities[89], together with the electrostatic state which was termed the vector potential. All were fundamental quantities for the expression of the energy of the field and for the electromagnetic field equation. Vector quantities replaced lines of force and not only were they conducive to Maxwell's geometrical approach but vectorial representation was for him a method of thinking because it required that we form a mental image of the geometrical features represented by the symbols. Because it was a valuable method of representing directed quantities in space Maxwell used these directed properties of vectors as a way of representing particulate polarization, where opposite parts of the particles had different polarity.

Although the emphasis on polarized particles was intended merely as an aid for understanding the theory, it also raised several problems with respect to its basic interpretive mechanisms. Maxwell's particulate representation of the medium made it difficult to distinguish between aether and matter. This together with the fact that he had no conception of the model of connection between matter and the aether raised further difficulties for providing an adequate account of the roles of material dielectrics and the aether. Although all the electromagnetic energy or force was _in principle_ located in the medium, Maxwell spoke of the aether as the limit of a dielectric of polarized particles. This of course contributed to the confusion between the aether and material dielectrics.[9] As a result, there was no independent variable representing the electromagnetic field when the theory was put in Lagrangian form. In addition he considered certain quantities involved in the equations for material dielectrics to be fundamental quantities of electromagnetism: quantities that were identified as vector potential of the electric current, electromagnetic momentum at a point and Faraday's electrotonic state.[96] Hence, Maxwell's theory of the electromagnetic field described the state of both the material and aetherial media.

Although the nature of displacement was quite different in _TEM_ as compared to earlier works this difference resulted in additional vagueness and ambiguity. Maxwell called the state of the medium giving rise to charge "displacement" and described it using expressions like polarization and displacement of electricity. However, the precise nature of these notions are left in doubt:

> Whatever electricity may be, and whatever we may understand by the movement of electricity the phenomena which we have called electric displacement is a movement of electricity in the same sense as the transference of a definite quantity of electricity through a wire is a movement of electricity [_TEM_,1, pp.166-67].

The problem was that displacement was crucial for maintaining any connection between electric waves and light waves. Faraday's experiments on magneto-optic rotation did suggest that the two types of waves were connected, but in order to explain this Maxwell required the additional assumption that there must be a connection between the motion of the aether that causes magnetic force and that which constitutes light. One could not give a mechanical account of the rotation of the plane of polarized light without an assumption relating these two kinds of motion. As a result Maxwell supposes that linear displacement of a point in the aether is the mechanism that gives rise to light. Given this assumption the laws of propagation of polarized light in a magnetic field agreed with the results of Verdet's experiments on the opposite rotation of polarized light by paramagnetic and diamagnetic substances.[87]

Finally, it is important to point out that the displacement current did not emerge as a result of his application of Lagrangian mechanics.[88] In fact Maxwell's Lagrangian treatment of electromagnetism becomes much less attractive once the displacement current is introduced. When dealing with closed conduction currents the generalized velocities that figured in his analysis were currents that could be directly measured and operationally defined. When the displacement current is included among the generalized velocities the situation changes. Because of the theoretical nature of the displacement current it is as if there were additional ropes influencing the motion of the mechanism responsible for ringing the bell, without their being accessible to the bellringers. Not only is the nature of the "hidden mechanism" involved in displacement unknown but no empirical or operational account of it could given.

6. Realism and Dynamical Explanation. Despite the fact that Maxwell was an instrumentalist in some cases, (particularly with respect to the notions developed in FL and the hypothesis of idle wheels), there are other instances where he seems to be a committed realist. For example, many commentators[27] have argued that throughout his career Maxwell remained firmly committed to the notion of an aether prevading space and to the fact that there was some rotation going on in the medium that was responsible for electromagnetic phenomena. Such a reading is not without textual support. However, the interesting thing to note regarding Maxwell's realism is that he was fully aware that none of his explanations of the aether could be considered a true description. He does claim in a letter to Faraday [1861, op. cit.] that even if his own theory is false we nevertheless had strong reason to believe that the electromagnetic and luminiferous aethers were one and the same. Although this seems to demonstrate a definite commitment to the aether as a causal mechanism it is

important to understand this remark in the context of Maxwell's overall views on the subject. In fact it was not unusual for Maxwell to adopt different points of view regarding the existence of the aether. For instance in IEM he claims that "there must be a medium or substance in which ...energy exists after it leaves one body and before it reaches [an] other" [2:866], while in a letter to Bishop Ellicott [1876] he refers to the aether as a "most conjectural scientific hypothesis".[20]

Although the aether and displacement played a central role in the physical understanding of the formal theory of electromagnetism the certainty Maxwell had achieved with respect to his field equations was limited to the mathematization of the phenomena, as opposed to the certainty of the physical description that accompanied the mathematical formalism. Indeed the distinction between mathematical and physical theories, between nature and the various ways nature is represented, is crucial in Maxwell's thought. And, in light of his work on the electromagnetic theory, these distinctions present a rather interesting mix of realist and anti-realist sentiments. Beginning with the discussion in DT and extending into the Treatise Maxwell's ideas about hypotheses often reflect the kind of agnosticism characteristic of the current attitudes of scientific anti-realists.[a] Although his writings indicate a kind of "in principle" committment to the aether and displacement they have the status of a "something I know not what": and although these entities provide the best possible explanation for the phenomena Maxwell feels that he is not at liberty to frame any hypotheses about their nature. He therefore remarks in the Treatise that all mechanical and electrical concepts are to be understood as illustrative. Despite this claim there is no doubt a distinction between the status of the (molecular vortex) aether and displacement and the illustrative role played by idle wheels. The former should be understood as at least a candidate for reality while the latter plays a merely heuristic role. In TEM Maxwell knows that he cannot justify any claims about the aether and displacement and rather than bring them forth as literally true descriptions about reality he classifies them as notions that perform an illustrative role in our understanding of the theory. Although it may be more accurate to characterize Maxwell's overall attitude toward scientific theories as that of an anti-realist, even if if one were to interpret his views as those of a realist the description would be markedly different from the characterizations of realism that we find in the literature. Although he could perhaps be described as possessing a limited belief in the aether it failed to figure, in any important way, in the development of the electromagnetic theory, at least at the later stages. As a result he remained completely agnostic about its precise role in accounting for the phenomena. Consequently his situation cannot be assimilated to the rather robust characterizations of theoretical entities condoned in current accounts of scientific realism.

I address this issue primarily because an insight into Maxwell's way of thinking about scientific theories provides some interesting perspectives on the realism issue in philosophy of science and particularly on the role of theory unification and consilience. The motivating factors behind Maxwell's use and understanding of the method of dynamical explanation reveal several interesting points about the nature of mathematical theories, their relationship to the physical phenomena they were designed to explain, and the kinds of epistemological conclusions we can draw on the basis of successful mathematizations.

There are two relatively obvious reasons that suggest why Maxwell adopted the method of dynamical theory. First of all he was unsuccessful in carrying out experiments (the nature of which I shall discuss in the next chapter) that were intended to test the electromagnetic theory. As a result his desire to move away from a method involving speculation about physical hypotheses in favour of a mathematical theory was strengthened. Dynamical theory or, more specifically, Lagrange's equations of motion, involves no assumptions concerning the existence of specific bodies or their properties. Maxwell's approach simply implied that all phenomena could be understood in terms of matter and motion.

However, the reliance on a general dynamical approach was also bound up with Maxwell's philosophical views about the nature of scientific knowledge, physical theory and the role of models.

In his address to the Mathematical and Physical sections of the British Association (1870) Maxwell was careful to distinguish between the structure of nature and the concepts we use to describe it:

> ...molecules have laws of their own, some of which we select as most intelligible to us and most amenable to our calculation. We form a theory from these partial data and we ascribe any deviation of the actual phenomena from this theory to disturbing causes. At the same time we confess that what we call disturbing causes are simply those parts of the true circumstances which we do not know or have neglected, and we endeavour in future to take account of them. We thus acknowledge that the so-called disturbance is a mere figment of the mind, not a fact of nature, and that in natural action there is no disturbance [<u>Papers</u>, 2, p.228].

It was this distinction between nature and its representations that formed the basis for much of Maxwell's ideas about models and analogies. Because these modes of representation were thought to be heuristic devices, or at best descriptions of what nature <u>may</u> be like, it was possible to propose several different and even contradictory models of the phenomena without a firm commitment to any particular one. And indeed this approach fit nicely with the methodology suggested by the Lagrangian formalism. Just as there are many ways of phrasing a sentence and an infinite number of ways of representing an operator by matrices. similarily, every formulation of a physical theory can be represented in a number of ways. It is also the case that different physical theories can be formulated in the same manner. For instance, insofar as the formal structure of the theory is concerned analytical mechanics, electromagnetism and wave mechanics can all be deduced from a variational principle; the result being that each theory has a uniform Lagrangian appearance.[1]

The Lagrangian equations of motion describe a mechanical system in which the free motions of a number of bodies are restricted by some form of connections among them. Lagrange introduced a set of generalized coordinates equal in number to the actual degrees of freedom of the system. This set was the result of a reduction in the number of variables by the number of constraints on the system (which of course was determined by the connections) and was chosen so as to correspond to quantities that we had information about. The state of the system can be determined provided that the displacements and velocities are known; but, the generalized coordinates can be <u>any</u> quantities that are sufficient to determine this state. So, velocities, moments and forces related to the coordinates in the equations of motion need not be interpreted literally in the fashion of their Newtonian counterparts. If we recall the example of the belfry, the ropes correspond to the generalized coordinates that determine the configuation of the system, while the mechanism responsible for the bellringing (the connection between the movement of the ropes and the ringing) remains unknown. Similarly, we can think of the field as a connected mechanical system with currents and integral currents as well as generalized coordinates corresponding to the velocities and positions of the conductors. We can have a quantitative determination of the field without knowing the actual motion, location and nature of the system.

One of the most interesting and useful features of the Lagrangain formalism is its degree of generality, which allows it to be applied to a variety of phenomena regardless of the nature of the physical facts involved. This is particularly useful because when the details of the system or mechanism of a particular process is unknown, as in the case of Maxwell's electromagnetic theory, or when the system is simply assumed to lack a mechanism we can always apply the Lagrangian method.[8] Consider the following example. Suppose we want to find the path of a light ray between two given points in a medium of known refractive index. In the qualitative interpretation of this problem we employ an analogy that compares light rays and particles, that is, we describe light rays <u>as if</u> they were trajectories of particles or entities endowed with a mass. However, in the quantitative or formal treatment no mention of masses, motions or forces is made. Instead a physically uninterpreted parameter t is introduced that allows us to specify optical length A as $A = \int_{t_1}^{t_2} L\,dt$.[9] Nothing about the formalism or the use of dynamical analogies allows us to infer that light rays <u>are</u> particle trajectories, or that optics can be reduced to mechanics.. The variables do not have an independent physical meaning and because the Lagrangian formalism can be applied to such a variety of different systems no conclusions about the nature of the system can be drawn. As was the case with Maxwell's theory, we cannot understand an analogy or formal identity as expressing an identity in kind or nature. This is especially true given the status of mathematical quantities within the electromagnetic theory. For example, the vector potential was used to denote a number of different physical quantities such as Faraday's electrotonic state as well as the electric momentum at a point:

> The whole history of this idea in the mind of Faraday...is well worth study. By a course of experiments, but without the aid of mathematical calculation, he was led to recognize the existence of something which we now know to be a mathematical quantity, and which may even be called a fundamental quantity in the theory of electromagnetism. But as he was led to this conception by a purely experimental path, he ascribed to it a physical existence, and supposed it to be a peculiar condition of matter, though he was ready to abandon this theory as soon as he could explain the phenomena by any more familiar forms of thought. Other investigators were long afterwards led up to the same idea by a purely mathematical path [<u>TEM</u>, sec. 540].

In keeping with this distinction and despite his commitment to the mechanical constraints of matter and motion Maxwell had a rather critical philosophical attitude toward the ontological status of matter. The form of a dynamical theory facilitated a distinction between matter as the quantity (mass) that is part of abstract dynamics and matter as it applies to the actual bodies of our perceptual experience. In the latter context it is that "unknown substratum against which Berkeley directed his arguments" which is never directly perceived by the senses. It is in a sense a presupposition, the supposed cause of our perception but as a thing in itself it is unknowable. Alternatively, as an abstract dynamical concept it enjoys the same status as a straight line:

> Why, then, should we have a change of method when we pass on from kinematics to abstract dynamics? Why should we find it more difficult to endow moving figures with mass than to endow stationary figures with motion? The bodies we deal with in abstract dynamics are just as completely known to us as the figures in Euclid. They have no properties whatever except those which we explicity assign to them [Papers, p.779].

Maxwell goes on to remark [p.781] that real bodies may not have a substratum but as long as their motions are related to each other according to the conditions laid down in dynamics we may call them dynamical or material systems.

Again, mindful of the distinction between nature and its representations and its importance in dynamical explanation Maxwell identified two conditions that must be adhered to in the development of scientific theories. One is the requirement of consistent representation which is simply the demand that any proposed hypothesis must be consistent with the fundamental principles of dynamics, specifically Newton's laws of motion and the conservation of energy. The other more stringent criterion is the condition of independent evidence. From the discussion in chapter 1 we know that the requirement of independent evidence can take on a variety of forms. In contexts like the unification of the aethers we use data from the domain of optics, a field supposedly independent from electromagnetism, as evidence for the reduction of the two media to one. However, there also exists the problem of providing evidence that is independent of the theory or theories involved, that is, evidence that supports the existence of a particular phenomena on the basis of experimental findings rather than on its place within any number of theoretical explanations. Consider the case of the displacement current. Although it was instrumental in facilitating the unification of electromagnetism and optics and, as such, could be seen as receiving confirmation from both domains there was no theory neutral, or experimental, evidence to suggest its existence. The inductive strength received from the unifying role played by displacement was insufficient to provide the support required for its acceptance.

The condition of independent evidence also influenced Maxwell's disregard for the method practiced by the French molecularists and now commonly referred to as hypothetico-deductive. Although he adopted the equations of Lagrange he downplayed the physical content of the theories:

> In forming dynamical theories of the physical sciences it has been a too

> frequent practice to invent a particular dynamical hypothesis and then by means of the equations of motion to deduce certain results. The agreement of these results with real phenomena has been supposed to furnish a certain amount of evidence in favour of the hypothesis.
>
> The true method of physical reasoning is to begin with the phenomena and to deduce the forces from them by a direct application of the equations of motion. The difficulty of doing so has hitherto been that we arrive, at least during the first stages of the investigation, at results which are so indefinite that we have no terms sufficiently general to express them without introducing some notion not strictly deducible from our premisses.
>
> It is therefore very desirable that men of science should invent some method of statement by which ideas, precise so far as they go, may be conveyed to the mind, and yet sufficiently general to avoid the introduction of unwarrantable details [Papers, 2, p.309].

The method of the molecularists violated the condition of independent evidence by assuming at the outset the existence of configurations of particles and specified forces between them. In addition, because dynamical hypotheses admit of an infinite number of solutions we require independent evidence as a way of determining the true solution. Alas it was because this was unattainable that one employed the method of Lagrange:

> ...when we have reason to believe that the phenomena which fall under our observation form but a very small part of what is really going on in the system, the question is not - what phenomena will result from the hypothesis that the system is of a certain specified kind? but - what is the most general specification of a material system consistent with the condition that the motions of those parts of the system which we can observe are what we find them to

be.[5]

Finally, and perhaps most importantly, an obvious advantage of the generality of the Lagrangian methods was its role in facilitating the process of unification. The framework based on the definition of a Lagrangian function exhibits a compact and analytic (deductive) character and consequently there is no difficulty in providing a unified treatment of electromagnetic and optical phenomena. Although Maxwell assumed the field to be a moving system with the motion being communicated from one part of the system to another by forces, the forces were eliminated from the equations of motion and the structure of the field was not specified. The physical model assumed the status of a dynamical analogy functioning as a heuristic device and was not to be understood as an explanatory hypothesis about the <u>actual</u> mechanisms involved. According to TEM electrical action was a phenomenon due to an unknown cause and subject only to the general laws of dynamics. The Lagrangian method allowed Maxwell to provide a mathematically unified account of the system without the excess of physical hypotheses. Although the unification began as a physical hypothesis about the nature of the electromagnetic and optical aethers this approach was abandoned in favour of a purely formal account. For Maxwell and his contemporaries the physical unification simply failed to provide the kind of evidence required for recognizing the aether and the displacement current as entities complete with properties and capable of being understood in a robust sense. The requirement of independent evidence needed to be realized: mere calculations and theoretical unification was not enough:

> ...if the study of two different branches of science has independently suggested the idea of a medium, and if the properties which must be attributed to the medium in order to account for electromagnetic phenomena are of the same kind as those which we attribute to the luminiferous medium in order to account for the phenomena of light, the evidence for the physical existence of the medium will be considerably strenghtened....IF [italics added] it should be found that the velocity of propogation of electromagnetic disturbances is the same as the velocity of light, and this not only in air but in other transparent media, we shall have strong reasons for believing that light is an electromangetic phenomenon, and the combination of the optical with the electrical evidence will produce a conviction of the reality of the medium similar to that which we obtain, in the case of other kinds of matter, from the combined evidence of the senses [TEM 781].

This requirement of independent evidence would not be fulfilled until Hertz's experiments on electromagnetic waves.

7. Summary and Conclusions. The use of models in Maxwell's work is, I think, representative of the approach facilitated by the embedding account discussed in chapter one; especially as it relates to the kind of anti-realism characterized by van Fraassen's semantic view of theories. Although Maxwell's approach can sometimes be associated with what we would typically refer to as classical instrumentalism, in other contexts he seemed intent on using models like the vortex aether model as an account of what nature may possibly be

like. Throughout Maxwell remained more or less agnostic about the reality and truth of his hypotheses and in later stages of his work focused on the more formal aspects of the electromagnetic theory achieved through the use of Lagrangian methods; an approach that required no commitment to any <u>specific</u> theoretical model.

One could of course argue that despite Maxwell's agnosticism about particular models and their details, he nevertheless remained commited to the <u>idea</u> of an aether and the fact that there was some form of rotary motion taking place within it. This kind of commitment however seems hardly enough to make Maxwell a scientific realist, especially in light of contemporary characterizations of the position. Recall that in chapter one (Cf. pp.34-35) I argued that one could conceivably retain a kind of limited realist position by maintaining a commitment to a general molecular theory or structure, the <u>details</u> of which change over time. For instance, we could interpret the continuity between the gaseous and liquid states of matter as indicative of a kind of ontological unity furnished by the molecular structure. This structure could be interpreted as providing a theoretical framework that could remain constant throughout the evolution and changes in the models suggested by the kinetic theory of gases. A similar approach could be utilized in the case of the aether and maybe even the displacement current although the latter case is slightly more difficult to motivate since there were other problems with the the displacement current that had to do with its role in the theory. However, as I argued a realism of this restricted form has none of the utility that the more robust accounts of scientific realism were designed to have. Instead one's beliefs and ontological commitments about theoretical structure/entities become generalized to belief in a "something I know not what"; a form of metaphysical speculation reminiscent of medieval accounts of science. In addition, if we are commited only to an entity devoid of properties then the realist position collapses, in a sense, into a position no different from anti-realism. When one attempts to explain or predict certain phenomena any appeal to theoretical entites or structures is at the same time an appeal to those structures as they are characterized by a particular theory; that is, we endow them with certain properties that in turn enable us to generate explanations. If one must remain skeptical about the properties then the explanatory power of the model or theory cannot be construed as evidence for realism. So, postulating certain structures while remaining agnostic about their properties allows us to maintain a restricted form of realism that serves no function within explanatory or predictive contexts. Moreover, the position presupposes that one can isolate entities independently of their theoretical descriptions; that we can somehow know or have good evidence that a particular entity exists while remaining skeptical about how they are characterized. Not only is this position extremely difficult to articulate but it would seem that the result, in practice, is no different from the anti-realist who is skeptical about both the structures and their properties. Instead a more plausible approach seems to be the view that theoretical models can provide a number of possible accounts of the phenomena without being forced to specify any particular one as the true depiction of the material world. And indeed such was the case with the electromagnetic theory. The entire theory was formulated in a way that precluded commitment to any specific account of the aether and hence as a theoretical structure it failed to occupy a place of importance within the final characterization of the theory. This is not to say that any one model is as good as any other or that we can neglect the search for better models once we have found one that is adequate. Instead it is simply to recognize the fallibility of our models and the extent to which they are only a partial description of the world. The fact that a theory can be articulated using a variety of different models and sometimes requires many different models in its application seems to be strong evidence against the view that models furnish us with a unique and true description of reality.

