On Some Patterns of Reduction

Clark Glymour

Philosophy of Science, Vol. 37, No. 3 (Sep., 1970), 340-353.

Stable URL:
http://links.jstor.org/sici?sici=0031-8248%28197009%2937%3A3%3C340%3AOSPOR%3E2.0.CO%3B2-C

Philosophy of Science is currently published by The University of Chicago Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ucpress.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Mon Jun 13 12:07:35 2005

ON SOME PATTERNS OF REDUCTION*

CLARK GLYMOUR

Princeton University

The notion of reduction in the natural sciences has been assimilated to the notion of inter-theoretical explanation. Many philosophers of science (following Nagel) have held that the apparently ontological issues involved in reduction should be replaced by analyses of the syntactic and semantic connections involved in explaining one theory on the basis of another. The replacement does not seem to have been especially successful, for we still lack a plausible account of inter-theoretical explanation. I attempt to provide one.

1. The Problem of Reduction. Whether or not the notion of reduction in the natural sciences should be assimilated to the notion of inter-theoretical explanation, it has been. Ernest Nagel, [9], was perhaps the first to insist that the apparently ontological issues involved in reduction should be replaced by an analysis of the syntactic and semantic connections involved in explaining one theory from another, and most philosophers of science, even his critics, have followed his lead. Thus far, the replacement does not seem to have been especially successful, for we still lack a plausible account of intertheoretical explanation. I shall try to provide one.

"The objective of reduction," writes Professor Nagel, ([9], p. 301), "is to show that the laws or general principles of the secondary science are simply logical consequences of the assumptions of the primary science."[1] Almost all contemporary discussions of inter-theoretical explanation have assumed that Professor Nagel is correct, or nearly so. Unfortunately, the stubborn fact is that if the aim of reduction is what Nagel says it is, then reduction is nearly always impossible. The paradigm cases of inter-theoretical explanation involve theories which are inconsistent; special relativity, we think, explains Newtonian mechanics, but the two theories are inconsistent; statistical mechanics explains thermodynamics, but again the two theories are inconsistent. The incompatibility between the aim of reduction according to Nagel and the actual relations among scientific theories has been brought home most forcefully by Paul Feyerabend, who concluded his criticism of the logical empiricists with the announcement that "... a formal and 'objective' account of explanation cannot be given" ([4], p. 95). Logical empiricists themselves, however, had noticed the difficulty. For example, after presenting a purely deductive account of reduction Kemeny and Oppenheim ([7], p. 17) make the following rather remarkable statement:

> ... the old theory usually holds only within certain limits, and even then only approximately. For example, in the reduction of Kepler's laws to Newton's we must restrict ourselves to the case of a large central mass with sufficiently small masses, sufficiently far

* Received June, 1968.

[1] Nagel uses "secondary science" or "secondary theory" for the theory which is explained, "primary theory" for the theory which does the explaining. I shall generally follow this usage, although sometimes I shall refer to the primary and secondary theories as the explanans and explanandum, respectively.

apart, around it. And even then the laws hold only approximately—as far as we can neglect the interaction of the planets. While these points are of fundamental importance, there is no way of taking them into account as long as we tacitly assume that our theories are correct. If we abandon this (contrary-to-fact) assumption then the problem of reduction becomes hopelessly complex.

Several writers, including Putnam [11] and Hempel [6], have dealt with the difficulty by claiming that the aim of reduction is not the deduction of the laws of the secondary theory from the primary theory but rather the deduction of laws which approximate those of the secondary theory. In some sense this is doubtless correct, but it leaves a great deal unsaid. When, for example, do one set of laws approximate another set of laws? Is no structural connection between two theories required in order for one of them to explain the other? Is it possible to have two theories each of which explains the same empirical phenomena but such that neither theory explains the other? Certainly, the notion of an approximate deduction is not by itself an adequate reply to Professor Feyerabend's challenge.

My intent is to show that if one gives up Nagel's view of the aim of inter-theoretical explanation, then it is possible to give a "formal" account of important aspects of that enterprise. Several years ago Wilfrid Sellars ([13], p. 71), remarked that "... theories about observable things do not explain empirical laws, they explain why observable things obey, to the extent that they do, these empirical laws." Professor Sellars' remark is incomplete, but it provides the orientation which is essential for understanding how one theory explains another theory or law. Inter-theoretical explanation is an exercise in the presentation of counterfactuals. One does not explain one theory from another by showing why the first is true; a theory is explained by showing under what conditions it *would be* true, and by contrasting those conditions with the conditions which actually obtain.

There are two aspects of explanation which it is customary and useful to separate —the syntactic, and the semantic. Syntactically, we are concerned with presenting the structural connections between two theories which arise when one theory explains the other. Semantically, we are concerned with the meaning relations among terms in the two theories. My primary concern in this essay is to characterize the structural connections arising in inter-theoretical explanation, but I will have something to say about semantics as well.

Syntactically, inter-theoretical explanation consists of the specification of certain conditions and the demonstration that under these conditions one theory entails a collection of sentences which are "isomorphic" to the laws of the other theory. How does this specification of conditions occur, and what is the nature of this isomorphism? I have found, in the physical sciences, two clear, and I think predominant, methods of generating sentences isomorphic to those of the secondary theory. The first is simply to conjoin some special assumptions to the primary theory, so that the special assumptions together with the laws of the primary theory entail a collection of sentences isomorphic to the laws of the secondary theory. Thus in explaining laws of physical chemistry from statistical mechanics, one might assume that molecules have certain properties, which in fact they generally do not have. Or, in explaining a law from Newtonian mechanics, one

might assume that only certain special forces act, although in fact the situation is much more complex. The special assumptions provide explicit conditions under which the secondary law or theory would hold—and thus one of the functions of inter-theoretical explanation is removal of *ceteris paribus* clauses which experience makes us impose on our laws and theories.[2] The explanatory role of such special assumptions is not, however, exhausted by the fact that they provide conditions under which the secondary theory would hold; equally important is the fact that these conditions are not satisfied in general. Explanation involves a contrast, usually implicit, between the contrary-to-fact special assumptions which do entail laws isomorphic to those of the secondary theory, and the true special assumptions for various cases—which generally entail the negations of the isomorphs of the laws of the secondary theory.

The second syntactic device is the taking of limits. Very often, the laws of the primary theory are not