A similar story can be told for the case of the displacement current. Although Maxwell experienced some difficulty in arriving at a workable model of the aether, the problems associated with the seemingly *ad hoc* postulation of the displacement current proved to be much worse. Although the latter did figure in the later developments of the electromagnetic theory (it was designated as a specific quantity) Maxwell remained agnostic about any qualitative account that could be offered. Despite these difficulties if we support the argument for unification or consilience as evidence for a realism about unifying structures then we ought to view the displacement current (and the aether) as real. This is especially true of the displacement current since its designated quantity provided the mechanism required for the derivation of the electromagnetic theory of light. Given Friedman's model this role would provide the displacement current with the kind of inductive support that was necessary for considering it a real physical process. However, as we know from the preceeding discussion there was no <u>independent experimental</u> evidence for its existence. The evidence that did exist was considered "independent" to the extent that it was accumulated via the explanatory power of the displacement current in unifying electromagnetic and optical processes. Hence the independence was a function of the <u>theory's</u> application in two different domains rather than independent evidencw for particular entities that results from a supposedly theory netural basis like experiment. This scenario raises some difficult problems for realism and the unification argument. As we saw in chapter one, if unification is the criterion for determining whether a particular entity or structure ought to be considered real, then the introduction of these entities becomes justified by a kind of bootstrap methodology where the legitimacy of the entities results from their unifying power which in turn strengthens their place within the theory. And, moreover, if unification is a theory relative notion then reality itself becomes contingent on the ability of particular theories to unify disparate bodies of data. None of these practices necessarily guarantees any degree of independent experimental evidence. Rather, there is a sense in which the theory generates its own evidence insofar as it determines what can be unified which in turn determines what is to count as real.

This problem of independent evidence was particularly crucial for the acceptance of Maxwell's electromagnetic theory. Despite the rather spectacular unification it achieved the unifying process failed to provide the basis for a realism about theoretical structures. The theory in its final form was in every sense a mathematization of the phenomena, with dynamical analogies and models being used to provide some understanding of what the possible physical interpretations that underlie the formalism may be like. The Lagrangian method, within which the theory was cast, provided the kind of unified approach that had been previously achieved through the use of rather problematic mechanical models. In the final analysis realism about the physical interpretation of the mathematical formalism was not required for the unification. In fact, the very structures that initially facilitated this extraordinary unification turned out to be the reasons for the definite lack of enthusiasm with which the theory was received. Hence Maxwell's agnosticism toward his physical models seemed an appropriate response.

One is led to conclude then that it is not the ability to unify that renders or gives us good reasons to think that particular physical entities are real; we can often provide a mathematically unified account in the face of problematic physical models. Some further condition is required. In the case of Maxwell's theory this condition was the existence of independent experimental evidence. It was not until Hertz's experiments on electromagnetic waves that Maxwell's theory was finally vindicated. Although the unification, the use of dynamical and physical analogies as well as various models of the theory played a major role in its discovery and development its <u>justification</u> lay in the experimental findings alone.

Footnotes

1. An argument of this form may have crucial implications for the spacetime case since the entities postulated by the theory do not seem to have a role that outstrips their place within the theoretical context. That is, there is no theory-neutral warrent for an ontology of space time points independently of their mathematical status in the theory. Unfortunately a discussion of the spacetime case would take us too far afield but I think it suffices to say that if a case can be made against the stronger claim (of unification as a criterion for realism in the context of a relatively straightforward physical theory) then it would apply _mutatis mutandis_ to the weaker case of spacetime theories. I would like to thank Philip Catton for discussion of this point.

2. Maxwell's view of models is particularly close, in many respects, to van Fraassen's understanding of the role of models as presented in his [1980]. Although some of Maxwell's models (unlike van Fraassen's) are purely instrumental in nature others he sees as giving a possible interpretation of the phenomena without commitment to their truth or reality.

3. Cf. Michael Faraday, _Experimental Researches in Electricity_ 3 Vols., London, 1839-1855. See especially Vol.1, para. 61.

4. _Ibid_. para. 72, 73.

5. Cf. _Experimental Researches_, "Electrostatic Induction"

6. _Experimental Researches_, 1, para. 1304. Faraday's concept of a line of force enabled him to provide a formulation of Ohm's law which states that the current produced will depend only on the facility of the cutting wire for conducting this force. This current will them depend directly on the electromotive force and vary inversely upon the resistance of the wire. For a discussion of Faraday's formulation of Ohm's law see Pearce Williams, _Origins of Field Theory_ New York: University Press of America, 1970.

7. _Experimental Researches_ 1, para.1231.

8. Faraday argued that by contiguous particles he did not mean paricles that touched one another but merely neighbouring particles. Hence there did exist a space between these particles, a space that seemed to indicate that some action at a distance must take place in order for the particles to transmit forces. Faraday's theory denied action at what he called "sensible distances" but the affirmation of action across "insensible distances" failed to offer any solution to Faraday's problem. The difficulties with Faraday's account were discussed by Robert Hare in an article entitled "A letter to Prof. Faraday, on certain Theoretical Opinions" _Philosophical Magazine_, 1840,

Vol. 27, p.45. A reply by Faraday can be found in _Experimental Researches_, 2, 252. For an account of the debate see P.M. Heimann, "Maxwell and the Modes of Consistent Representation", _Archive for the History of Exact Sciences_, Vol.6, 1970, pp.171-213; Mary Hesse, _Forces and Fields_, New Jersey: Littlefield, Adams & Co. 1961, Chapter VIII.

9. _Experimental Researches_ 3, para. 447.

10. _Experimental Researches_, 3, 447-452, especially pp..449-50. See also Barbara Gusti Doran, "Field Theory in 19th-Century Britain, _Historical Studies in the Physical Sciences_ Vol.1, pp. 133-260. For a discussion of Faraday's experiments on magneto-optic rotation as well as para and diamagnetism see Pearce Williams, op. cit., Chapter IV.

11. Having discovered that light was affected by the magnetic lines of force Faraday was determined to show that all matter must show some reaction to these magnetic lines. This was not a new hypothesis; Couloumb had proposed that all bodies were magnetic like iron only to a lesser degree. Although Faraday essentailly agreed that magnetic force must have some effect on all matter he disagreed that it was the same effect exhibited by iron. After testing several substances Faraday found that none reacted indifferently to the magnet. Everything acted either like iron, setting along magnetic curves, or like glass setting across magnetic curves. Those that behaved like iron Faraday called paramagnetic and those that behaved like glass were termed diamagnetics. Several theories explaining diamagnetism were put forth [Cf. Williams, _op. cit_. pp.96-110] but it was Faraday who finally explained the phenomenon. He had concluded that diamagnetism could not be explained by the forces of the particles of diamagnets and hence by diamagnetic polarity; instead the explanation was to be found in the magnetic line of force itself. Iron and paramagnetics conduct the lines of force very well. Hence in a magnetic field the lines of force tend to converge on the iron intensifying the field within the iron itself by concentrating the lines of force. Because diamagnetics are poor conductors of the lines of force the lines tend to spread out around the diamagnetic. So, if a diamagnetic were placed in a medium more diamagnetic than it, then the lines of force would converge on it and it would act like a paramagnetic substance. Similarily if a paramagnetic substance was placed in a medium more paramagnetic than it, it would act like a diamagnetic.

Modern physics tells us that diamagnetism is caused by the motion of electrons in atoms around the nuclei. An orbiting electron produces a magnetic field in the same way as an electric current flowing in a coil of wire. If an external magnetic field is applied, the electrons change their orbits and velocities so as to produce a magnetic field that opposes the applied field. If a diamagnetic substance is placed in a non-uniform field it tends to move from the stronger to the weaker part of the field. Diamagnetism is a very weak effect and is sometimes totally masked by paramagnetism or ferromagnetism.

12. Experimental Researches, para. 2640- 2701.

13. Faraday pointed out that the electrostatic lines of force would terminate on charges with opposite charges being at opposite ends of each line. This idea was not however associated with the polarization of molecules of matter. He emphasized that no condition of quality or polarity had yet been discovered in the lines of electrostatic force. So, even though electrostatic lines of force had charged ends there was no polarization with in the lines. This was at variance with his abandoned theory of molecular polarization where lines of force represented polarized particles. In his new theory lines of force seem to represent a physical state or are at least associated with a state of the medium (dielectric). See Experimental Researches 3, para. 3249.

14. For details of these experiments see Williams, op. cit. pp. 111-12.

15. See footnote 13 for an explanation of these experiments.

16. Also in view of Faraday's earlier speculations about light rays being vibrations of the lines of force the aether was no longer needed as the carrier of light.

17. Cf. ER, 3, para 3273.

18. In Faraday's field theory force could not exist without a medium and the force was to be found in the medium, not in the body from which it originated. The magnetic lines of force were propogated in empty space, a space that also transmited gravitational force. It was the emergence of the principle of conservation of energy in the 1840's that strengthened Faraday's attack on action at a distance. Action at a distance supposedly led to the creation of force (or energy) which was in opposition to the conservation principle. If body a existed in an otherwise uninhabited universe then according to action at a distance theorists it would have no force associated with it. If a second body b appeared a would attract it, resulting in a being suddenly endowed with force; force that is seemingly created out of nothing. On a field theory a sets up gravitational lines of force (a strain) in space. B then simply reacts to this strain; there is no creation of force. The energy of the system is in the medium; the bodies merely detect these strains and react to them. Consequently the action is not at a distance but rather where the body is. Cf. Pearce Williams, op. cit.

19. See Silvanus P. Thompson, The Life of William Thomson, Baron Kelvin of Largs, 2 Vols., London, 1910 (cf. Vol. 1, pp.146-49). The original paper can be found in William Thomson, Reprint of Papers on Electrostatics and Magnetism, London, 1872, pp.1-14. It is interesting to note that Thomson published the paper under the name P.Q.R.

20. The analogy was described in a letter to Faraday dated August 6, 1845. The text of the letter is reprinted in Thompson, op. cit., 1, pp.146-49. See also pp. 143-44, 210-16 and 134. For a nice discussion of the importance of the potential fuction see Doran, op. cit.

21. Thomson's Papers give us a good indication about his views on models and their role in the construction of physical theory. See also P.M. Harman, Energy Force and Matter Cambridge: Cambridge University Press, 1982.

22. See Thomson, Papers, op. cit., 1, pp. 76-80.

23. Cf. Harman, op. cit.

24. See S. Thompson, Lord Kelvin, op. cit. pp.203-4.

25. Papers, pp.340-405

26. The potential function was a kind of physico-geometrical notion used by Green and later by Thomson to show that Couloumb's electrical action at a distance and Faraday's notion of action by contiguous particles in the intervening medium lead to the same mathematical theory. Indeed it was an important feature in formulating an alternative to action at a distance accounts. As a mathematical apparatus it could be applied to continuous media as well as discrete mass points. As a result, using the potential function one could explain forces between mass points in terms of continuous action. For a detailed explanation see Doran, op. cit. p.167 n.88.

27. See Doran, op. cit. for details, especially pp. 174-75.

28. Papers, p.341.

29. William Thomson, "Dynamical Illustrations of the Magnetic and Helidoidal Rotary Effects of Transparent Bodies on Polarized Light", Proceeding of the Royal Society, 8 (1856).

30. By using the concept of molecular motion developed in his dynamical theory of heat as well as Rankine's theory of heat as a vortical motion of aethereal atmospheres surrounding molecular neuclei, Thomson argued that the magneto-optic rotation discovered by Faraday could be explained as the result of a vortical motion of the aether. Thomson did not adopt Rankine's

account of molecular vortices and chose to remain agnostic about the specific physical details of his account. For a discussion of Rankine's work and its influence on Thomson and Maxwell see W.J.M. Rankine, Miscellaneous Scientific Papers, W.J. Millar (ed.) London, 1881, pp. 16-48; 150-155; 234-281. See also, Doran, op. cit.; and D.F. Moyer, "Energy, Dynamics and Hidden Machinary: Rankine, Thomson and Tait, Maxwell"; Studies in the History and Philosophy of Science, vol. 8, 1977, pp.251-268.

31. Maxwell, Papers, 1, p.160.

32. See "On the Uniform Motion of Heat in Homogenous Solid Bodies and its Connection with the Mathematical Theory of Electricity" printed in Reprint, op. cit. pp.1-14.

33. Papers, 1, p.157.

34. Ibid.

35. The chief proponent of action at a distance theories was Wilhelm Weber who in the 1840's developed a reduction of the laws of electricity and magnetism to consequences of the laws of mechanics. Weber's theory was based on the interaction of charged particles through distance forces depending on their relative positions, velocities and accelerations. This together with his Amperean theory of molecular currents enabled him to provide a unified account of electricity, magnetism, electromagnetism and electromagnetic induction. Maxwell disliked Weber's theory due to its many startling assumptions and despite its many accurate results considered it only a mathematical speculation that ought to be compared to other accounts. Cf. Maxwell's letters to Thomson May 15th and Sept. 13th 1855, printed in Larmour, op. cit. Maxwell felt that his analogy, though not nearly as complete as Weber's theory, could nevertheless be justified on the basis of its mathematical equivalence to the action at a distance formulation. "It is a good thing to have two ways of looking at a subject, and to admit that there are two ways of looking at it". Papers, 1, pp.207-8.

36. The model provided a kind of mathematical isomorphism between the equations, of percolative streamline flow and the equations describing electric or magnetic lines of force.

37. Maxwell Library, Cambridge, Mss. 7655 as quoted in P. Heimann "Maxwell and the Modes of Consistent Representation" Archive for the History of the Exact Sciences, Vol. 6, 1970.

38. Cf. Heimann, note 84.

39. In a letter to Thomson (Sept. 13/1855) Maxwell remarked that he was setting down his theory for the sake of acquiring knowledge sufficient to guide him in devising experiments. See Origins of Clerk Maxwell's Electrical Ideas, J. Larmour (ed.) Cambridge (1937) p.18.

40. Papers, 1, pp.155-56.

41. Ibid.

42. Ibid.

43. Ibid. p.205.

44. Letter to Thomson, Sept. 13/1855 in Larmour, op. cit.

45. Papers, 1.p.203.

46. For details see Pearce Williams, op. cit.

47. Cf. Daniel Seigel, "Maxwell's Electromagnetic Theory" in P.M. Harman (ed.) Wranglers and Physicists Manchester: Manchester U. Press (1985).

48. "On Physical Lines of Force" in Papers, 1, pp. 451-513.

49. No tendancies of the magnetic lines to repel each other and contract along their lengths were derivable from the flow analogy; hence, it could not explain magnetic forces. By assuming that the magnetic line of force represented the axis of a molecular vortex, the vortex theory was capable of explaining these phenomena without difficulty. It was possible to demonstrate that centrifugal forces caused by the rotational motion of the molecular vorticies would tend to make each vortex tube expand in thickness, thereby tending to increase the spacing between magnetic lines. Moreover, because of the incompressibility of the fluid in the vortex tubes theory would tend to shrink in length causing the magnetic lines to have a corresponding tendancy to contract along their lengths. Cf. PL, p.467 as well as Seigel, op. cit. pp.186-87.

50. That Maxwell felt less than convinced about the strength of his position is evident from his remark on p. 488 of PL:

> Those who look in a different direction for the explanation of the facts may be able to compare this theory with that of the existence of currents flowing freely through bodies, and with that which supposes electricity to act at a distance... The facts of electromagnetism are so complicated and various, that the explanation of any number of them by several different hypotheses must

> be interesting...to all those who desire to understand how much evidence the explanation of the phenomena lends to the credibility of the theory...

51. Maxwell describes his model in the following way:' The electromagnetic medium is divided into cells spearated by partitions formed of a stratum of particles which play the part of electricity. When the electric particles are urged in any direction they will, by their tangential action on the elastic substance of the cells, distort each cell and call into play an equal and opposite force arising from the elasticity of the cells. When the force is removed the cells will recover their form and the electricity will return to its former position [PL, p.492]. Displacement satisfies the phenomenological law $R = -4\pi E^2 h$ with h being the electric displacement per unit volume. R is the electromotive force on an elastic sphere in equilibrium and the constant E is the coefficient depending on the nature of the dielectric (elasticity or rigidity). It is possible to interpret E in this way because the theory of elasticity yields equation relating force and strain in terms of the elastic parameters. h is the displacement or the sphere's distortion if r is the value of the electric current due to displacement, $r = dh/dt$. In calculating h we integrate over the surface of a single spherical cell (δS) of volume V.

$$h = \int_S \delta S \; 1/2 \; pt \sin\theta / V.$$

Not only is displacement introduced in terms of a model for polarization, it is inserted into an experimental equation expressing the relation between polarization and force without any theory about the internal mechanism of dielectrics. In addition it receives a mathematical definition which establishes it as an electric dipole moment per unit volume. Cf. Joan Bromberg, "Maxwell's Electrostatics" <u>Americal Journal of Physics</u> (36) 1968 as well as A.F. Chalmers, "Maxwell's Methodology and his Application of it to Electromagnetism" <u>Studies in the History and Philosophy of Science</u> Vol.4, No.2 pp.129-31.

52. Cf. Bromberg, <u>op. cit.</u> as well as "Maxwell's Displacement Current and his Theory of Light" <u>Archive for the History of the Exact Sciences</u>, 4, 1968.

53. Cf. Bromberg's article in <u>AHES.</u>

54. Cf. Bromberg, <u>Ibid</u>. p.226. Earlier in the paper Maxwell relates to the magnetic properties of the medium by $Q = \mu/\pi$ where is magnetic permeability. Hence $V = \sqrt{\pi m/\mu} = \sqrt{E^2/\mu}$ and in air $\mu = 1$ and E = 310,740,000,000 mm/sec.

55. In an earlier paper Weber and Kohlrausch (1857) determined the ratio between the electroststic and electrodynamic units of charge, a ratio having the dimensions of the velocity of propogation of electric action. Weber's constant referred to

electrodynamic rather than electromagnetic units and so his ratio came out as $\sqrt{2}$ x the velocity of light. In a letter to Thomson Dec. 10/1861 Maxwell stated what he had previously written to Faraday, that he was unaware of Weber's result and that he had made out the equations before he had "any suspicion of the nearness between the two values of the velocity of propogation of magnetic effects and that of light".

56. Duhem was the first to point this out in his Les Theories Electriques de J.C. Maxwell Paris: 1902 pp.211-12. Maxwell's definition of m is as follows: Let ξ, λ, ζ be the displacements of any particle of the sphere in the directions of x, y, and z. Let P_{xx}, P_{yy}, and P_{zz} be the stresses normal to planes perpendicular to the three axes and let P_{yz}, P_{zx}, and P_{xy} be the stresses of the distortion in the planes yz, zx, and xy. m can now be defined as $P_{xx} - P_{yy} = m(d\xi/dx - d\lambda/dy)$, & c. Cf. RL, p.493. Duhem concluded that Maxwell was guilty of conciously falsifying one of the fundemental formulae of elasticity, implying that Maxwell had constructed his model in order to justify an electromagnetic theory of light that he already knew to be in error. However, as Chalmers op. cit. points out, Maxwell cannot be justly accused of defining a coefficient of rigidity that differed from the customary one in order to conceal a faulty derivation because he had defined the constant in precisely the same way twelve years earlier. See Maxwell's "On the Equilibrium of Elastic Solids" reprinted in Papers, 1; pp.30-73.

57. Cf. Bromberg AHES op. cit.

58. Cf. Siegel, op. cit.

59. Cf. Sir Edmund Whittaker, A History of the Theories of the Aether and Electricity. New York: Thomas Nelson and Sons (1951) Ch. VIII. For a more detailed mathematical treatment of the problems with the displacement current see Alfred O'Rahilly, Electromagnetic Theory, New York: Dover (1965) Vol.I, Ch.III.

60. There are numerous reasons discussed in the literature for this move. See for example, Bromberg, op. cit.; A. Bork, "Maxwell, Displacement Current and Symmetry" Americal Journal of Physics, (31) 1963; D. Siegel, "Completeness as a Goal in Maxwell's Electromagnetic Theory", Isis, (66) 1975.

61. According to earlier theories a current employed in charging a condenser was not closed; instead it terminates at the coatings of the condenser where the charge accumulated.

62. My discussion follows Bromberg, op. cit. An alternative account if offered by Chalmers op. cit. who argues that Bromberg is wrong in attributing an inconsistency to Maxwell. He claims that her interpretation arises from a failure to appreciate the distinction Maxwell makes between electromotive force and

electric tension. However, for what seem to me to be obvious reasons, namely, the difficulty involved in giving an <u>electro-mechanical</u> theory of light, Maxwell seems guilty of interpreting $R = -4\pi E^2 h$ as both an electrical equation and an equation representing elasticity. Once the mechanical apparatus of Maxwell's model has been abandoned (in DT and <u>TEM</u>) the negative sign in the equation disappears.

63. $E = -4\pi E^2 D$ is vector notation for equation (105)PL $R = -4\pi E^2 h$ where R is the 2-component of electromotive force (i.e. electric intensity), h is the 2-component displacement and E is a constant varying with the nature of the dielectric.

64. Equation (9) in PL gives this quantity. $\frac{1}{4\pi}\left(\frac{d\beta}{dx} - \frac{d\alpha}{dy}\right)$ represents the strength of an electric current parallel to z through unit of area, and if we write $p = \frac{1}{4\pi}\left(\frac{d\gamma}{dy} - \frac{d\beta}{dz}\right)$; $q = \frac{1}{4\pi}\left(\frac{d\alpha}{dz} - \frac{d\gamma}{dx}\right)$; $r = \frac{1}{4\pi}\left(\frac{d\beta}{dx} - \frac{d\alpha}{dy}\right)$ then p,q,r will be the quantity of electric current per unit area perpendicular to the axes of x, y and z respectively. In order to correct this equation of electric currents for the effect due to the elasticity of the medium we differentiate $R = -4\pi E^2 h$ (the equation connecting electromotive force with electric displacement) with respect to t giving us $dR/dt = -4\pi E^2\, dh/dt$, which shows that when electromotive force varies so does displacement. Because variation of displacement is equal to a current this current must be taken into account and added to r yielding the equations (112) $p = \frac{1}{4\pi}\left(\frac{d\gamma}{dy} - \frac{d\beta}{dz}\right) - \frac{1}{E^2}\frac{dP}{dt}$; $q = \frac{1}{4\pi}\left(\frac{d\alpha}{dz} - \frac{d\gamma}{dx}\right) - \frac{1}{E^2}\frac{dQ}{dt}$; $r = \frac{1}{4\pi}\left(\frac{d\beta}{dx} - \frac{d\alpha}{dy}\right) - \frac{1}{E^2}\, dR/dt$.

65. It is natural to assume that the equation (112) represent the addition of the displacement current (dh/dt) to the ordinary current $r = \frac{1}{4\pi}\left(\frac{d\beta}{dx} - \frac{d\alpha}{dy}\right)$ ... is the same as $r = \frac{1}{4\pi}\left(\frac{d\beta}{dx} - \frac{d\alpha}{dy}\right) + \frac{dh}{dt}$ and differentiating (112) with respect to x, y, and z respectively and substituting we get $\frac{de}{dt} = \frac{1}{4\pi E^2}\frac{d}{dt}\left(\frac{dP}{dx} + \frac{dQ}{dy} + \frac{dR}{dz}\right)$ (114); hence $e = \frac{1}{4\pi E^2}\left(\frac{dP}{dx} + \frac{dQ}{dy} + \frac{dR}{dz}\right)$ (115), the constant being omitted because e = o when there are no electromotive forces.) yielding $p = \frac{1}{4\pi}\left(\frac{d\gamma}{dy} - \frac{d\beta}{dz} - \frac{1}{E^2}\frac{dP}{dt}\right)$; $q = \frac{1}{4\pi}\left(\frac{d\alpha}{dz} - \frac{d\gamma}{dx} - \frac{1}{E^2}\frac{dQ}{dt}\right)$; $r = \frac{1}{4\pi}\left(\frac{d\beta}{dx} - \frac{d\alpha}{dy} - \frac{1}{E^2}\frac{dR}{dt}\right)$.

In parts I and II of PL (pqr) is used to represent conduction current density; however, in part III Maxwell uses r to denote only the z component of the displacement current [p.491] and states that displacement current must be added to r in order to get pqr ... the electric currents in the directions of x,y, and z [p.496]. Similarily in (112) α, β, γ are designated as the components of magnetic intensity and P,Q,R are the electromotive forces. The ambiguity surrounding the interpretation of pqr arises from Maxwell's formulation of what he calls the equation of continuity (113). If e is a quantity of free electricity in unit of volume then $\frac{dp}{dx} + \frac{dq}{dy} + \frac{dr}{dz} + \frac{de}{dt} = 0$ is the equation of continuity. However, (113) does not relate the total current to the volume density of electricity as one might expect but instead relates the conduction current: a result that implies conduction alone give rise to magnetic effects. On the other hand if pqr is interpreted as total current it implies that total currents are not closed; a result that violates on of the structural presuppositions of Maxwell's model. The requirement that all currents are closed is necessary is electricity is to obey the same condition as an incompressible fluid.

66. Cf. Papers, PL, p.486. See also Siegel in Cantor and Hodge (eds.) Conceptions of the Aether n.51.

67. Campbell and Garnet, p.340.

68. Memoir and Scientific Correspondence of the Late Sir George Gabriel Stokes J. Larmour (ed.) Cambridge: CUP (1907) Vol. II, p.26.

69. This generalized dynamical approach was also used by Thomson and Tait in their Treatise on Natural Philosophy (1867) where they integrated the energy approach with the Lagrangian formalism.

70. See Whittaker op. cit. for discussion.

71. I will say more about the specifics of the Lagrangian method in Section VI of the paper. The calculus of variations deals with the mathematical problem of minimizing an integral, allowing particular results to be established without taking into account the infinity of tentatively possible paths. We do this by restricting our mathematical experiment to paths (variations of the actual path) that are infinitely near to the actual path. Most problems in physics can be solved by vectorial methods but in more complicated situations the variational approach is superior. Because the variational approach satisfies the principle of invariance for all natural phenomena it allows us complete freedom in choosing the appropriate coordinates for our problem; our processes and resulting equations remain valid for an arbitrary choice of coordinates. The philosophical and mathematical value of the variational method is the freedom of choice and of arbitrary coordinate transformations, an option that greatly facilitates the formulation of and solutions to the differential equations of motion. If we hit on a certain type of coordinates called "ignorable" coordinates then a partial integration of the basic differential equations is at one accomplished; and, if all our coordinates are ignorable then the problem is completely solved. Hence the problem of sloving differential equations can be seen as a problem of coordinate transformation. Rather than integrating the differential equations of motion directly we try to produce and increasing number of "ignorable" coordinates. There was no systematic way to do this in Lagrangian mechanics but the later developments by Hamilton and Jacobi introduced cannoncial equations that had much wider transformation properties. As a result we could produce a complete set of ignorable coordinates by solving one single

partial differential equation. It is perhaps important to point out that the basic difference between the variational and vectorial methods is that the latter (Newton's approach) does not restric the nature of a force while the former assumes that the acting forces are derivable from the "work function", a scalar quantity. Hence because frictional forces have no work function they lay outside the realm of variational principles. For a good treatment of this topic see C. Lanczos, The Variational Principles of Mechanics, Toronto: University of Toronto Press (1970).

72. W. Thomson, "On the Possible Density of the Luminiferous Medium, and on the Mechanical Value of a Cubic Mile of Sunlight" Transactions of the Royal Society of Edinburgh (1854) p.57.

73. These experimental facts include: 1. The induction of electric currents by the increase or diminuation of neighbouring currents according to the changes in the lines of force passing through the circuit.
2. The distribution of magnetic intensity according to the variations of a magnetic potential.
3. The induction (or influence) of statical electricity through dielectrics.

74. For details see F. Everitt, James Clerk Maxwell Physicist and Natural Philosopher, p. 103.

75. This mechanical analogy would be spelled out in a more formal way using the Lagrangian mechanics of the Treatise.

76. (p,q,r) was to be understood as conduction current density. The general equation (A) [p.554] states that the variations of the electrical displacement must be added to the currents p,q,r to get the total motion of electricity, which we may call p', q', r' so that p' = p+df/dt; q' = q+dg/dt; r' = r+dh/dt. f,g,h denoted the electric displacements parallel to x,y,z respectively. Contrast this with Maxwell's rather ambiguous interpretation p,q,r discussed earlier. This equation (A) together with the equation of currents(C) $d\gamma/dy - d\beta/dz = 4\pi p'$; $d\alpha/dz - d\gamma/dx = 4\pi q'$; $d\beta/dx - d\alpha/dy = 4\pi R'$ p.557 implies that the total current is circutal and that a displacement current gives rise to a magnetic field. By contrast in PL some aspects of Maxwell's model implied that it was the conduction current alone that gave rise to magnetic effects. For a detailed account (presented in vector notation) of the relationship between conduction and total currents and the difficulties raised by Maxwell's model see Bromberg, op. cit.

77. In DT Maxwell replaces $4\pi E^2$ with K such that R = Kh. The value of E is derived from Maxwell's new electrostatic equation divD = -e.

78. This liberation from mechanical influences allowed the law to figure nicely with the generalized form of Ampere's law.

79. Not only did Maxwell wish to identify charge with a discontinuity in displacement in a dielectric medium but he also insisted that displacement involved a movement of electricity "in the same sense as the transferrence of a definitequantity of electricity through a wire is a movement of electricity".

80. Cf. Bromberg, <u>AJP</u>, <u>op</u>. <u>cit</u>.

81. The method of dynamical explanation that Maxwell developed in the treatise was one that had been used by Thomson and Tait in the <u>Treatise on Natural Philosophy</u>. Thomson and Tait's method emphasized dynamical concepts as opposed to the more straightforward Lagrangian method that provided a purely mathematical formalism. In the latter case reference to concepts like momentum, velocity and energy was avoided by replacing them with symbols in the generalized equations of motion. The dynamics of Thomson and Tait was based on a theorem relating the variation in a system due to implusive forces acting in an infinitesimal time increment to the kinetic energy of the system. This facilitated the derivation of a generalized equation of motion.

82. Cf. "Thomson and Tait's Natural Philosophy" <u>Papers</u>, 2, pp.776-785, n.58.

83. We can briefly describe some features of Lagrangian mechanics in the following way: The generalized coordinates x_1, x_2, ... of a system are a set of parameters which, in conjunction with their generalized velocities (time derivatives) $\dot{x}_1$, $\dot{x}_2$, ..., specify the state of the system. The number of generalized coordinates equal the number of degrees of freedom of the system. If T and V represent the kinetic and potential energy then the Lagrange equations of motion can be written $d/dt\,(\delta T/\delta \dot{x}_i) - \delta T/\delta x_i = -\delta V/\delta x_i$ where $-\delta V/\delta x_i$ is the generalized force Q_i. If the functions T and V are known, the equations of motion enable the time development of a system to be deduced from knowledge of the generalized coordinates and velocities at some instant. My summary is taken from Chalmers <u>op</u>. <u>cit</u>.

84. Maxwell defined polarization in terms of equal and opposite charges at opposite ends of a particle; as a forced state of the medium. In keeping with PL dielectric polarization was represented by electric displacement with variations constituting electric currents. According to the account provided in <u>TEM</u> particles of a magnet were polarized and the theory of magnetic polarization was developed in a manner analagous to the theory of dielectric polarization. The difference between PL and <u>TEM</u> was that in the latter context displacement was defined by the motion of electricity, a quantity of charge crossing a specified area,

1. Nearly all physical laws can be derived from variational principles from which differential equations mathematically identical to Lagrange's are deduced. The most straightforward changes in representation are produced by coordinate (point) transformations $Q_i = Q_i(q_i, t)$. In Lagrangian formulations every transformation of generalized coordinates and momenta $[Q_i = Q_i(q_i, p_i, t),\ P_i = P_i(q_i, p_i, t)]$ generates the corresponding representation. Conditions of invariance are also imposed (given certain specifications about physical equivalence and significance). The representation of a physical theory is really then a study of its transformation properties with the various representations pictured as many different mappings of the same object. For a detailed discussion of this topic see M. Bunge, "Lagrangian Formalism and Mechanical Interpretation" in American Journal of Physics, Vol.25, 4, (1957).

2. The formalism provides us not only with a way of expressing a dynamical system but also with a method for framing equations of motion. In fact it can be applied in quantum field theory in connection witn elementary particles.

3. P_1 & P_2 are points in a medium of known refractive index $n(x,y,z)$. Using Fermat's principle we claculate $A = \int_{P_1}^{P_2} n(x_i)ds$. Using the methods of Lagrange we introduce a physically uninterpreted parameter t that figures in the formula for optical length, i.e. $A = \int_{t_1}^{t_2} L dt$ where $L(x_i, \dot{x}_i) = n(x_i) \cdot (\sum \dot{x}_i^2)^{1/2}$.

4. I realize of course that I am ignoring one of the most widely used methods in establishing conclusions based on inductive evidence. Scientific realists want to argue that if we can successfully extend the analogy into new domains and provided the comparisons are relevant and close enough then the analogical argument should give us good reasons for accepting the conclusions we draw from it. Wesley Salmon presents an argument of this kind in his Causal Explanation and the Causal Structure of the World, Princeton (1984). He argues that the move from observables to unobservables can be sanctioned by analogical argument. The argument is based on the induction that if all effects $E_1 \ldots E_n$ have been the result of cause $C_1 \ldots C_n$ then we can infer that E_{n+1} is the result of C_{n+1}. Salmon cites some additional evidence for the legitimacy of analogical arguments including the similarities between human and animal research as well as the analogy between Newton's inverse square law and Couloumb's theory of electrostatics.

Maxwell's examples in FL and PL characterize an analogy between two different branches of physics by showing how the same mathematical formalism can be applied to each. Similarily Thomson showed that there was an analogy between the theory of heat and electrostatics on the basis that both can be described by the same equation if one reads "temperature" for "potential" and "source of heat" for "positive electric charge".

Consequently the theory of heat was used as a model for the field theory of electrostatics. But, because we don't know the mathematical structure of nature and because nature can be described in a variety of mathematical ways the most we can say is that there is a resemblance between the numerical consequences of our experimental outcomes and the numerical features of the model. In the case of atomic particles we employ a model based on dynamics and electrostatics because deductions from this type of model yields numbers comparable to experimental measurements. Salmon points out that our model may provide a better analogy than anything else with which we are acquainted and on that basis can lend plausibility to the conclusions of our inductive inferences; but plausibility (being such a loosely construed notion) is surely not enough when the issue is evidential support.

According to Maxwell the method for recognizing real analogies rests on experimental identification. (I shall discuss this in Chapter 3.) If two apparently distinct properties of different systems are interchangeable in appropriately different physical contexts they can be considered the same property. This is a species of what Hesse calls substantial identification, which exists when the same entity is found to be involved in apparently different systems. An example would be identification in PL of the aetherial media of electromagnetism and light on the basis of the numerical value of the velocities of transmission of transverse electromagnetic waves and of light. This kind of analogical argument plays an important role in consilience, but, ~~as I~~ argued in chapter 1 the latter cannot be easily accommodated within realist epistemology. For more discussion on analogies see M.Hesse "Logic of Discovery in Maxwell's Electromagnetic Theory" in Giere and Westfall (eds.) Foundations of Scientific Method Bloomington: Indiana U.Press.

5. Cf. Papers, 2, p.781.

6. I will have more to say about bootstrap methodologies in chapter three where I discuss experimentation and its role in providing independent evidence.

rather than in terms of the rolling particle motion.

85. Cf. Sec. 11 & 12.

86. Cf. Doran, op. cit. and Alfred Bork, "Maxwell and the Vector Potential" Isis, 58 (1967) pp. 210-22.

87. Cf. TEM sec. 809. Maxwell had discussed Verdet's result in 1862 in the fourth part of PL as well as in a letter to Thomson Dec. 1861. Maxwell aregued that "we must admit the diamagnetic state to be the opposite of the paramagnetic" because the vorticies "revolve in the opposite direction". Although this result agreed with Weber's theory that paramagentic and diamagnetic phenomena were due to opposite states, Maxwell maintained that this did not require us to admit Weber's theory of molecular electric currents; a theory that involved action at a distance explained by the mutual of electric particles in motion. It did however require a break with Faraday's theory of magnetism that he adopted in part I of PL; a view that endorsed the idea of lines of force instead of molecular polarization. Maxwell adopted this theory of polarization in part III of PL according to which the molecules of the dielectric are said to be polarized. Maxwell retained this idea throughout DT and TEM together with the significance of Verdet's experiments on the opposite rotation of polarized light be para and diamagneitc substances.

88. Cf. Chalmers op. cit.

89. Cf. Siegel, Bromberg and others.

90. Quoted in Campbell and Garnett, op. cit.

Chapter III

Experiment and Theory: Hertz and the Electromagnetic Theory

<u>1. Introduction.</u> My aim in this chapter is to provide an analysis of the relationship between Maxwell's theory and the experiments that led to its confirmation. The relationship between theory and experiment or theory and evidence has attracted considerable debate in philosophy of science, or for that matter in all of philosophy. Questions concerning the nature of evidence and the degree to which a certain piece of evidence supports or justifies the claim that a theory or belief is true is the focus of epistemological concerns about science and other areas of human inquiry. Up to this point I have focused on aspects of theory acceptance, specifically consilience and the more general features of theory unification and reduction, that many realists see as providing evidence for the truth of the theory in question. I have attempted to counter these realist claims with philosophical argument supplemented by an historical case study, both of which show I believe the philosophical inadequacy of the realist position as well as its failure to explain actual historical cases. The examination of the development of Maxwell's electromagnetic theory revealed that <u>internal</u> aspects of the <u>theory</u> such as its ability to unify hitherto independent phenomena were of little importance in deciding its superiority over its rivals. As a methodological rule it would seem that we not only require

that the theory be able to accommodate experimental facts of various kinds but there must be some independent experimental evidence for the existence of the entities postulated by the theory. In the case of the electromagnetic theory further independent evidence confirming the reality of the displacement current was required to establish or secure its success within the theory as a unifying entity or structure.

In what follows I shall examine the relationship between theory and evidence or experiment especially as it pertains to the electromagnetic theory. One of the major experimental difficulties facing Maxwell's theory was its inability to explain the well-established phenomena of reflection and refraction. Although this was problematic it was the lack of independent evidence for displacement that proved to be perhaps the most serious obstacle for the theory. Despite Maxwell's spectacular success in unifying electromagnetic and optical phenomena it was Hertz's reformulation of the theory, his omission of the displacement current and the experimental detection of electromagnetic waves that were the crucial factors resulting in the final acceptance of Maxwell's theory.

It is also important to recognize that the way in which empirical evidence bears on theory is not always straightforward, as Hertz's experiments show. My discussion of these experiments and their relationship to the electromagnetic theory will reveal some of the tinkering' and looseness of fit that characterizes the theory-experiment relationship.

Very briefly then the focus of this chapter is an examination of the issues involved in the confirmation of Maxwell's theory as they pertain to the more general questions about the relationship between theory and experiment. Section two provides a summary of some of the experimental difficulties encountered by Maxwell's theory. I also discuss some experiments involving electromagnetic phenomena prior to Hertz's work. Because these experiments were considered insignificant even by those acquainted with Maxwell's theory they reveal the importance of the theoretical context in the assimilation of experimental findings. In sections three and four I discuss Hertz's experiments and his theoretical interpretation of Maxwell's account of electromagnetism. I conclude with an analysis of Hertz's results, what aspects of theory they could be said to confirm and how Hertz himself viewed the relationship between theory and evidence.

2. Experimental Problems of the Electromagnetic Theory.

Although the status of the displacement current raised severe difficulties for Maxwell's theory there were other experimental problems that may have figured in its less than enthusiastic reception. The goal of TEM was to inquire into the nature of the propagation of electromagnetic effects. Maxwell remarked that because we are unable to conceive of propagation in time except as either the flight of a material substance through space or as the propagation of a condition of motion or stress in a medium already existing in space, a great many theories:

> ...lead to the conception of a medium in which propagation takes place, and if we admit this medium as an hypothesis, I think it ought to occupy a prominent place in our investigations, and that we ought to endeavour to construct a <u>mental representation</u> of all the details of its action, and this has been my constant aim in this treatise [<u>TEM</u>, V.2, p.493].

Because Maxwell had no knowledge of the mechanism involved in the aether the best he could do was to formulate equations that described the relationships between electric and magnetic phenomena and construct dynamical analogies that would provide some physical interpretation of what the medium might be like. Once the value of an electrical constant had been established Maxwell concluded that the velocity of electromagnetic waves, within reasonable limits of experimental error, was equal to the velocity of light. Although the velocity of light waves had been measured experimentally the value for the velocity of Maxwell's electromagnetic waves emerged solely from his theoretical calculations. Maxwell did claim empirical support for this aspect of his theory but any such evidence seems to have been <u>indirect</u> at best.[6]

An additional problem concerns the transversality of the electromagnetic waves. If Maxwell's four basic electromagnetic equations[7] are correct then transverse electromagnetic waves corresponding to an acceptable mathematical solution to the equations ought to occur in nature. In "Dynamical Theory" and in the "Treatise" Maxwell struggled with the problem of obtaining a condition of transversality for wave propagation. His arguments attempted to show that the components of the vector potential were propagated as transverse waves: that div A = 0 (where A is the vector potential) was the condition for transversality. Unfortunately the arguments Maxwell used to justify his claim were presented in several different forms and the account given in DT was different from the one used in TEM.[4] In DT he claimed that the equations of the electromagnetic field showed that only transverse vibrations could be propagated that in optical science too we can only account for transverse vibrations. "Both sciences were at a loss when called upon to affirm or deny the existence of normal vibrations" [DT, p.582]. Stokes had called attention to this problem in his 1862 British Association Report[5] and remarked that any satisfactory theory of the optical aether would have to explain the propagation of only transverse vibrations. He further claimed that:

> That theory should point to the necessary existence of such a wave consisting of strictly normal [i.e. longitudinal] vibrations, and yet to which no known phenomenon can be referred, is bad enough; but in the present theory the vibrations are not even strictly normal, except for waves in a direction perpendicular to any one of the principal axes. [op. cit. p.256].

Similarily, in a reply to a paper of Green's[6] that suggested the suppression of longitudinal waves, Stokes commented that the only way of circumventing the difficulty would be to make the "perfectly gratuitous assumption that the medium, though perfectly transparent for the more nearly transversal vibrations, is intensely opaque for those more nearly normal" [Ibid. p.258]. In the final analysis Stokes found that no theory of the optical aether had successfully explained the absence of longitudinal waves.

It appears that Maxwell shared Stokes' concern about the problems of the optical aether, namely, the transversality condition and the phenomena of reflection and refraction. In a letter to Stokes in 1864[7] Maxwell expressed concern over this problem and concluded by remarking that the conditions at a surface for reflection and refraction may not be the same for the period of vibration of light and for experiments made at leisure.

In his discussion of Maxwell, Whittaker [op. cit.] remarks that it is somewhat strange that having successfully employed the electromagnetic theory to explain the propagation of light in isotropic media, in crystals and in metals, Maxwell should have omitted to apply it to the problem of reflection and refraction. This is even more surprising because the study of optics and crystals had already revealed a close analogy between the electromagnetic theory and MacCullagh's elastic solid theory. In order to explain reflection and refraction electromagnetically one simply had to interpret the time flux of the displacement of MacCullagh's aether ($\partial e/\partial t$) as the magnetic force and curl-as the electric displacement. In MacCullagh's theory the difference between the contiguous media could be represented by a difference in their elastic constants and so in the electromagnetic case it could be represented by a difference in their inductive capacities. It would appear however that there is a relatively straightforward reason why Maxwell neglected to use this approach in explaining relfection and refraction; a reason that has to do primarily with the difficulties involved in specifying a workable model of the aether.

Cauchy's[8] 1830 study of crystal optics had been the first application of the theory of ordinary elastic solids to the optical aether.[9] He attempted to deduce the principles of optics by taking the differential equations of motion of an elastic solid to be the equations representing the propagation of light. In addition he made some initial assumptions about whether the vibrations constituting light are parallel or at right angles to the plane of polarization and about the density of the aether in different material bodies. Cauchy's study revealed that the boundary conditions that yield optical laws were not those of an ordinary elastic solid. Hence the laws of optics and the elastic constants that he used proved extremely difficult to justify in any physical way.

In 1837 MacCullagh[10] used the opposite procedure by starting with principles of optics and using geometrical methods in attempting to derive the equations of motion for the medium in which Fresnel's formulas of reflection hold. Because his initial assumptions were similar to those made by Cauchy, MacCullagh's optical theory also contradicted the requirements of an ordinary elastic solid and hence failed to provide any justification for the elastic constants.[11] As a result MacCullagh was forced to admit that "the constitution of this aether and the laws of its connection with the particles of bodies are completely unknown."[12] A few years later in 1839 MacCullagh was able to account for Fresnel's laws by using a purely dynamical method introduced by Green. The method consisted of establishing the potential energy function required for the action of the medium and then deriving the laws of optics directly from it. In this newly formulated dynamical theory the elasticity and potential energy of MacCullagh's aether depended only on the rotational displacements of its elements from equilibrium resulting from a strain. Although the boundary conditions were different from those of an ordinary elastic solid this account proved to be insufficient as well. In 1862 Stokes objected that the vector used by MacCullagh to represent the light disturbance could not be be the displacement in an elastic medium that was similar to an elastic solid. Hence MacCullagh's vector was incompatible with the form deduced by Green. The stumbling block was clearly the lack of a mechanical (or any other) model that could illustrate the mode of action of a rotationally elastic aether.

It is not surprising then, given Maxwell's difficulty in accounting for the connection between the aether and matter and in providing a model of the aether that could specify the nature of its motion, that he was reluctant to apply his electromagnetic theory to reflection and refraction; especially if such an application were to involve MacCullagh's account. Although the final form of Maxwell's theory was independent of any specific model of the aether he did emphasize that there was evidence that some phenomenon of rotation seemed to be going on in the magnetic field [TEM 2:470 Sec.31]. Because Maxwell had linked the electromagnetic variables to a material interpretation of the aether (even though he remained agnostic about the specifics of that interpretation) he was unable to envisage the aether as merely a collection of electromagnetic properties. It was not until 1880 when Fitzgerald was successful in formulating a purely electromagnetic analogue of MacCullagh's theory incorporating Maxwell's electromagnetic theory of light; that reflection and refraction were successfully explained.[13]

Unfortunately for Maxwell there seemed to be little if any direct experimental evidence for his theory. As I noted above what Maxwell took to be empirical evidence lacked the the independence typical of experimental tests. Maxwell had arrived at the velocity of propagation for electromagnetic waves in a purely theoretical manner without any experimental evidence concerning the nature or existence of these waves. In addition, his calculation required the postultation of a displacement current which was also without experimental justification. *If* these theoretical postulations were rooted in experimental findings Maxwell would have certainly been in a more secure position but as it stood the evidence reduced to arguments from coincidence and this seemed a poor basis on which to accept the theory.

The support for Maxwell's theory was further weakened by the fact that similar equations for the field could be obtained from action-at-a-distance theories. The strength of the Maxwell's theory had centred around the fact that its electrical consequences were in accord with all the known observations. Moreover, the field equations accounted for the experimentally observed electromagnetic forces and from these equations one could deduce the laws of electromagnetism. But now these successes could be claimed by the rival theory as well. The chief defender of the modified action-at-a-distance account was Helmholtz who had revised Maxwell's electromagnetic account into one that construed the aether as a polarized dielectric. The polarization was explained in terms of the particulate structure of the medium with electricity being the result of the motion of charged particles acting at a distance. Maxwell had tried to eliminate action-at-a-distance by using the notion of polarization in a dielectric based on contiguous action; Helmholtz's account of polarization only served to diminish its importance.[14] As Helmholtz remarked:

> The two theories are opposed to each other in a certain sense, since according to the theory of magnetic induction originating with Poisson, which can be carried through

> in a fully corresponding way for the theory of dielectric polarizations of insulators, the action at a distance is diminished by the polarization, while according to Maxwell's theory on the other hand the action at a distance is exactly replaced by the polarization...It follows...from these investigations that the remarkable analogy between the motion of electricity in a dielectric and that of the light aether does not depend on the particular form of Maxwell's hypotheses, but results also in a basically similar fashion if we maintain the older viewpoint about electrical action at a distance.[15]

In fact, because Helmholtz's theory was based on the most general form of expression for the force between two elements of current (consistent with Ampere's experiments) it had the added advantage of being even more general than Maxwell's,[16] thus giving it greater explanatory power. In 1870, prior to the publication of the Treatise, Helmholtz had already proposed an explanation of reflection and refraction, something that Maxwell's theory had thus far failed to do.[17] Although Helmholtz did not provide the mathematical details of the solution to the problem of refraction they were supplied a few years later by Lorentz.[18] These refinements also led to an explanation of the phenomenon of dispersion, something that had been hitherto unaccounted for by Maxwell's theory.[19]

These diffitulties, together with the fact that no <u>direct</u> experimental evidence for the claims that energy is radiated from an electromagnetic circuit or that electromagnetic effects are propagated in time, left Maxwell's theory in a rather precarious position. Furthermore, there could be no direct proof of the electromagnetic theory of light since light could not be produced in the lab by setting up electrical or magnetic vibrations of the necessary frequency. The disparity between the frequency of vibrations attainable in laboratory circuits and the established frequency range of light was far too great.[20] Maxwell himself remarked on the difficulty in comparing "the results of our sluggish electrical experiment with the alterations of light, which take place billions of times in a second."[21] He had little else to say regarding this particular problem but he did suggest a possible test for the displacement current, an hypothesis that was also without direct evidence:

> According to this view, the current produced in discharging a condenser is a complete circuit, and might be traced within the dielectric itself by a galvanometer properly constructed. I am not aware that this has been done, so that this part of the theory, though apparently a natural consequence of the former has not been verified by direct experiment. The experiment would certainly be a very delicate and difficult one.[22]

Although there is no evidence that Maxwell himself undertook to provide experimental evidence for his theory one of his students, the first to work in the Cavendish laboratory, M. Hicks, did attempt an experiment to measure the velocity of propagation of electromagnetic waves. Unfortunately the experiment was unsuccessful.[23]

The interesting feature about the experimental status of electromagnetic waves was that although Maxwell had been unsuccessful in their detection there were several experiments, done prabr to and even during 1880, whose results were considered to be extraneous due to lack of theoretical knowledge. As a result none of the findings was associated with the results predicted by Maxwell's theory. As early as 1842 Joseph Henry had noticed that a one-inch spark could magnetize needles over 30 feet away. This led to the formulation of a rather remarkable analogy by Henry - the hypothesis of an electrical plenum. He claimed that on the basis of the experiment:

> ...it would appear that the transfer of a single spark is sufficient to disturb perceptibly the electricity of space throughout at least a cube of 400,000 feet of capacity; and when it is considered that the magnetism of the needle is the result of the difference of two actions, it may be further inferred that the diffusion of motion in this case is almost comparable with that of a spark from a flint and steel in the case of light.[24]

Henry was the first to arrive at the idea that the Leyden jar discharge was oscillatory. This was done by studying the magnetization of steel needles by a jar. When the needles were removed from helices they were not always magnetized in the right direction.[25] Henry noted that this phenomenon seemed at odds with the views connecting electricity and magnetism and so concluded that the discharge of the Leyden jar was not correctly represented by the single transfer of an imponderable fluid from one side of the jar to the other. Instead the phenomenon required that one:

> ...admit the existence of a principle discharge in one direction and then several reflex actions backward and forward, each more feeble than the preceding until

> equilibrium is obtained. All the facts are shown to be in accordance with this hypothesis, and a ready explanation is afforded by it of a number of phenomena, which are to be found in the older works on electricity, but which have until this time remained unexplained.[24]

So, although Henry was the first to discover the oscillatory character of the spark no attempt was made to expand on its theoretical implications until Kelvin investigated and explained the phenomenon in 1853. Of course Henry could not have gone on to extend his experimental findings in any significant way without the advantage of a background theory but it is odd that his results were more or less ignored by those interested in electromagnetic phenomena. Henry's work ought to have been well known abroad; he had made presentations before the British Association and had papers published in Philosophical Magazine in 1842 and 1847. In addition his work on the oscillatory discharge and induction over a distance was reported by August de la Rive in A Treatise on Electricity, a volume that Maxwell offered as a reference in the Treatise.

In addition to Henry's findings in 1871 Elihu Thomson noticed sparks from metallic objects in the vicinity of a Ruhmkorff coil. With the necessary adjustments the sparks could be detected from as far as the sixth floor of an observatory. A similar discovery using a magnetic vibrator relay was made by Thomas Edison in 1875. In 1879 David Hughes noticed that he could receive signals in a microphone reciever while the same phenomena was discovered by Amis

Dolbear in 1882.[87] At the time such effects were interpreted as either induction effects, the manifestation of a new principle or, alternatively, as a new aetheric force that was as distinct from electricity as heat and light. Hughes, especially, thought that he had discovered a new principle. He noticed that his microphone reacted whenever the current flowing in a nearby coil was interrupted. This occured regardless of whether the microphone circuit was connected to the coil. He concluded that the loose contact in his receiver was in fact receiving eminations from the sparks accompanying the interruptions. In addition he confirmed that a similar effect was obtained from sparks produced electrostatically. The following year in 1880 he called together a group of noted scientists including Stokes to demonstrate his findings. He showed that signals sent from a spark transmitter could be detected at distances of up to 500 yards by a microphone contact connected to a telephone through which the signals were heard.[88] Hughes maintained that the signals were transmitted by electrical waves in the air but thought that they resulted from conduction. Unfortunately he was discouraged from persuing his line of thought by Stokes who convinced him that he was observing induction effects. Of course there was a sense in which Stokes was correct in claiming that the results were not due to conduction. The tragedy was the lack of insight on Stokes' part. Surely someone so well acquainted with Maxwell's theory ought to have realized the theoretical importance of Hughes' results, namely, that he had discovered standing waves produced by interference between incident and reflected waves. Initially the failure to see these and other experimental findings as significant had to do with the lack of mathematical and theoretical expertise, especially in the cases of Edison, Dolbear and E. Thomson. But even for the more informed such as Stokes, Maxwell's theory was far from being considered the received view. Due to its rather poor reception it is unlikely that any connection between the theory and experimental phenomena would have even been suggested much less seen as conforming instances. Hughes did publish an account of his work but not until long after the same results had been established by Hertz.[29]

One experiment that did have some bearing on the relationship between the electric field and the propagation of light was performed by John Kerr in 1875. He showed that when dielectrics are subjected to a powerful electrostatic force they acquire the property of double refraction. This was due to the similarity between their optical behavior and that of uniaxal crystals whose axes are directed along the lines of force. The most important experiments conducted at the time however were those of Helmholtz and his student Schiller in 1875-76.[30] These experiments emphasized the importance of discovering the nature of the effects produced by the translatory motion of electric currents. The idea that the convection of electricity is equivalent to a current was first stressed by Faraday in 1838 and later by Maxwell in the Treatise [sec. 768-70]. Although the experiments by Schiller and Helmholtz showed that any satisfactory theory must take into account some action in the dielectric, thereby ruling out theories based on direct action at a distance, they made no specification as to whether Helmholtz's theory of the dielectric could be seen as superior to Maxwell's.

In Germany in the 1870's the electrodynamic theories of both Weber and Neumann shared the fundamental physical assumption that electrodynamic actions are instantaneous and at a distance. They differed in their formulations of these assumptions and about the nature of electricity. Neumann's theory emphasized electrodynamic potential and was independent of any atomistic assumptions unlike Weber's theory which construed electricity as fluids of particles of two signs and possessing mechanical inertia. On this latter interpretation the particles interacted through a force or potential similar to gravitational attraction. Helmholtz wanted to contrast these German views of electrodynamics with those of Maxwell by exposing their experimental differences. (As I noted above Helmholtz's theory included those of Weber, Neumann and Maxwell as special cases.) What Helmholtz succeeded in establishing was the fact that the three theories agreed in their predictions about electrodynamic phenomena associated with closed currents but differed with respect to the predictions of phenomena associated with the oscillatory surgings of electricity in unclosed currents.

In an attempt to decide among the theories Helmholtz devised an experiment that was performed by H.H. Rowland in 1876.[31] The experiment confirmed that under certain conditions charges borne by moving electrical bodies exert magnetic actions. Unfortunately these magnetic actions were of no help in deciding between theories because their cause could be accounted for in two different ways. One involved Weber's theory and the displacement of charges through the motion of their ponderable carriers. The other took account of the variation of dielectric polarization of air or aether in a fixed volume of space resulting from the ~~motion~~ of electric force. In order to reach a decision regarding these two possibilities further experiments were proposed by Helmholtz in 1879; experiments in which polarization effects could be produced with the (macroscopic) motion of charges. Hence, these experiments had the distinct advantage of requiring no mechanical assumptions to produce the predicted effect. It was this technique that gave way to Hertz's detection of electromagnetic effects by purely electromagnetic means.

3. Hertz's Experiments and the Vindication of Maxwell's Theory.

3.1 Hertz's Pre-experimental Work. Hertz's first work on the electromagnetic theory began as early as 1878 when he won the Berlin Philosophical Faculty prize for work on unclosed currents. Helmholtz had set a problem that dealt with an implication of Weber's theory. Quite simply, when

3 of/de 3

oscillations of electricity are set up in an unclosed circuit, Weber's hypothetical electrical interia should manifest itself in a retardation of the oscillation. Using some methods suggested by Helmholtz Hertz proved that the inertia of electricity is either zero or less than a very small value; a conclusion that supported Helmholtz's disagreement with Weber's theory. A second problem set by Helmholtz in 1879 (the one referred to in section II) dealt specifically with the behavior of unclosed circuits in Maxwell's theory. One of the main assumptions of Maxwell's theory was that changes in dielectric polarization result in electromagnetic effects in the same way that conduction currents do. What was missing, and what Helmholtz wanted, was either an experimental test of the existence of such effects or of the electromagnetic production of dielectric polarization. Hertz declined to carry out the experiments on the grounds that the equipment (Leyden jars producing oscillations and open induction coils) did not appear capable of producing observable effects. It was not until 1886 that Hertz attempted to tackle the problem.

In 1884 prior to his experimental research Hertz published a theoretical paper entitled "On the Relations between Maxwell's Fundamental Electromagnetic Equations and the Fundamental Equations of the Opposing Electromagnetics".[38] Here he formulated a principle for the unity of electric forces which stated that electric forces having their origin in inductive actions are equivalent to equal and equally directed forces from an electrostatic source.[33] On the basis of this principle he proceeded to deduce a formulation of electrodynamics valid only for closed currents. In contrast to Helmholtz's experimental emphasis on unclosed currents Hertz showed that a theoretical decision about competing theories could be made on the basis of predictions for closed currents. Hertz deduced Maxwell's equations by an alternative means and presented them in a symmetrical form without assuming the existence of a dielectric aether. Instead, the propagation of electric and magnetic forces could be seen as resulting from the structure of the forces and potentials as expressed in the equations.

This of course was the method used by Riemann in 1858 and Lorentz in 1867 to associate electrical and optical phenomena with one another. The vector potentials showed themselves to be quantities propagated with a finite velocity - that of light - according to the same laws as the vibrations of light and heat. Hertz however did not share the commitment to the theory of retarded potentials and sought to remove these potentials from his formulation of the equations thereby yielding a much simpler account. As Hertz himself points out:

> [The] contrast between the two kinds of forces disappears as soon as we attempt to determine the propagation of these forces themselves i.e. as soon as we eliminate the vector potentials from the equations...These same equations connect together the electric and magnetic forces in empty space...whatever the origin of the forces.[34]

Although the system of forces given by Hertz's equations was the same as Maxwell's he was forced to acknowledge that there were other possible formulations (particularly Helmholtz's) that could also give exact solutions. He did however admit that if the choice was between the usual system of electromagnetics (direct action at a distance) and Maxwell's the latter was to be preferred for several reasons: one of being that the system of electromagnetic action of closed currents founded on direct action at a distance was incomplete. To overcome the problem it was necessary to introduce different kinds of electric force or admit the existence of actions which had not been taken into account. Maxwell's system, on the other hand, did not contain within itself what Hertz called "the proof of its incompleteness" [Misc. Papers p.289]. Secondly, any attempt to complete the usual system of electromagnetism resulted in laws that were very complicated to use. And, if we accepted the assumptions of such an account we would arrive at forces that were in fact the same as those required by Maxwell's theory. But, given this scenario we ought to prefer the latter because it provided the simplest exposition of the results.

The two specific points of interest with respect to Hertz's 1884 paper are his lack of attention to physical hypotheses and his reluctance to endorse Maxwell's theory for what seem to be reasons pertaining to a lack of theoretical uniqueness.[35] Although Maxwell's equations were, unlike those of other theories, compatible with physical assumptions shared by all electrodynamic theories, Hertz seemed reluctant to offer more than a tentative commitment based on simplicity.

Hertz did not think it possible to deduce a rigid proof that Maxwell's system was the only suitable account from the premisses available to him. Instead he formulated a deduction beginning with general fundamental premisses that were also admitted in the opposing system of electromagnetics and concluded by remarking that "the exact may be deduced from the inexact as the most fitting from a given point of view, but never as the necessary".[36] Despite his tentative acknowledgement of the superiority of Maxwell's equations Hertz made no mention of Maxwell's interpretation of those equations or his accompanying denial of action at a distance. In fact Hertz's formulation completely ignored the notion of a medium as the supporter of the field. Although the paper was based on a purely formal approach it had important implications for Hertz's later experimental work. It presented the view that evidence for the propagation of electromagnetic forces should be sought in electromagnetic as opposed to optical phenomena; the latter being the approach favoured by Maxwell.

3.2 Experimental Researches Pre 1888. At the time Hertz began his experimental work Oliver Lodge had been carrying out a series of experiments on the discharge of small condensors which led him to the observation of oscillations and waves in wires. Some years earlier Fitzgerald had predicted the possibility of such waves and attempted to discover the conditions for producing them. Hertz remarked that his experiments were not influenced by that research for he only knew of it subsequently.[37] However, it is important to point out that the method of producing electrical oscillations in a conductor had been known for quite some time prior to Hertz's work. Thomson and Helmholtz had both pointed it out and Schiller (1874) had used that same method to determine the inductive capacity of glass.[38] However, before one could demonstrate this propogation empirically and make any measurements on electrical waves it was necessary to produce waves of a sufficiently short period and devise a method for detecting them. This required that one be able to produce the phenomenon of transmission, a condition that was probably unknown to Maxwell and his contemporaries and something which required especially high frequencies.[39] Hertz remarked that he thought it impossible to arrive at knowledge of electromagnetic waves with the aid of theory alone for their observation through experiments depended not only on their theoretical possibility but on a surprising property of the electric spark that could not be forseen by any theory.[40]

The experiments done in 1886 provided Hertz with the methods he thought necessary for solving the Berlin Academy problem. Using the discharge of a small Leyden jar or induction coil one could produce a spark in a secondary circuit provided it could spring across a spark gap. These experiments enabled Hertz to produce more rapid electrical disturbances than had been previously avaliable to physicists.[41] Hertz interpreted the phenomena by showing that the dimensions of the secondary circuit were such as to make its free period of electric oscillation approximately equal to the period of oscillations in the primary circuit. When the disturbance passed from one circuit to another by induction it would be intensified in the secondary circuit by the effect of resonance (provided the circuit had the proper dimensions). This mechanism supplied the method for detecting electrical effects in air at a distance from the primary disturbance. One needed only a suitable detector to observe the propagation of electric waves in air.

In the spring of 1887 Hertz published two papers, numbers IV and V in EW. According to the results of the former paper, ultra-violet light displayed the property of increasing the sparking distance of the discharge of an induction coil as well as those of other discharges.[42] Hertz offered no theoretical explanation for this effect but at the end of the latter paper he remarked that when the primary spark is illuminated it looses its power of exciting rapid electrical disturbances. Sparks induced in a secondary conductor vanish when an arc of light is started and the primary spark looses its crackling sound. Because the spark is particularly sensitive to light from a second discharge the oscillations always cease if we draw sparks from the opposing faces of the knobs by means of a small insulated conductor.

Paper V dealt with, among other things, the determination of forces at greater distances. When Hertz placed the primary conductor in one corner of a 12 x 14 m. room sparks could be perceived in the farthest parts of that same room. However, in neighbouring rooms the action could not be perceived due to the damping effects of the walls. On the basis of his observations Hertz determined that at distances beyond three meters the force is everywhere parallel to the primary oscillation (this being the region where electrostatic force is negligible and only the electromagnetic force is effective). At the time of these experiments it was agreed that the electromagnetic force of a current element was inversely proportional to the distance in contrast to the electrostatic force (the difference between the effects of the two poles) which was thought to be inversely proportional to the third power of the distance. What Hertz found especially surprising then was the continual increase of the distance for which he could perceive the action. He also noted that in the direction of the oscillation the action becomes weaker more rapidly than in the perpendicular direction. There were also some areas where the direction of the force could not be determined for it appeared to have approximately the same strength in all directions. However, because the force could not act simultaneously in all directions it was thought to assume different directions in succession. The best explanation that Hertz could think of for this phenomenon was that electrostatic and electromagnetic forces were propagated with a different velocity. Any difference between them of course implied a finite rate of propagation for at least one; the first indication of a finite rate of propagation for electrical actions.

Hertz spent the summer of 1887 investigating the electromagnetic influence of insulators using this new method of oscillations. During the course of his experiments he noticed that the phenomena (the sparks induced in the resonating circuit by the primary disturbance or oscillator) could be modified by the presence of a large mass of insulating substance. The object was to test for the presence of electromagnetic actions that were thought to accompany electrical disturbances in insulators (where electrostatic actions were already known to exist).[43] Hertz tested several substances and concluded that the observations constituted a real electromagnetic effect as opposed to one arising from electrostatic forces or from currents resulting from residual conductivity. Hertz went on to conduct some experiments that he thought would establish a finite rate of propogation of the electric forces. Although he was unsuccessful in this attempt Hertz did not feel that this negative result ought to weigh against his otherwise positive results.

Having reported his successes to the Berlin Academy Hertz now seemed ready to tackle the problem designed to test

Maxwell's electromagnetic theory. The assumptions required to arrive at Maxwell's equations were the following:

> 1) Changes of dielectric polarization in non-conductors produce the same electromagnetic forces as currents that are equivalent to them;
>
> 2) Both electromagnetic and electrostatic forces are capable of producing dielectric polarizations; and
>
> 3) In all these respects air and empty space behave in the same manner as other dielectrics.

The first assumption had been shown to be correct and while Hertz was working on an experimental proof of the second it occurred to him that the third hypothesis contained the core of the Faraday-Maxwell approach. Hertz had already shown that when variable electric forces act with insulators the polarizations corresponding to these forces exert electromagnetic effects. If it could be shown that the first two hypotheses (or their consequences) were correct for any given insulator it would follow that electromagnetic waves of the kind predicted by Maxwell could be propagated in the insulator with a finite velocity; one that may differ from that of light. However, it was a somewhat different question whether the same forces in air were also accompanied by polarizations capable of producing electromagnetic effects. If they were then it would immediately follow that electromagnetic actions were propogated with a finite velocity. Although Hertz saw no way of testing the first two hypotheses for air both would be proved if he could demonstrate waves with a finite rate of propagation in air.

3.3 The 1888 Experiments on Electromagnetic Waves in Air and Wires. While Hertz was attempting to formulate an experimental approach it occurred to him that it might be possible to proceed with a test even if the velocity under consideration was greater than that of light. Using the rapid oscillations of a primary conductor he produced regular waves in a straight wire. A secondary conductor would then be exposed to the influence of the waves propagated through the wire and to the direct action of the primary conductor which was propagated through air. Both actions were made to interact with the interferences being produced at different distances from the primary circuit. This would enable Hertz to find out whether the oscillations of the electrical force at certain distances exhibited a retardation of phase in comparison with the oscillations in the neighbourhood of the primary circuit. The experiments showed that the inductive action was propagated with a finite velocity greater than the propagation of electric waves in wires, with the ratio being 45:28. Hence, it followed that the absolute value of the first was of the same order as the velocity of light while the value for the latter velocity was 2×10^{10} cm./sec.[44] Not only were Hertz's conclusions incorrect but the events leading up to the experiment itself provide a series of stumbling blocks and frustrations.

When Hertz initially set up the apparatus and found that the phases of interference were different at different distances he concluded that the alternation would correspond to an infinite rate of propagation in air. This proved to be extremely discouraging for Hertz and as a result he gave up experimenting for a number of weeks. Finally he came to the conclusion that it would be equally as important to continue with his research and prove that Maxwell's theory was false as it would be to prove it correct. The important point was to reach an outcome that was "definite and certain" [EW, p.9]. When Hertz began to examine his results he found that the sequence of the interferences could not be "harmonised with the assumption of an infinite rate of propagation" [Ibid.]. Instead Hertz found it necessary to assume that the velocity was finite but greater than that in the wire. At the time Hertz knew that this difference in velocities was highly improbable but he saw no reason to doubt the experimental result. It was always possible that the motion in the wire might be retarded by some unknown cause like an essential inertia of the free electricity [Ibid.]. Hertz felt it necessary to describe even his failings in order to show that he did not wish to establish a "preconceived idea in the most convenient way by a suitable interpretation of the experiments". However, as I mentioned above, not only did he encounter difficulties in reaching his conclusions but the final results of the experiments proved to be problematic.

One major error was that the time of oscillation had been overestimated. Poincaré [1890] showed that the period calulated by Hertz was $\sqrt{2}$ x the true period, making the velocity in air equal to that of light x $\sqrt{2}$.[45] Hertz's reliance on the calculation resulted from its agreement with the experiments of Siemens and Fizeau[46] and claimed that this error affected only the form of the research and not its substance.

The other more serious difficulty was the claim that velocities were different in air and wire. In the introduction to EW Hertz acknowledges the possibility of disturbances that may have had some effect on the iron stone located a short distance from the wire. However, at the time the experiments were performed Hertz had several reasons for thinking his conclusion correct. If the waves in the wire were the same speed as those in air (and of light) then lines of electric forces must be perpendicular to the wire. Hence a straight wire traversed by waves could not exert an inductive action on a neighbouring wire parallel to it. When Hertz found that this action did take place he concluded that the velocity of the waves could not be the same as those of light. In addition Hertz assumed that if lines of force were perpendicular to the wire it could be shown by means of a simple calculation that the energy propagated by a wave in a single wire would be logarithmically infinite. Hence he concluded a priori that such a wave would be impossible. Lastly Hertz found that alterations in the shape of the wire produced noticible effects on the velocity and on that basis concluded that there must be some obscure cause that produced the retardation; a cause that could also operate in straight wires. Although Hertz did not consider these conditions to be decisive in 1893 he does remark that in 1888 he had found them extremely convincing [Ibid.].

These same results were confirmed in a further series of experiments reported in paper VIII, "On Electromagnetic Waves in Air and their Reflection" [EW, pp. 125-136]. Although Hertz acknowledged that he may have been influenced in the interpretation of the results by a desire to find agreement with his prior experiments, he also offered a possible explanation of the discrepancy. He supposed that the difference in velocities may have been the result of special conditions of resonance in the room that he used. Although later experiments by Sarasin and de la Rive found results that agreed with the theory (they worked in a much smaller room) Hertz interpreted these as showing that his experiments were subject to local variations causing the phenomena to be different if the reflecting walls and rooms are different. Hence, only under certain conditions do the wave lengths have the values required by theory. Because Hertz's experiments contained "sources of error that were not understood" he felt that they could not be brought forward as arguments against a theory that was "supported by so many reasons based on probability" [Ibid., p.13]. Although there was a sense in which the experiments done by Sarasin and de la Rive restored faith in the theory while casting doubt on the validity of his own results, Hertz did not see them as decisive.

The interesting feature of Hertz's discussion in paper VIII is his lack of concern for the conflict between the experimental results and the theory. He was willing to think of the discrepancy as a minor difficulty compared to the many results that were found to be in agreement with the Faraday-Maxwell account. In claiming this, Hertz is careful to stress the independent nature of the experimental results despite their obvious agreement with Maxwell's theory:

> I have described the present set of experiments, as also the first set on the propagation of induction, without paying special regard to any particular theory; and indeed the demonstrative power of the experiments is independent of any particular theory. Nevertheless, it is clear that the experiments amount to so many reasons in favour of that theory of electromagnetic phenomena which was first developed by Maxwell from Faraday's views. It also appears to me that the hypothesis as to the nature of light which is connected with that theory now forces itself upon the mind with still stronger reason than heretofore...
>
> That Maxwell's theory, in spite of all internal evidence of probability, cannot dispense with such confirmation as it has already received, and yet may receive, is proved - if indeed proof be needed - by the fact that electric action is not propogated along wires of good conductivity with approximately the same velocity as through air. Hitherto it has been inferred from all theories, Maxwell's included that the velocity along wires should be the same as that of light [EW, p.136].

Shortly after this Hertz attempted to resolve the discrepancy on the basis of a discovery indicating that for short waves the difference between the velocities tended to disappear. Unfortunately Hertz had no method for testing this hypothesis using long waves (for he had no access to rooms that were large enough to accommodate the experiment). Nevertheless, he remarked that he had "little doubt that [decisive experiments would] decide in favour of equal velocities" [Ibid.]. A few years later in 1893 Sarasin and de la Rive carried out experiments in the Great Hall of Rhone and proved conclusively the equality of velocities in air and wires, thereby establishing agreement between theory and experiment.[48]

Having finished the experiments on the reflection of waves in March 1888 Hertz directed his attention to devising a clearer theoretical treatment of them. Up to this point he had interpreted the experiments from Helmholtz's perspective on electromagnetism. Hertz had previously shown that in a special limiting case the equations of Helmholtz's theory are the same as those of the much simpler Maxwell account. (Here Helmholtz's dichotomy between electromagnetic and electrostatic force vanishes leaving only one force propagated with the velocity of light). While testing his results against the assumptions of the theory Hertz ran into some difficulty with coiled wires. When a wire was rolled into a spiral and the velocity was measured along the axis of the spiral the wave moved much more slowly. When the velocity was measured along the wire itself it moved more rapidly. A similar phenomenon occurred with crooked wires.[49] Maxwell's theory was unable to account for this; according to the theory the propagation along the axis of the spiral must take pl~~ace with th~~e velocity of light for every form of conductor.[50] But again Hertz was not disuaded:

> In our endeavour to explain the observations by means of Maxwell's theory we have not succeeded in removing all the difficulties. Nevertheless, the theory has been found to account most satisfactorily for the majority of the phenomena; and it will be acknowledged that this is no mean performance. But, if we try to adapt any of the older theories to the phenomena, we meet with inconsistencies from the very start..., unless we reconcile these theories with Maxwell's by introducing the aether as dielectric in the manner indicated by von Helmholtz [<u>EW</u>, p.159].

In addition to continuing his research on propagation along wires [<u>EW</u>, pp.160-171] ~~Hertz~~ also performed a series of experiments [pp.172-185] using concave mirrors that enabled him to concentrate the action of an electric oscillation and make it perceptible at greater distances. By doing this he was able to carry out a variety of elementary studies commonly performed on light and radiant heat. These experiments gained immediate approval and were frequently repeated and confirmed. Of particular interest is Hertz's remark that a considerable part of their approval was due to philosophical reasons. The experiments were seen as a refutation of forces acting at a distance in the domain of electricity. For some time action-at-a-distance had provided an acceptable method of explanation in scientific contexts even though it was looked less favourably upon in common sense or philosophical settings.

The final series of experiments[51] attempted to show that magnetic force was also propagated with a finite velocity.

Theoretically this was not necessary since electric waves were also waves of magnetic force; however, the importance for Hertz was to detect the presence of magnetic force in the waves of electric force. The plan was to observe the mechanical forces which the waves exerted in ring-shaped conductors. In keeping with theoretical predictions Hertz found that the electric and magnetic forces were the same order of magnitude, with 'the preponderence of one over the other being determined by the proportions of the neighbouring parts of the ring and the fixed conductors [p.193]. The more these resembled infinitely thin wires the greater the magnetic force while the broader the surfaces the greater the electric force. In addition Hertz had hoped to make some observations on waves in free air. This last attempt was frustrated and Hertz was only able to examine the effects produced by waves travelling along wires. As a result Hertz concluded that these findings (researches on waves in wires) could not be used to decide between Maxwell's theory and action-at-a-distance accounts; the mechanical actions produced by waves in wires could be regarded as the result of both the electrification of the wires and the currents flowing in them. On the other hand, if one accepted the theoretical view that waves in wires were a form of waves in air then the experimental result posed no difficulty.

In summary then, Hertz's experiments established several connections between electric and light waves. By passing electric waves through huge prisms of hard pitch Hertz demonstrated that electric waves refract in the same manner as light waves. Electric waves were polarized by directing them through a grating of parallel wires and then diffracted by interrupting them with a screen with a hole in it. He then reflected them from the walls of the room in order to obtain the interference between the original and reflected waves. Hertz noticed that conducting materials would cast a shadow when placed in the path of the electrical rays while insulators did not stop the ray. Hertz remarked that "it passes right through a wooden partition or door, and it is not without astonishment that one sees the sparks appear inside a closed room." [Cf. Paper VI, pp.95-106].

In the introduction to EW Hertz claims that his experiments prove, for the first time, the propagation in time of a supposed action-at-a-distance:

> This fact forms the philosophic result of the experiments; and, indeed, in a certain sense the most important result. The proof includes a recognition of the fact that the electric forces can disentangle themselves from material bodies, and can continue to subsist as conditions or changes in the state of space. The details of the experiments further prove that the particular manner in which the electric force is propagated exhibits the closest analogy with the propagation of light. Indeed that it corresponds almost completely to it. The hypothesis that light is an electrical phenomena is thus made highly probable [EW, p. 19].

Hertz goes on to point out that the (approximately) equal velocity is only one element among many others that support the hypothesis; however, to give a strict proof would require experiments on light itself.

Although Hertz maintained that the accomplishments of the experiments had been achieved independently of the correctness of particular theories, he did acknowledge the connection between the experiments and the theories they were disigned to test [Ibid.]. Until now Maxwell's theory had surpassed its rivals in its unity but it was unable to displace them insofar as its strength "depended solely on the probability of its results and not on the certainty of its hypotheses" [Ibid.]; hypotheses that had no basis in experiment. Hertz's results changed this situation insofar as the experiments were seen as confirming tests of the fundamental hypotheses of Maxwell's electromagnetic theory.

As we have seen from this brief discussion of the experiments the relationship that they bore to Maxwell's theory is by no means straightforward. Similarily, the extent to which they provide a proof of the fundamental assumptions is something that needs a closer analysis. In the following section I discuss Hertz's understanding of the electromagnetic theory and the extent to which his interpretation sheds a different light on what it meant for the electromagnetic theory to be confirmed.

4. Hertz's Theoretical Interpretation of Maxwell's Equations.

Although Hertz saw his experiments as confirming Maxwell's theory it is important to clarify how Hertz interpreted the fundamental hypotheses of the theory and to what degree he shared Maxwell's rather speculative physical interpretation of these hypotheses. There is a sense in which this problem seems relatively straightforward. Hertz appears to share Maxwell's idea that electromagnetic phenomena are caused by polarizations in a dielectric medium filling space. However, in his reformulation of the theory some of Maxwell's physical quantities, especially displacement, have been abandoned and in the spirit of Maxwell's own account no mechanical hypothesis about of the structure of the aether is forthcoming.

While working on his experiments Hertz had been guided by Helmholtz's work but found that in the limiting case of Helmholtz's theory (which leads to Maxwell's equations) the physical basis of the theory disappears when action at a distance is disregarded. In the attempt to formulate a consistent theoretical account Hertz began with Maxwell's equations (unlike in his 1884 paper where he deduced them) and eliminated the portions of the theory that he viewed as dispensable: those that did not affect "any possible phenomena" [Ibid.]. Because Hertz saw Maxwell's account, its representation as a limiting case of Helmholtz's theory and his own formulation, as leading to the same equations he considered them all to be special cases of Maxwell's theory, even though they were not a precise representation of Maxwell's views. Hence his famous remark that "Maxwell's theory is Maxwell's system of equations" [Ibid.].

In order to expose what he took to be the difficulties with Maxwell's formulation of the theory Hertz began by distinguishing four ways that action could be transmitted across free space, ranging from direct attraction to indirect attraction. The first category regarded the attraction of two bodies as due to a kind of spiritual "affinity" between them with the force exerted by each being bound up with the presence of the other body. This was the conception of pure action at a distance presented by Coulomb's inverse square law. The second category involved the notion of a spiritual "influence" rather than an "affinity". Each of the bodies continually strove to excite attractions of definite magnitude and direction at all surrounding points. These strivings varied from point to point and filled the surrounding space. The acting body was the seat and source of the force with no change occurring in the space itself. This was the standpoint of the potential theory. The third conception built on the second by assuming that forces induced changes in the surrounding space which in turn gave rise to new distance forces. Hence the attractions between separate bodies depended partly on action at a distance and partly on the polarization of small parts of the medium. This was Helmholtz's conception of electromagnetism. In the limiting case we could decrease the part of the energy that has its seat in the electrified bodies and instead seek the whole of the energy in the medium. Because no energy corresponded to the electricities on the conductors the distance forces became infinitely small. As a result there was no free electricity; instead it behaved like an incompressible fluid travelling in closed currents. Although the mathematical treatment of this limiting case leads to Maxwell's equations Hertz is quick to point out that "in no sense must this be taken as meaning that the physical ideas on which it is based are Maxwell's ideas" [Ibid., p.24]. The fourth and final category dealt with 'the pure conception of action through a medium.' The action of material bodies on one another was caused by changes in space. The polarizations were not seen to be the result of distance forces; instead the existence of these forces as well as the electricities from which they proceed were denied. The explanation of the nature of the polarizations was deferred leaving the expressions of electricity and magnetism with no value beyond that of abbreviations [Ibid., p.25].

The mathematical treatment of the fourth point of view was regarded by Hertz as coinciding completely with the limiting case of the third. However, there was a fundamental difference with respect to their physical interpretations. The fourth category denied the existence of distance forces viewing polarizations as fundamental while the third saw them as the cause of polarizations. Although Hertz regarded the fourth standpoint as Maxwell's, he indicated that some of Maxwell's general explanations left room for doubting whether he wished to discard the notion of distance forces entirely. Maxwell sometimes spoke of the force or displacement in a dielectric that caused the particles to be charged positively or negatively with electricity, while other statements appeared to contradict this claim. He also assumed that electricity exists in conductors, that its movement forms closed currents with the displacements in the dielectric and that it behaves like an incompossible fluid; statements that fit more with the third category than the fourth. Hertz believed that these apparent contradictions were the result of a misunderstanding about Maxwell's use of qualitative physical notions. He offered the following interpretation:

> Maxwell originally developed his thory with the aid of very definite and special conceptions of electrical phenomena. He assumed that the pores of the aether and of all bodies were filled with an attenuated fluid which could not exert forces at a distance. In conductors this fluid moved freely, and its motion formed what we call an electric current. In insulators the fluid was confined to its place by elastic forces, and its "displacement" was regarded as being identical with electric polarization. The fluid itself, as being the cause of all electrical phenomena, Maxwell called "electricity". [Ibid., p.27].

Hertz went on to point out that Maxwell did not eliminate all of his physical conceptions, some remained that were derived from his earlier ideas. Hence the word electricity took on a double meaning in Maxwell's work. It denoted a positive or negative quantity that forms the starting point of distance forces as well as an hypothetical fluid from which no distance forces can proceed. However, as mentioned in Chapter II, when Maxwell wrote DT and the Treatise he abandoned many of his physical ideas in favour of casting the theory in Lagrangian form. What Hertz failed to mention is that the physical ideas associated with the formal account of the theory had, for Maxwell, the status of an analogy. This is especially true of electricity, the nature of which Maxwell acknowledged to be unknown [Cf. Chapter II].

What Hertz attempted in his theoretical papers of 1891[32] was a reformulation of Maxwell's theory using a limited number of physical conceptions. Hertz was uneasy with Maxwell's formulation insofar as it began with the assumption of direct action at a distance, investigated the laws according to which hypothetical polarizations of the aether varied under the influence of distance forces and concluded that although the polarizations do vary the cause is not distance forces. Hertz felt that this approach suggested that something was wrong with either the final result or the method used to arrive at it [p.196]. In addition there were a number of superfluous physical and mathematical ideas that were significant only in the context of a direct action-at-a-distance account. One such idea was the "dielectric displacement (polarization) in free aether as distinguished from the electric force which produces it, and the relation between the two - the specific inductive capacity of the aether" [p.196]. These distinctions were meaningful only if it were possible to remove the aether from a space while allowing the force to remain. Again, such ideas may have been applicable in FL and PL but they had no direct role once the theory had been cast in Lagrangian form. A further superfluous idea was the mathematical quantity, vector potential. The vector potential had allowed for the replacement of distance forces by magnitudes which were determined at every point in space by conditions at neighbouring points. Once it was possible to identify forces as magnitudes there seemed little point in retaining potentials. In additon, as mentioned in Chapter II, the vector potentials had been associated with a number of physical notions, including displacement. Hence there seemed no reason to retain them once the corresponding physical quantities had been eliminated.

Although Hertz developed his formulation of the electromagnetic theory along the lines discussed in category four he simply postulated the fundamental equations rather than deriving them from a mechanical model of the aether. In fact, neither the ideas outlined in category four nor any of the others implied that a mechanical account of the structure of the aether could be given. Instead Hertz proposed the symmetrical relations between the electric force E and the magnetic force H in the free aether (forces and polarizations were assumed to be identical and c represented the speed of light) yielding the following equations:

$$1/c\ \partial H/\partial t = -\mathrm{curl}\ E \qquad \mathrm{div}\ H = 0$$

$$1/c\ \partial E/\partial t = \mathrm{curl}\ H \qquad \mathrm{div}\ E = 0$$

Hertz emphasized at the outset of the paper [pp.195-197] that his intention was to give the fundamental ideas and the formula that connects them. Although explanations were added to the formulas they were not to be considered proofs; instead they were to be regarded merely as facts derived from experience [p.197].

In the second paper Hertz addressed the problem of bodies in motion taking into account the fact that disturbances of the aether that arise with the motion of ponderable matter must exhibit some effects; even though we have no knowledge of them. Unfortunately the failure to have such knowledge while adhering to the position was equivalent to saying that questions concerning motion cannot be considered without the introduction of arbitrary assumptions about the motion of the aether [p.242]. In fact, all available knowledge seemed to indicate that the motion of the aether produced no effects [pp.241-243]. One could assume that the aether moved independently of matter (even in its interior): but to adapt the electromagnetic theory to this hypothesis one must also suppose that the aether and matter are independent at every point in space. However, as Hertz pointed out, a consideration of the phenomena to be investigated revealed that none required the hypothesis of the independent motion of aether and matter because no indication of the magnitude of the relative displacement had been obtained. As a result these phenomena were also consistent with the hypothesis that denied displacement and assumed that the aether moved with matter. This latter view included the possibility of taking into account only one medium filling space. It was the view adopted by Hertz even though he admitted that:

> ...a theory built on such a foundation will not possess the advantage of giving to every question that may be raised the correct answer, or even of giving only one

> definite answer; but it at least gives possible answers to every question that may be propounded, i.e. answers which are not inconsistent with the observed phenomena nor yet with the views which we have obtained as to bodies at rest [p. 243] (emphaisis added).

Hertz concluded his second paper by remarking that the only value he attatched to his theory of electromagnetic forces in moving bodies was its systematic arrangement [p.268]. Although the theory enabled one to provide a complete treatment of the electromagnetic phenomena in moving bodies it involved certain arbitrarily imposed restrictions; restrictions that Hertz thought were "scarcely probable" in relation to the "actual facts of the case" [Ibid.].

In his attempt to eliminate any arbitrary assumptions from the theory Hertz acknowledged its rather abstract and colourless appearance [p. 28]. However, he did remark that if one wanted to lend more colour to it it was possible to aid one's powers of imagination by using concrete representations of concepts such as electric current, polarization etc. But, one must remain cautious and not confuse the:

> simple and homely figure as it is represented to us by nature, with the gay garment which we use to clothe it. Of our own free will we can make no change whatever in the form of the one, but the cut and colour of the other we can choose as we please [p. 28].

As mentioned above Hertz, like Maxwell, seems committed to the idea that the medium is the seat of force; however, this commitment is not accompanied by a corresponding physical hypothesis. In the supplementary notes to EW Hertz elaborated on the findings described in paper X. In note 24 [p.275] he claimed that the experiments discussed in that paper proved that in the case of rapid variations of current the changes penetrate into the wire from without. He goes on to remark:

> It is thereby made probable that in the case of a steady current as well, the disturbance in the wire itself is not, as has hitherto been assumed, the cause of the phenomena in its neighbourhood, but that, on the contrary, the disturbances in the neighbourhood of the wire are the cause of the phenomena inside it [p. 275].

Of particular interest is Hertz's remark in the next sentence:

> That the disturbances in the wire are connected with a regular circulation of material particles, or a fluid assumed ad hoc, is a hypothesis which is neither proved not disproved by our experiments; they simply have nothing to do with it. We have neither any right to oppose this hypothesis, nor have we any intention of doing so, on the ground of the experiments here described [p.275] (emphasis added).

This reluctance to sanction an hypothesis about the nature of the aether is further evidenced in the second of his theoretical papers. Although Hertz accepted, for the purposes of the paper, that the aether was mechanically dragged by moving bodies he did so for what could be construed as essentially pragmatic reasons. First, within the domain of electromagnetic phenomena there was nothing incompatible with the idea of a dragged aether. Secondly, a denial would require the additional complication that two sets of electric and magnetic vectors would be required by each point in space; one for the moving body and one for the aether. So, although Hertz accepted the hypothesis of the aether drag he was clearly skeptical about its viability. He remained convinced that a correct theory should distinguish between the conditions of the aether at every point and those of embedded matter [p. 268]. Unfortunately this knowledge was unavailable and any attempt at providing such an account would require the addition of "arbitrary hypotheses" to the theory; something Hertz was unwilling to do.

How then should one characterize Hertz's commitment to the field theoretic account of electromagnetism? As we saw above his axiomatic formulation of Maxwell's theory eliminated the concept of displacement as something independent of electric force and thereby avoided the need to explain the connection between polarization of the aether and displacement of electricity. But, it was the symmetric relations between magnetic and electric forces as expressed in Hertz's equations that formed the essential core of Maxwell's theory - the laws governing the electromagnetic field. Like Maxwell Hertz interpreted "the energy of the field" literally but denied that a mechanical model of the aether was necessary for explaining Maxwell's theory. Although Maxwell had denounced his mechanical model of the aether prior to the Treatise the physical notions of displacement as well as the vortex model were retained, even if only as analogies describing very tentative possibilities.

Hertz was considerably more anxious to banish as many physical notions as possible from the foundations of the theory and in doing so he seems to leave little in the way of even a possible physical interpretation of what the nature of the aether/field might be.

Some commentators including Doran [1975] and Harman [1983] suggest, implicitly or explicitly, that Hertz was commited to a mechanical view of the aether. Although Doran acknowledges Hertz's denial that a mechanical model of the aether was necessary to explain Maxwell's theory she remarks that Hertz retained the mechanical postulates in his attempts to represent a continuous aether [op. cit., p.233]. Meanwhile Harman remarks that Hertz "remained committed to the belief that the electromagnetic waves were produced by an aether whose parts were connected by a mechanical structure"[p.11]. However, the important issue here which both commentators fail to emphasize, is the provisional character of Hertz's acknowledgement of the aether and aether drag. If one were to call Hertz's attitude one of commitment then, to an even greater extent than Maxwell's, it was a commitment to a "something I know not what"; an hypothesis whose consequences could be characterized by a set of equations but an entity whose essence was unknown.

Given this interpretation of Hertz's account of electromagnetism how should we view his experiments vis a vis their supposed confirmation of Maxwell's theory. Hertz himself remarked at the end of his introduction to the experimental section of EW that the object and result of the experiments was to test and confirm the fundamental hypotheses of the Faraday-Maxwell theory. Fitzgerald in an address to the British Association in 1888 remarked that Hertz's experiment proves the aethereal theory of electromagnetism. However, what Hertz's experiments seem to confirm is the validity of Maxwell's theory insofar as the equations provide an accurate representation of electromagnetic phenomena. More specifically the experiments showed that electromagnetic waves were finitely propagated in air. This however was by no means a confirmation of all of Maxwell's physical or even mathematical ideas. Hence, the extent to which one could interpret Hertz's experiments as proof of the reality of the aether was certainly limited. The nature of the relationship between the electromagnetic theory and the aether remained a puzzle.[53] Electromagnetic phenomena may have been shown to have their origin in field but the nature of that field described as an aether was unknown.

In conclusion then even Hertz's experiments failed to support realism about the aether, the structure that was initially seen to be the unifying mechanism for electromagnetics and optics. Displacement had been disregarded and hence the need for experimental evidence to justify its place in the theory was no longer necessary. In what follows I discuss the relationship between theory and experimental evidence and whether an epistemology of experiment is strong enough to support an account of scientific realism.

Notes

1. This is a term used by Andy Pickering to describe the way in which scientists interact with their experimental environment in order to achieve successful results. For more on this notion of "tinkering" see his forthcoming paper "Experiment and the Real", PSA 1986, Vol. 2.

2. See Thomas K. Simpson, "Maxwell and the Direct Experimental Test of his Electromagnetic Theory", Isis, Vol. 57, 4, 1966 pp. 411-432. As Simpson points out, "What Maxwell measured as empirical support for his theory was the ratio of the values of a single quantity in units of the two systems of measurement fundamental in Maxwell's formulation of the electromagnetic theory: the 'electrostatic' and 'electromagnetic' systems of units" (Cf. p.412). Maxwell's argument revealed that the velocity of propagation could be obtained from this ratio.

3. The equations in their final form constitute a series of connecting vector quantities that apply to any point in a varying electric or magnetic field. They are: curl $H = \delta D/\delta t + j$; div $B = 0$; curl $E = -\delta B/\delta t$ and div $D = \rho$. H is the magnetic field strength, D is the electric displacement, t is time, j is current density, B is magnetic flux density (or magnetic induction - the magnetic flux passing through a unit area of a magnetic field in a direction at right angles to the magnetic force), E is the electric field strength and ρ is the volume density of charge. Using these equations Maxwell was able to show the interrelationship between electricity and magnetism. Whenever a varying electric field exists there is also a varying magnetic field induced at right angles and vice versa. The two together form an electromagnetic field with each field vector obeying a wave equation. For an interesting discussion of the derivation of Maxwell's equations see M.S. Longair Theoretical Concepts in Physics. Cambridge: Cambridge University Press, 1984.

4. For a formal presentation of these differences see Heimann op. cit. especially n. 191 p.205.

5. Sir George Gabriel Stokes, "Report on Double Refraction" in Report of the Thirty-Second Meeting of the British Association for the Advancement of Science. London, 1863, pp. 253-282.

6. George Green, "On the Laws of Reflection and Refraction of Light" in Transactions of the Cambridge Philosophical Society. Vol. 7, 1837, pp.1-24.

7. See Sir George Garbiel Stokes, Memoir and Scientific Correspondence 2 Vols. Cambridge: Cambridge University Press, 1907 and reprinted in the Sources of Science series,

No. 139, pp.25-26.

8. Augustin Louis Cauchy, "Memoire sur la theorie de la lumiere", Memoir de l'Academie. Vo. 10, 1830, pp.293-316.

9. Cf. Doran, op. cit. Doran provides an excellent discussion of the elements involved in the early attempts at a demechanization of the aether.

10. James MacCullagh, "Short Account of Some Recent Investigations Concerning the Laws of Reflection and Refraction at the Surface of Crystals" in the British Association Report 1835, p. 37. Se also his paper entitled "On the Laws of Crystalline Reflection" Phiosophical Magazine Vol. 10, pp. 42-45.

11. MacCullagh criticized Cauchy's theory because it failed to guarantee continuity of the displacement at the interface of the two media. To counter this MacCullagh chose to adopt the assumption that optical vibrations are parallel to the plane of polarization and not perpendicular.

12. MacCullagh, op. cit., 1837.

13. In Fitzgerald's model Maxwell's electric displacement represented the rotation of MacCullagh's elastic solid and magnetic force represented the velocity of aether streams. It was Fitzgerald's intention to abolish material ideas about the structure of the aether.

14. Cf. Doran, op. cit.

15. Hermann von Helmholtz, Wissenchaftliche Abhandlungen 3Vols. Leipzig, 1882-1885, Vol. 1, pp.545-628. Quoted in Woodruff "The Contributions of Hermann von Helmholtz to Electrodynamics" Isis, Vol. 59, 1968, pp.300-311.

16. Cf. R.T. Glazebrook, James Clerk Maxwell and Modern Physics London: Cassell & Co., 1901.

17. Cf. Whittaker, op. cit., p.304.

18. The details can be found in Lorentz's inaugural dissertation Over de theorie der teruqkaatsing en breking van het licht Arnhem 1875. See also Zeitschrift fur Math. und Phys. xiii, 1877, pp.1-205.

19. Dispersion is the phenomenon associated with the differences between the degree to which light rays of different colours are bent; for instance, violet light is refracted more strongly than red light. It is due to dispersion that one experiences a rainbow. Maxwell's theory failed to account for dispersion because of the way in which K, the numerical constant characterizing the medium and matter, was interpreted. The

bending of a light ray falling on a piece of glass under a given incidence was thought to be governed by the value of K for glass. According to Maxwell's theory K was determined solely by the glass and not by the color of light. Hence, all light rays should be bent to the same degree and dispersion should not occur. To resolve this problem one had to assume that K is not only a characteristic of the glass (matter) but that it should vary according to the color of the light. In the presence of matter Maxwell's theory could be considered only an approximation. Once again it was Maxwell's inability to differentiate between the aether as a carrier of light waves and matter that led to difficulties with the theory. It was Lorentz's attempt to establish a conception of the aether as a substance distinct from matter that enabled him to overcome these difficulties. For a discussion of Lorentz's innovations see Tetu Hirosige "Origins of Lorentz's Theory of Electrons and the Concept of the Electromagnetic Field" Historical Studies in the Physical Sciences vol. 1, 1969, pp.151-209.

20. Cf. Thomas Simpson, "Maxwell and the Direct Test of his Electromagnetic Theory", Isis, Vol. 57, 4, No.190, pp.411-432.

21. See Maxwell's article entitled "Ether" reprinted in Papers 2, p.772. Others including Fitzgerald and Oliver Lodge discussed and worked on the problem but there is no evidence that Maxwell became involved to any extent. Cf. Simpson, op. cit.

22. "On the Method of Making a Direct Comparison of Electrostatic and Electromagnetic Force" (1868) in Papers 2, p.139.

23. For details of the experiment see Charles Suskind, "Observations of Electromagnetic Wave Radiation before Hertz" Isis Vol. 55, No. 179, 1964, pp. 32-42 as well as A History of the Cavendish Laboratory 1871-1910 London: Longmans, Green and Co., 1910, p.19.

24. Quoted in Suskind, op. cit. The original appears in the Proceedings of the American Philosophical Society 1842, 2: 193-196.

25. This phenomenon was referred to as "anomalous magnetization". The thing Henry found puzzling was not the magnetization but the fact that the currents had no simple direction. Cf. Oliver Lodge, Modern Views on Electricity London: MacMillan & Co., 1907, appended lecture III, pp.387-409.

26. The Scientific Writings of Joseph Henry. Washington: The Smithsonian Institute, 1886, Vol. 1, p. 201.

27. Both Dolbear and Hughes had invented this new type of microphone which was similar to a telephone receiver.

28. This was later reported in The Electrician, Vol. 43, 1899, pp.40-41.

29. For details of some of this early experimental work see Whittaker, Simpson [1966] and Suskind, op. cit. as well as Salvo D'Agostino, "Hertz's Researches on Electromagnetic Waves" in Historical Studies in the Physical Sciences, Vol. 6, 1975, pp.261-323.

30. Cf. Whittaker, Vol. 1, Chapter 10 as well as Philosophical Magazine, i, 1875 pp. 337, 446; viii, 1879, pp. 85, 229; xiii, 1882, pp. 153, 248.

31. Rowland's experiment consisted of an electrified body (a disk of ebonite) coated with gold leaf and capable of turning rapidly around a vertical axis between two fixed plates of glass each of which was gilted on one side. The gilt faces were earthened while the disk recieved electricity from a point placed near its edge. Each coating of the disk formed a condense with the plate of glass that was nearest to it. An astatic needle was positioned above the upper condensor plate almost over the edge of the disk. When the disk was rotated a magnetic field was produced. What the experiment showed was that the convection current produced by the rotation of a charged disk produces the same magnetic field as an ordinary conduction current flowing in a circuit that coincides with the path of the convection current. Details summarized from Whittaker, op. cit.

32. Reprinted in Hertz's Miscellaneous Papers London: MacMillan & Co., 1896, pp.273-290.

33. This principle was thought to be analagous to the unity of magnetic force predicted by Ampere and holding between ponderomotive forces exerted by a magnetic pole and those exerted by an electric current. Ibid.

34. Ibid.

35. Hertz's brand of agnosticism is more fully developed in The Principles of Mechanics New York: Dover, 1957. In what follows I shall discuss how Hertz's approach differs in important ways from the thoroughgoing instrumentalism of someone like Mach.

36. Ibid. The distinction between deductions that are only the "most fitting" rather than "necessary" is discussed by Hertz in the context of conclusions derived from the principle of conservation of energy. The mode in which such conclusions are deduced is the most fitting from the usual standpoint of electromagnetics because it corresponds exactly to the accepted proposition from which Helmholtz (1847) and Kelvin (1848) deduced induction from electromagnetic action. Hertz points out that other methods may be possible:

..for just as in that proposition so we have in ours made tacit assumptions besides the principle of conservation of energy. That proposition also is not valid if we admit the possibility that the motion of metals in the magnetic field may of itself generate heat; that the resistance of conductors may depend on that motion; and other such possibilities [Ibid., n.1].

37. Electric waves being Researches on the Propagation of Electric Action with Finite Velocity Through Sapce D.E. Jones (trans.). New York: Dover, 1962, Cf. Introduction p.3.

38. Cf. Glazebrook, op. cit.

39. Cf. Simpson, op. cit. The connection with high frequencies was made initially by Wilhelm von Bezold in a paper entitled "Researches on the Electrical Discharge" in Phil. Mag., 1870, 40 pp.42ff. In his introduction to Electric Waves [EW] Hertz acknowledges von Bezold's paper (repr inted in EW pp.54-62) which gave many of the same results that Hertz himself had found. Hertz's explanation as to why the results had been ignored was that the title suggested that the paper related to matters quite apart from electric oscillations or, alternatively, because von Bezold had described his communication as preliminary.

40. EW, p.3.

41. It is interesting to note that in the intoduction to EW Hertz claims that his convictions about testing Maxwell's theory were strengthened by finding that the oscillations he was dealing with were regular. He then refers to the paper describing the course of the investigation, "On the Very Rapid Oscillations", pp.29-53. Here he claims that there are several problems that he has not solved [p.49] and further remarked that although the nature of the observed electric disturbances manifested themselves as oscillations they didn't exhibit the characterisitic of perfectly regular oscillations. Later in a reference to the fifth paper "On the Action of a Rectilinear Electric Oscillation upon a Neighbouring Circuit" Hertz remarked that he was surprised by their variety and regularity.

42. For example, if the knobs of an induction coil are drawn so far apart that sparks can no longer pass, and if an arc of light is started at a distance of 1 or even 4 meters the sparking begins simultaneously and stops when the light goes out.

43. Hertz's method was as follows: The primary conductor exciting oscillations acts inductively on a secondary conductor with the induced disturbances observed by inserting a spark-gap. Both conductors are then set to the same period of oscillation with the secondary one brought very near to the primary one. The position must be such that any forces

acting on the secondary conductor are neutralized keeping it free from sparks. If this equilibrium is upset sparks resume. The system of induction balance also indicated a charge when large insulating masses are brought close to it. Because of the very rapid oscillations the quantities of electricity displaced in insulators by dielectric polarization are of the same order of magnitude as those that are set in motion by conduction in metals. Cf. EW, pp.95-96.

44. Hertz determined the velocity of electric waves in wires by observing the distance between nodes of stationary waves in the sire and calculating the period of the primary oscillation. Cf. "On the Finite Velocity of Propagation of Electromagnetic Actions" in EW, pp.107-123. Se also Whittaker, op. cit., Chapter ten.

45. Henri Poincare, Comptes Rendus, 111, 1890, p.322. Hertz also mentions this correction in the Introduction to EW written in 1893.

46. Cf. EW, p.114.

47. Reported in Comptes Rendus, 112, p.658.

48. Hertz acknowledges these experiments in the introduction to EW and remarks that he submits to these results with as much readiness as he initially felt hesitation in submitting to their earlier results; experiments that he felt to be no more conclusive than his own.

49. Cf. EW, p.159.

50. This of course requires certain conditions: that the resistance of the conductor does not come into consideration and that the dimensions of the conductor perpendicular to the axis are negligible in comparison with the wave length. Although both conditions are satisfied by the coiled wires the result is not borne out. Cf. EW, p.159.

51. Discussed in "On the Mechanical Action of Electric Waves in Wires", EW, pp. 186-194.

52. "On the Fundamental Equations of Electromagnetics for Bodies at Rest" reprinted in EW, pp.195-240 as well as a second paper "On the Fundamental Equations of Electromagnetics for Bodies in Motion", pp.241-268.

53. For a discussion of Maxwell's thoughts about the aether and aether drag see his article "Ether " in Papers, op. cit., as well as Matter and Motion, Cambridge: Cambridge University Press, 1877.

Chapter IV

Experimentation and Scientific Realism

1. Introduction. Traditionally, experimental evidence was thought to provide the kind of independent evidence that would justify belief in entities and processes postulated by theory. Regardless of whether these entities occupied a fundamental place in the theory (as was the case with the displacement current) they could not be sanctioned on the basis of theoretical considerations alone. Independent experimental evidence for the existence of the displacement current for instance would have done a great deal to secure the viability of Maxwell's account of electromagnetism. And, as we have seen, the fact that it was supposedly confirmed by electromagnetic and optical phenomena (in virtue of its unifying role) was simply not enough to warrant its acceptance.

In the last chapter I discussed the relationship between Hertz's experiments and Maxwell's theory. However, in addressing this issue it might be helpful to consider some of the current literature on experimentation and approximation and how it does or does not apply to this case. Minimally one expects that a methodology or epistemology of science explains the practice of science itself. In the preceeding chapters I argued that scientific realism fails in this regard and that a defence of realism based on methodological factors such as unification, is insufficient to establish the conclusions required by some of the more familiar articulations of the position. The ontological commitment characteristic of most forms of scientific realism includes belief in entities endowed with specific properties which enable us to successfully predict and explain other phenomena. A realism that denies this kind of specificity is one committed to a realm of entities or agents, the nature of which we have no knowledge. We can summarize the forgoing in the form of a dilemma. On the one hand we have a strong realism that upholds belief in entities and the truth of theories that describe them. However, as we have seen this kind of realism is difficult to maintain; the truth of our theories cannot be justified on the basis of theoretical virtues like explanatory power or unification and moreover the practice of science itself provides good reasons for thinking that our current theories are not "true" in any strict sense. Alternatively we can opt for a weaker form of realism; one that is agnostic about the truth of our theories but is nevertheless committed to a realm of entities postulated by these theories. The main problem with this account is that it requires a separation of the entities from the theory that describes them. Unfortunately it is the entities together with their properties than enable us to predict and explain phenomena. Without the accompanying theory this form of realism loses its significance. Not only do the entities require a very specific articulation in order to figure in scientific explanations but it is difficult to see how one could have any evidence for the existence of a particular entity independently of theory. Because our commitments are determined by a theoretical picture of the world it is difficult if not impossible to separate ontological commitments from theorizing.

Hacking's [1983] position has some of the characteristics of the weakened form of realism discussed above. In what follows I examine some of the details of Hacking's account. I argue that although it has considerable intuitive appeal it fails to meet the demands required of an epistemology of science. Before addressing Hacking's claims I look briefly at some of the philosophical problems raised by an "experimental" realism and how they apply to the case of Maxwell and Hertz.

2. Maxwell and Hertz - The Philosophical Problems of Experimentation. In the last chapter we saw that Hertz's results were seen, both by himself and others, as confirming Maxwell's electromagnetic theory despite the discrepancies and lack of experimental work on optical phenomena. However it is particularly interesting to consider the nature of the conclusions drawn on the basis on the experimental results especially in light of the theoretical agnosticism displayed by both Hertz and Maxwell. Because the experiments concentrated solely on electromagnetic rather than optical phenomena they "made probable" [EW, p.19] the hypothesis that light is an electromagnetic phenomena. Hertz concluded that

the way in which electrical force was propagated was analogous to the propagation of light. Nevertheless the analogy was so close that the two phenomena could be said to correspond almost completely to one another. Hertz saw his reformulation of Maxwell's equations as merely expressing the relations that hold between electric and magnetic forces rather than an account of their nature:

> After these equations are found, it no longer appears expedient to deduce them (in accordance with the historical course) from conjectures as to the electric and magnetic constitution of the aether and the nature of the acting forces, - all these things being entirely unknown. Rather, it is expedient to start from these equations in search of such further conjectures respecting the constitution of the aether [EW, p.201].

This approach was in keeping with Maxwell's view that one should emphasize relations between phenomena rather then the things so related.[1]

We must be cautious however not to interpret this emphasis on the relational aspects of the electromagnetic theory as indicating a kind of positivism or instrumentalism similar to that of Mach or the earlier critics of LaSage's aether model.[2] Although Hertz claims that his theory has value only in light of its systematicity, unlike traditional instrumentalists he acknowledges the possible existence of a medium but has both physical and philosophical reasons for thinking his theory improbable. First he realized that even the most abstract formulation of the theory ought to be able to provide a mechanism for distinguishing between matter and an aetherial medium: something neither his account nor Maxwell's could do. Secondly, as mentioned above, Hertz believed that there were some questions that were simply unanswerable:

> ...Can we by our conceptions, by our words, completely represent the nature of anything? Certainly not...with the terms "velocity" and "gold" we connect a large number of relations to other terms; and between all these relations we find no contradictions which offend us. We are therefore satisfied and ask no further questions. But we have accumulated around the terms "force" and "electricity" more relations than can be completely reconciled amongst themselves. We have an obscure feeling of this and want to have things cleared up. Our confused wish finds expression in the confused question as to the nature of force and electricity. But the answer which we want is not really an answer to this question. It is not by finding out more and fresh relations and connections that it can be answered; but by removing the contradictions existing between those already known, and thus perhaps by reducing their number. When these painful contradictions are removed, the question as to the nature of force will not have been answered; but our minds, no longer vexed, will cease to ask illegitimate questions.[7]

By rendering certain questions illegitimate or unanswerable he avoided the problem encountered by Maxwell of explaining the nature of electricity.[8] Instead Hertz accepted polarization as a primitive notion, a phenomenon that was expressed by electric force. These epistemic limitations did not entail a denial of so-called theoretical entities but rather an acknowledgement of our inability to uncover their true nature.

The differences between Hertz and Maxwell's approach to physical theory had mainly to do with their approach to models and the expression of theoretical ideas. Hertz was concerned with providing only the essentials, stripping Maxwell's theory of many quantities and ideas that he thought unnecessary for formulating the basic electromagnetic equations. But again this was not because he denied the existence of a reality represented by the mathematical formalism. Instead, he wanted to draw a sharp distinction between reality and our representations of it. We are perfectly free to supplement our mathematical theories with physical interpretations and images but we must keep in mind the arbitrariness of the latter. Hertz saw mathematics as the simplest and most accurate way to represent what the so-called known facts were - that is, the relations among phenomena. By restricting the number of physical interpretations we could achieve a degree of logical consistency that would aid in preventing confusionss about the phenomena that result in illegitimate questions. Mathematics would also provide a framework for a possible physical interpretation.

Although Maxwell shared Hertz's skeptical attitudes toward uncovering the true nature of reality his approach to physical theory differed from that of Hertz. His distaste for purely formal theories and his obvious awareness of the problems associated with physical hypotheses led him to adopt a method using the Lagrangian formalism supplemented by models and analogies. However in the use of <u>of</u> these models

Maxwell was both sensitive and faithful to the same distinction Hertz sees as important - the one between reality and our representations of it. It is because Maxwell sees models as imperfect representations that he takes the liberty of using several sometimes contradictory ones in attempting to explain a physical process. He is considerably less concerned with presenting a mathematical account stripped of unnecessary physical concepts than with attempting to provide some physical insights as to what the phenomena might possibly be like. All of Maxwell's attempts to use analogies (and models) were for the most part provisional: some merely aided in theory construction while others were thought to provide a possible interpretation of physical reality but not a true theory. This of course is not to suggest that Maxwell was not concerned with consistency. In his article on the atom [op. cit.] he remarked that "all that is necessary to form a correct mathematical theory of a material system is that its properties shall be clearly defined and shall be consistent with each other". Of course consistency with the phenomena and general dynamical principles is also necessary. However, because analogies (and models) describe possibilities they can take a variety of forms just as a mathematical theory can have a variety of physical interpretations.

In light of the agnosticism of Hertz and Maxwell what can we conclude about the theory-experiment relationship in this context? First the experiments did prove the finite propagation of electromagnetic waves in space showing that such waves do not require material bodies for their propagation. This is a slightly different formulation of the "very probable hypothesis" mentioned by Maxwell [DT, p.564] according to which electric and magnetic phenomena could be seen as the motion and strain of one and the same medium. Hertz does not explicitly mention the medium except to assert the existence of polarizations. But, recall that for Hertz polarization is a primitive physical concept for which he offered no causal explanation. Obviously the idea of polarization presupposes that something, an aether perhaps, is polarized; but Hertz has little to say about this:

> ...The explanation of the nature of the polarizations, of their relations and effects, we defer, or else seek to find out by mechanical hypotheses; but we decline to recognize in the electricities and distance forces which have hitherto passed current a satisfactory explanation of these relations and effects [EW, p.25].

The existence of polarizations was not something directly proved by experiment. Instead it constituted part of Hertz's theoretical account: a view that he acknowledged to be highly improbable. So, although no conclusions about the aether could by drawn, the experiments did show a stronger analogy between light and electrical phenomena than did the relationships that were evident prior to the experiments.

When examining the relationship between Maxwell's theory and Hertz's experiments we see that some predictions were borne out while others conflicted with the experimental outcomes and were simply left unaccounted for. As Hertz himself remarked, "the theory has been found to account most satisfactorily for the majority of the phenomena, and it will be acknowledged that this is no mean performance" [emphasis added, p.159]. The crucial point was that Maxwell's theory fared better than any of its competitors and, difficulties aside, to that extent one could think of it as being relatively confirmed.

This case raises some interesting problems for current accounts of the use of experimental evidence as a defence of scientific realism. Such accounts must of course take into consideration the approximative nature of the theory-experiment relationship as well as the ambiguities that are often involved in the interpretation of experimental evidence. However, as I shall argue, once these issues are acknowledged it becomes increasingly difficult to motivate scientific realism as an epistemological position.

3. <u>Scientific Realism as an Epistemology of Science.</u> In a paper entitled "The Function of Measurement in Modern Physical Science" Kuhn discusses what he calls "the fifth law of thermodynamics" - that no experiment quite gives the expected numerical result.[5] Consequently what scientists supposedly seek is not necessarily "agreement" but "reasonable agreement". However in specifying what constitutes reasonable agreement there seems to be no consistently applied or applicable external criterion. As Kuhn points out this notion of reasonable agreement not only

varies from one science to another but varies within any one part. In other words the criteria for "reasonable agreement" are decided on the basis of reasonable agreement within the scientific community. This is not to say that there is anything necessarily wrong with such a procedure but only that we ought to recognize it as part of the scientific process, part of what constitutes theory confirmation. This raises potential difficulties for expressing the relationship between theory confirmation in the relative sense described above and the corresponding notion that a highly confirmed theory is a true theory (unless of course one interprets truth as a relative or conventional notion); but, let me leave that issue for the moment and focus on the problem of disagreement.

Many of the philosophical accounts of the theory-experiment relationship assume that the correspondence between the experimental results and reality is relatively unproblematic. Although we all more or less accept the theory-ladenness of observation we nevertheless feel confident that given two competing theories and a series of experiments that can supposedly test the predictions made by each we would have no trouble deciding which of the two theories was confirmed. However, the search for fractionally charged quarks presents an interesting example of the naivite of this particular view.

An important consequence of the quark model of strongly interacting particles was that quarks would have fractional charge ± 1/3e or ± 2/3e (instead of integral multiples of e- the charge of the electron). When several experimental searches for these quarks were found to be unsuccessful the confinement hypothesis was proposed, claiming that free quarks could never be observed. It was assumed that they existed, bound together in triplets to make a baryon or nucleon, or in pairs consisting of a quark or anti-quark bound together to make a meson. The interesting point about these expriments is that while pursuing similar programmes Morpurgo was unsuccessful in detecting fractional charge while William Fairbank was successful in finding these charges.[6] Although Fairbank's conclusions were thought to be the result of an experimental artifact no explanation of such an effect was offered.[7] So, both Fairbank and Morpurgo were successful in establishing conclusions that were diametirically opposed. In each case the data were consistent with the proposed hypothesis. (In the Morpurgo case all charge was thought to be quantized in units of the charge of the electron while the Fairbank hypothesis postulated isolated particles carrying integral fractions of that charge.) Hence there was no reason to think that there was a technical difficulty with the apparatus. Morpurgo's initial discovery of continuously distributed charge confirmed neither hypothesis leading to a modification of the apparatus which later produced outcomes in accordance with the no fractional charge account. This case shows that even in experimental contexts the uniqueness condition characteristic of realism can lead us into difficulties. In Chapter I we saw that one can specify many theoretical models for a given set of data, here we find that in experimental contexts no one set of data presents itself unequivocally as the candidate for what is to be considered real. But uniqueness is not the only problem that besets "experimental" realism. Hacking's [1983] arguments for the reality of theoretical entities focus on their ability to be manipulated. However, he is careful to distinguish experimenting on an entity from manipulating it; the former does not commit you to believing that the entity exists but the latter does. For instance he writes that:

> We are completely convinced of the reality of electrons when we regularly set out to build-and often succeed in building - new kinds of devices that use various well-understood causal properties of electrons to interfere in other more hypothetical parts of nature [p.265].

Of course Hacking does not claim that reality is constituted by human manipulability, instead he claims that the best evidence we have for the reality of a postulated entity consists of measuring or "otherwise understanding its causal powers" [p.274]. And the best evidence we have for this understanding is our ability to build machines that are fairly reliable, machines that are designed on the basis of specified causal knowledge. Hacking proposes an "entity realism" that denies any corresponding realism about theories. He claims that our ability to manipulate electrons depends only on using well-understood low-level causal properties, or home truths, as he calls them [*Ibid*.]. Although there may be a number of theories, models, formalisms and methods involving electrons there is no reason to suppose that the intersection of these is a theory: "even if there are a lot of shared beliefs there is no reason to suppose they form anything worth calling a "theory" [*Ibid*.].

Hacking's remarks on theories are somewhat puzzling. Surely the causal properties of electrons, regardless of whether or not they can be considered low-level, presuppose some account of electron theory. If the difficulties he envisions stem from a proliferation of theoretical accounts then he is correct in thinking that we ought not be realists about theories. However, if there is a set of common beliefs then this would suggest that there is at least some theory-neutral account of the basic properties of the electron. However, theory-neutrality entails only that the specified properties are not ones that enable us to decide between competing theoretical accounts. There may not be one theory of the electron, but everyone uses some theory in the description of its causal properties. Perhaps Hacking wants to claim that because these "facts" about electrons have become so well-entrenched they cease to be theoretical. But this is not an argument against a realism about theories, rather it is only because the theory itself has become generally accepted and is no longer subjected to question that we are less mindful of it. Such is the case for many of our "theories" about macro objects. However, if this is

Hacking's position then he seems committed to an experimental/theoretical distinction not unlike the observable/unobservable distinction that he criticizes van Fraassen for holding.

If for the sake of argument we were willing to grant that Hacking's "home truths" [p.265] about electrons and their causal properties do not constitute a "theory" then it would seem that the only way these truths could be initially established is by inference on the basis of our engineering successes in building machines that utilize the specified causal properties. Hacking claims that we do not infer the reality of electrons after we make the instruments; in designing the apparatus we rely on these home truths [p.265]. But again this is to dismiss the problem of how the knowledge of causal properties is established. However, it could be argued that Hacking is simply uninterested in this question. One way of characterizing his position is to see it as a transcendental argument. Given that the practice of physics consists (in part) of building very complex machinery what must we presuppose in order to make this process an intelligible one? The very fact that we build these machines commits us to a belief in the entities that the machine is designed to manipulate. Hence, the position addresses the grammar of manipulation instead of speaking to confirmation theory and accompanying epistemological concerns. The difficulty with this argument is that the very process of manipulation is immersed in theory. That we are manipulating entities like electrons is itself a theory and as such it is at most empirically adequate.

Consider for a moment the very straightforward process of detecting electrons in a cloud chamber. The cloud chamber itself is nothing more than a box in which air can be supersaturated with water vapours. Any disturbance of the air will result in the water vapour condensing to form droplets. Because an electron is a charged particle if it is travelling fast enough it can knock other electrons off gas atoms: that is, it can ionize them. Hence an electron moving in a cloud chamber leaves a trail of ions which act as disturbances around which water droplets form. This in turn causes one to see a cloudy trail in the chamber. Furthermore one can determine whether the particle in the chamber is an electron by observing how it moves. Because electrons are charged they experience forces in the presence of electric fields causing their tracks to look a certain way. These "facts" about electrons figure importantly in establishing the causal properties that Hacking speaks about. However, although they are commonly known characteristics of the electron we nevertheless require a scaffolding of theory in order to provide such information and in the design and operation of the cloud chamber. Electromagnetic theory tells us the conditions under which electrons can be stripped off atoms, how to use electric and magnetic fields to force electrons into the cloud chamber and whether the electron will ionize the gas atom. It also tells us how to produce a field in the chamber and together with mechanics tells us the shape of the electron's path and cloud trail. Mechanics also provides the information about the electron's collision with the gas atom. This together with other information from heat theory and atomic theory allows us to make claims about electrons and how they behave in specific circumstances.

If we insist on being anti-realists about these theories while maintaining a realism about electrons it is difficult to articulate what this realism would commit us to: what kinds of claims could be made about the entities in question. On the other hand, if we acknowledge that the causal properties of electrons are specified by the aforementioned theories then we seem forced to accept these theories as true. Our knowledge about entities, even the so-called experimental ones, and their causal properties is inextricably linked to the theories that specify those properties. In summary then given that manipulation need only presuppose an empirically adequate theory one can quite consistently acknowledge the importance of the process while remaining agnostic about the entities involved. However, adopting a realist position advocating belief in entities also requires belief in some theory even if only a low level one. This in turn involves difficulties associated with realism about theories; the epistemological problem of distinguishing between those theoretical claims that can be classified as "home truths" and those that cannot. Regardless of whether Hacking's account is intended as a kind of transcendental argument or as a pragmatic realism analagous to Moore's proof of the external world, it fails to escape or answer the problems posed by theory realism.

Hacking argues that if we look at the practice of science rather than being concerned with philosophical questions about theory verification we can see that traditional accounts of the relationship between theory and experiment are essentially misguided. However, it isn't immediately clear that scientific practice is faithful to Hacking's model. A particularly interesting feature of experimental and theoretical science is that the so called "reality" of many of the postulated entities is established through a kind of bootstrapping procedure: a method that is heavily dependent on theory in order to get off the ground.

Consider the following situation. Earlier I mentioned the difficulties surrounding the experiments determining quark charge. Failure to isolate quarks led to the confinement hypothesis. Although the quark model had been extremely successful in explaining various data the quark model of hadron masses incorporated assumptions that were incompatible with special relativity. Primarily on the basis of symmetry considerations the existence of a fourth heavy quark called charm was postulated. It was difficult to see how the existence of a fourth quark would solve the field theoretic problem. However, the crucial feature of the charmonium model was the difference in weight between the charmed quarks and the other three, namely the up, down and strange particles.* According to the model the quarks are so heavy that they moved nonrelativistically as opposed to the original quark model of hadrons which postulated relativistic motion. In analogy with the atomic system of an electron orbiting a positron (its antiparticle) the charmed quark was thought to (nonrelativistically) orbit its antiparticle under the influence of a central potential. However, as discrepancies between data and predictions occurred hyperfine splittings were introduced into the potentials to accommodate the recalcitrant data. If it was in fact the case that these charmed quarks were heavy then one ought to be able to give as much credence to the charmonium spectrum as to the atomic spectrum. And, in a classic inference to the best explanation, if the charmonium model fit the data then it seemed reasonable to accept charmed quarks as real entities on a par with electrons and positrons. Moreover, because charmed quarks were simply a new flavour of the original lighter quarks the latter ought to have the same status. As it turned out the charmonium model did agree with much of the experimental data and its overwhelming success endowed the original quark model with even greater reality. Although the reality of quarks was assumed in the search for charm nevertheless each contributed to the overall success of the other. There is a sense in which this example seems to grant Hacking his point. Quarks are real otherwise the search for charmed quarks makes no sense. Although the bootstrapping process enhances the status of the entities as "real" and allows us to take into account the causal properties of some entities in the discovery of others. It is important to point out that many of the entities involved in such procedures are themselves hypothetical, their reality is assumed for the purposes of experimentation. But, in accepting these entities as real we also accept a substantial amount of theory which figures in this bootstrapping process.

Although much of this discussion focuses on the independence of experiment from theory Hacking maintains that he is primarily interested in denying what he calls the "strong version" of a thesis about the theory-experiment relationship - that there must be a conjecture under test in order for an experiment to make sense. To the extent that Hacking's position involves claims about the new or anomalous phenomena then I am in complete agreement. However, quite often these phenomena stand as events or objects to be explained, and to the extent that we gain an understanding of them we do so when they are brought under the umbrella of some existing theory or when they spawn entirely new theories. As we saw in chapter three it was a lack of theoretical knowledge together with a failure to see Maxwell's theory as a viable option that resulted in much important experimental work going unnoticed. In fact much of the groundwork necessary for Hertz's experimental work on electromagnetism had been undertaken prior to 1886. It was the theoretical context in which Hertz's experiments were performed that rendered them significant. The idea that "much truly fundamental research preceeds any relevant theory whatsoever" [p. 158] seems true only to the extent that the experimental "doings" need not necessarily require agreement with the current body of theory in order for them to be sustained.*

As I understand it Hacking's position on the theory independence of experimental doings bears a certain similarity to the way in which most of us think of our interaction with the world of physical objects. To the extent that we accept the theory ladeness of observation we acknowledge the epistemic input involved in our characterizations of daily experience. And although we subscribe to certain implicit theories about the objects of experience our "doings" with respect to these objects are seen, for the most part, to be relatively theory-neutral. It would seem that Hacking is trying to characterize some experimental work in much the same way. Because of the way in which we "interact" with electrons they cease to be theoretical entities whose existence we acknowledge for the purposes of saving the phenomena [p. 262]. But can this comparison be successfully made? The answer has to do with the role and status of theory in each case. If one can successfully show, as I have tried to do, that our knowledge about electrons and other entites does depend on theory then the argument fails unless there is some account of how "home truths" come to achieve their status.

What then are the implications for an account of scientific realism based on experimentation or manipulation. The difficulties in sustaining such a position are many - the problem of competing intepretations about experimental data, the discrepancies involved in the relationship between evidence and theory; quite simply the fallibility and corrigibility of experimental results.[11] Both approximations and discrepancies were obvious in the confirmation of Maxwell's theory, hence one is unable to call the theory "true" in the traditional philosophical sense of word. The extent to which the limits of "reasonable agreement" between theory and experiment are the subject of pragmatic negotiation leaves in doubt the prospects for a strong form of scientific realism justified on the basis of experiment.

In conclusion it seems as though even experimental evidence as a whole is not the sort of thing that need compel belief; if belief or some degree thereof encompasses the idea of truth. This is not to say however that criteria enabling us to make choices between competing theories cannot be given. Franklin [1986] discusses a set of strategies which can supposedly provide independent and reasonable justification for rational belief in experimental results. Several of these strategies use the argument that observation of predicted behavior of known phenomena argues for the proper operation of the apparatus which in turn validates the observations or measurements. Other strategies include the use of different experiments to validate observations; all of which, together with an apparatus based on a wellcorroborated theory, serves to strengthen our belief in experimental results and the theories they confirm.

At the level of practice both Franklin's and Hacking's approaches are intuitively appealing. The difficulty arises when we attempt to forge an agreement between a theory or philosophy of science and the practice of science itself. At the level of epistemology, logic and semantics we seem unable to accommodate the intuitive notions of approximate truth and relative confirmation. Realist epistemology and logic deal with truth and falsity in a way that seems to rule out the kinds of approximations involved in scientific practice. Our scientific theories represent, more or less, our best efforts given the current state of knowledge. The special nature of science as a constant revisionary process poses problems for the kind of true belief characterisitic of realist epistemology.

In a recent paper entitled "Do Scientists Need a Philosophy?" Gerald Holton describes some attitudes characteristic of the scientific process. He quotes from a paper of Sheldon Glashdw's on the partial symmetries of weak interactions; a paper which would be the basis for his 1979 nobel prize:

> The mass of the charged intermediaries must be greater than zero, but the photon mass is zero - surely this is the principle stumbling block in any pursuit of the analogy between hypothetical vector bosons and photons. It is a stumbling block we must overlook.[11]

These kinds of speculative proposals are part and parcel of the risk taking nature of scientific activity: an activity which for the most part can be described only in pragmatic terms. This is not to say that all decisions regarding theories are exclusively pragmatic. There is a point at which some empirical evidence simply compels us to belief, where global skepticism becomes a meaningless issue if we are to engage in human inquiry. But as I mentioned above this is not characteristic of all experimental evidence. If we wish to develop an epistemology of science that adheres to traditional notions of truth and correspondence while providing an explanatory account of practice then the constraints of scientific realism render the position an unsuitable candidate.

Footnotes

1. Everitt, op. cit., p.102 remarks that Maxwell's decision to replace the vortex model of electromagnetic and optical processes as described in PL with an analysis of the relations between two classes of phenomena [DT and Treatise] was a concretization of William Hamilton's views on the relativity of knowledge; that all human knowledge is of relations between objects rather than objects themselves. Maxwell was very strongly influenced by Hamilton while he was a student at the University of Edinburgh. Hamilton's emphasis on the analogical or metaphorical aspects of the relational account of knwoledge is central to much of his work on electromagnetism and the kinetic theory. See also Olson, op. cit.

2. LeSage's aether model was the only prima facie adequate model of gravitational interaction to emerge from eighteenth century debates on the subject. According to LaSage's account gravity was explained by postulating the existence of streams of aethereal corpuscles flowing into the earth from all directions. Maxwell in his article on the atom in the Encyclopedia Britanicca [reprinted in Papers, 2:474] remarked if the theory could be shown to be consistent with the facts in other respects, it may turn out to be a "royal road into the very arcana of science". However, the model met with staunch opposition from many of LaSage's contemporaries. Boscovich, Euler and the french astronomer Bailly all objected to it on grounds that it postulated unobservable entities which were thought to be illegitimate in scientific inquiry. Cf. Laudan, "The Medium and its Message" in Cantor and Hodge, op. cit., pp.157-184, for details of the LaSage-Hartly debate. A similar situation occurred with Mach's rejection of the atomic hypothesis. The reality of atoms was rejected for philosophical (epistemological) rather than physical reasons.

3. H. Hertz, The Principles of Mechanics, D. Jones and J. Walley (trans.) New York: Dover, 1956, pp.7-8.

4. Although he made a commitment to neither Maxwell had proposed two seemingly contradictory representations of electricity. Hence the import of Hertz's remarks in PM. See chapter II above for a discussion of Maxwell's views on the subject.

5. Thomas Kuhn, The Essential Tension, Chicago: University of Chicago Press, 1977, pp.178-224.

6. For a discussion of these experiments see Andrew Pickering "Against Correspondence - The Constructivist View of Experiment and the Real" forthcoming in PSA Proceedings, Vol. 2, 1986 as well as "The Hunting of the Quark", Isis, Vol. 72, 1981, pp. 216-236. Pickering's book Constructing Quarks

Chicago: University of Chicago Press, 1984 provides an excellent discussion of the sociological factors relevant to discoveries in high energy physics.

7. Cf. Allan Franklin, "The Epistemology of Experiment" The British Journal for the Philosophy of Science, Vol. 35, 1984, pp.381-401 as well as a chapter with the same title in his recent book The Neglect of Experiment, Cambridge: Cambridge University Press, 1986.

8. The charmonium model referred to mesons that were made up of a charmed quark and a charmed anti-quark. It was so named because scientists saw it as analagous to the positronium system containing an electron orbiting a positron. For an interesting discussion on the development of the charm hypothesis see A. Pickering, "The Role of Interests in High Energy Physics" in K. Knorr (ed.) The Social Process of Scientific Investigation, Dordrecht: D. Reidel, 1981. See also Burton Richter, "From the psi to charm: The experiments of 1975 and 1976" Reviews of Modern Physics, Vol. 49, No. 2, April 1977, pp.251-266.

9. A good example is the Balmer formula for atomic spectra. At the time the formula was proposed it failed to be explicable by any of the current theories; yet it was used successfully for several years. It was not until Bohr's 1913 model of the atom that the Balmer formula was fully incorporated into a theoretical context.

10. For an excellent discussion of these problems see Allan Franklin, "Experiment and the Development of the Theory of Weak Interactions: Fermi's Theory", unpublished manuscript.

11. Cf. Gerald Holton, The Advancement of Science and its Burdens, Cambridge: Cambridge University Press, 1986, p.[illegible].

Bibliography

Achinstein, Peter (1968), Concepts of Science. (Baltimore: Johns Hopkins University Press).

Achinstein, Peter and Hannaway, Owen (1985), Observation, Experiment and Hypothesis in Modern Science. (Cambridge, MA: M.I.T Press).

Asquith, P. and Nickles, T., (eds.) (1982), P.S.A. 1982, Volume 2. (East Lansing, MI: Philosophy of Science Association).

Balzer, W.; Pearce, D.A.; and Schmidt, H. J. (1984), Reduction in Science: Structure, Examples and Philosophical Problems. (Dordrecht: D. Reidel Publishing Co. Inc.).

Bork, Alfred (1963), "Maxwell, Displacement Current and Symmetry", American Journal of Physics 31: 854-59.

-------- (1967), "Maxwell and the Vector Potential", Isis 58: 210-22.

Boyd, Richard (unpublished), Realism and Scientific Methodology.

-------- (1985), "Lex Orandi est Lex Credendi" in Images of Science, P. Churchland and C.A. Hooker (eds.). (Chicago: University of Chicago Press) pp.3-34.

Bromberg, Joan (1968), "Maxwell's Displacement Current and his Theory of Light", Archive for the History of the Exact Sciences 4: 218-34.

-------- (1968), "Maxwell's Electrostatics", American Journal of Physics 36: 142-51.

Brush, Stephen (1976), The Kind of Motion We Call Heat. Books 1 and 2. (Amsterdam: North Holland Press).

-------- (1983), Statistical Physics and the Atomic Theory of Matter. (Princeton: Princeton University Press).

Bunge, Mario (1957), "Lagrangian Formulation and Mechanical Interpretation", American Journal of Philosophy 25, 4 (April): 211-218.

Butts, R.E. (1973), "Whewell's Logic of Induction", in Foundations of Scientific Method: The Nineteenth Century. R. Giere and R. Westfall (eds.). (Bloomington: Indiana University Press) pp. 53-85.

________ (1977), "Consilience of Inductions and the Problem of Conceptual Change in Science", in Logic, Laws and Life. (Pittsburgh: Pittsburgh University Press) pp.71-88.

Campbell, Lewis and Garnett, William (1884), The Life of James Clerk Maxwell. (London: Macmillan and Co.).

Cantor, G. and Hodge, M. (1981), Conceptions of the Ether. (Cambridge: Cambridge University Press).

Cartwright, Nancy (1983), How the Laws of Physics Lie. (Oxford: Oxford University Press).

Chalmers, A.F. (1973), "The Limitations of Maxwell's Electromagnetic Theory", Isis 64: 464-83.

________ (1973), "Maxwell's Methodology and his Application of it to Electromagnetism", Studies in the History and Philosophy of Science 4, 2: 107-64.

Collie, C.H. (1982), Kinetic Theory and Entropy. (New York: Longman Group).

Crowe, Michael (1985), A History of Vector Analysis. (New York: Dover Books).

Churchland, Paul and Hooker, C.A. (eds.) (1985), Images of Science. (Chicago: University of Chicago Press).

D'Agostino, Salvo (1975), "Hertz's Researches on Electromagnetic Waves", Historical Studies in the Physical Sciences 6: 261-323.

Demopoulos, William (1982), "Review of The Scientific Image", Philosophical Review 91: 603-7.

Doran, Barbara (1975), "Field Theory in Nineteenth Century Britain", Historical Studies in the Physical Sciences 6: 133-260.

Dupre, John (1983), "The Disunity of Science", Mind 92, 367 (July): 321-346.

Everitt, C.W.F. (1974), James Clerk Maxwell: Physicist and Natural Philosopher. (New York: Scribner's).

Feyerabend, Paul (1981), Realism, Rationalism and Scientific Method: Philosophical Papers, Volume I. (Cambridge: Cambridge University Press).

Fisch, Menachem (1985), "Whewell's Consilience of Inductions", Philosophy of Science 52, 2 (June): 239-255.

Franklin, Alan (1979), "The Discovery and Nondiscovery of Parity Non-conservation", Studies in the History and Philosophy of Science 10, 3: 201-57.

________ (1986), The Neglect of Experiment. (Cambridge: Cambridge University Press).

________ (unpublished a), "Experiments, Theory Choice and the Duhem-Quine Problem".

________ (unpublished b), "Experiments and the Development of the Theory of Weak Interactions: Fermi's Theory".

Friedman, Michael (1983), Foundations of Space-Time Theories. (Princeton: Princeton University Press).

Gitterman, M. and Halpern, V. (1981), Qualitative Analyses of Physical Problems. (New York: Academic Press).

Glazebrook, R.T. (1901), James Clerk Maxwell and Modern Physics. (London: Cassell and Co. Ltd.).

Glymour, C. (1980), "Explanation, Tests, Unity and Necessity", Nous 14: 31-50.

________ (1980), Theory and Evidence. (Princeton: Princeton University Press).

Hacking, Ian. (1983), Representing and Intervening. (Cambridge: Cambridge University Press).

Harman, P. M. (1982), Energy, Force and Matter. (Cambridge: Cambridge University Press).

________ (1985), "The Natural Philosophy of James Clerk Maxwell", in Wranglers and Physicists. P. Harman (ed.). (Manchester: Manchester University Press), pp. 202-24.

Heimann, P. M. (1970), "Maxwell and the Modes of Consistent Representation", Archive for the History of the Exact Sciences 6: 171-213.

________ (1971), "Maxwell, Hertz and the Nature of Electricity", Isis 62: 149-157.

Hertz, Heinrich (1896), Miscellaneous Papers. D.E. Jones (trans.). (London: Macmillan and Co.).

________ (1956), The Principles of Mechanics. D.E. Jones and J.T. Walley (trans.). (New York: Dover Books).

________ (1962), Electrical Waves Being Researches on the Properties of Electric Action with a Finite Velocity Through Space. D.E. Jones (trans.). (New York: Dover Books).

Hesse, M. (1965), Forces and Fields. (Totowa, NJ: Wittlefield, Adams and Co.).

________ (1968), "Consilience of Inductions", in The Problem of Inductive Logic., I. Lakatos (ed.). (Amsterdam: North Holland Press) pp. 232-46.

________ (1971), "Whewell's Consilience of Inductions and Predictions", The Monist 55: 520-24.

________ (1980), "The Hunt for Scientific Reason", in P.S.A. 1980, Volume II., P. Asquith and T. Nickles (eds.). (East Lansing, MI: Philosophy of Science Association) pp. 3-21.

________ (1973), "Logic of Discovery in Maxwell's Electromagnetic Theory", in Foundations of Scientific Method., R. Giere and R. Westfall (eds.). (Bloomington: Indiana University Press) pp.86-114.

Holton, Gerald (1986), The Advancement of Science and its Burdens. (Cambridge: Cambridge University Press).

Hooker, C.A. (1975), "Global Theories", Philosophy of Science 42: 152-79.

Howson, Colin (1976), Methods and Appraisal in the Physical Sciences. (Cambridge: Cambridge University Press).

Hartkamper, A. and Schmidt, H.J. (eds.) (1981), Structure and Approximation in Physical Theories. (New York: Plenum Press).

Jones, L.W. (1977), "A Review of the Quark Search Experiments", Reviews of Modern Physics 49, 4 (October): 717-52.

Kinchin, A.I. (1949), Mathematical Foundations of Statistical Mechanics. (New York: Dover Books).

Klein, Martin (1974), "The Historical Origins of the Van Der Waals Equation", Physica 73: 28-47.

Knudsen, Ole (1976), "The Faraday Effect and Physical Theory 1845-1873", Archive for the History of the Exact Sciences 15: 235-281.

Kuhn, Thomas (1977), The Essential Tension. (Chicago: University of Chicago Press).

Lanczos, Cornelius (1970), The Variational Principles of Mechanics. (Toronto: University of Toronto Press).

Laudan, Larry (1971), "William Whewell on the Consilience of Inductions", The Monist 65, 3: 368-91.

________ (1981), "The Medium and its Message", in Conceptions of the Ether. G. Cantor and M. Hodge (eds.). (Cambridge: Cambridge University Press) pp.157-86

Lodge, Oliver (1907), Modern Views on Electricity. (London: Macmillan and Co.).

Longair, M.S. (1984), Theoretical Concepts in Physics. (Cambridge: Cambridge University Press).

Maxwell, James Clerk (1884), "Are There Real Analogies in Nature", in The Life of James Clerk Maxwell. L. Campbell and W. Garrett (eds.). (London: Macmillan and Co.) pp. 235-44.

________ (1954), Treatise on Electricity and Magnetism. 2 volumes. (New York: Dover Books).

________ (1965), The Scientific Papers of James Clerk Maxwell. W. D. Niven (ed.). (New York: Dover Books).

McCormach, Russell (1972), "Heinrich Hertz", in Dictionary of Scientific Biography. C. Gillespie (ed.). (New York: Scribner's) pp.340-350.

Morrison, P. and Morrison, E. (1957), "Heinrich Hertz", Scientific American (November): 98-106.

Moyer, Donald (1978), "Continuum Mechanics and Field Theory: Thompson and Maxwell", Studies in the History and Philosophy of Science 9: 35-50.

________ (1977), "Energy, Dynamics, Hidden Machinery: Rankine Thomson and Tait, Maxwell", Studies in the History and Philosophy of Science 8: 251-68.

Nagel, Ernest (1979), The Structure of Science. (Indianapolis: Hackett Publishing Co.).

Nickles, Thomas (1973), "Two Concepts of Intertheoretic Reduction", Journal of Philosophy 70, 7 (April): 181-201.

Olsen, Richard (1975), Scottish Philosophy and British Physics 1750 - 1880. (Princeton: Princeton University Press).

O'Rahilly, Alfred (1965), Electromagnetic Theory. 2 volumes. (New York: Dover Books).

Pickering, Andrew (1980), "The Role of Interests in High Energy Physics", in The Social Process of Scientific Investigation.. K.D. Knorr (ed.). (Dordrecht: D. Reidel Co.).

-------- (1984) "Against Putting the Phenomena First", Studies in the History and Philosophy of Science. 15, 2: 85-117.

-------- (1984), Constructing Quarks. (Chicago: University of Chicago Press).

-------- (forthcoming), "Experiments and the Real", in PSA 1986. A. Fine and P. Machamer (eds.). (Lansing, MI: Philosophy of Science Association).

Putnam, H. (1975), Mind, Language and Reality: Philosophical Papers, Volume II. (Cambridge: Cambridge University Press).

-------- (1978), Meaning and The Moral Sciences. (Boston: Routledge and Kegan Paul).

Richter, Burton (1977), "From Psi to Charm: The Experiments of 1975 and 1976", Review of Modern Physics 49, 2 (April): 251-66.

Salmon, Wesley (1970), "Bayes Theorem in the History of Science" in Historical and Philosophical Perspectives of Science. R. Stewer (ed.). (Minneapolis: University of Minnesota Press) pp. 66-86.

-------- (1984), Scientific Explanation and the Causal Structure of the World. (Princeton: Princeton University Press).

Segre, Emilio (1983), From Falling Bodies to Radio Waves. (New York: W.H. Freeman and Co.).

Shaffner, Kenneth (1967), "Approaches to Reduction", Philosophy of Science 34, 2 (June): 137-47.

-------- (1972), Nineteenth Century Aether Theories. (New York: Permagon Press).

Siegel, Daniel (1975), "Completeness as a Goal in Maxwell's Electromagnetic Theory", Isis 66: 361-68.

________ (1981), "Thomson, Maxwell and the Universal Aether in Victorian Physics", in Conceptions of the Ether. G. Cantor and M. Hodge. (Cambridge: Cambridge University Press) pp.239-68.

________ (1985), "Mechanical Image and Reality in Maxwell's Electromagnetic Theory", in Wranglers and Physicists. P. Harman (ed.). ((Manchester: Manchester University Press) pp. 180-201.

Simpson, T.K. (1966), "Maxwell and the Direct Experimental Test of his Electromagnetic Theory", Isis 57, 190: 411-432.

________ (1970), "Some Observations on Maxwell's Treatise on Electricity and Magnetism", Studies in the History and Philosophy of Science 1, 3: 249-63.

Spector, Marshall (1978), Concepts of Reduction in Physical Science. (Philadelphia: Temple University Press).

Stokes, Sir George (1907), Memoirs and Scientific Correspondence. J. Larmour (ed.). (Cambridge: Cambridge University Press).

Suppes, P. (1961), "A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences", in The Concept and Role of the Model in Mathematics in the Natural and Social Sciences. H. Freudenthal (ed.). (Dordrecht: D. Reidel Publishing Co.) pp.163-177.

Suskind, Charles (1964), "Observations of Electromagnetic Wave Radiation Before Hertz", Isis 55, 179: 33-42.

Tabor, David (1979), Gasses, Liquids and Solids. (Cambridge: Cambridge University Press).

Tricker, R.A.R. (1966), The Contributions of Faraday and Maxwell to Electrical Science. (New York: Permagon Press).

Turner, Joseph (1955), "Maxwell on the Method of Physical Analogy", British Journal for the Philosophy of Science 6: 226-38.

________ (1956), "A Note on Maxwell's Interpretation of Some Attempts at Dynamical Explanation", Annals of Science 11, 3: 238-45.

Van Fraassen, Bas (1980), The Scientific Image. (Oxford: The Clarendon Press).

Whewell, W. (1847), The Philosophy of the Inductive Sciences Founded on their History, Second Edition, 2 volumes. (London: John W. Parker).

-------- (1968), William Whewell's Theory of Scientific Method. R.E. Butts (ed.). (Pittsburgh: Pittsburgh University Press).

Williams, L. Pearce (1980), The Origins of Field Theory. (New York: University Press of America).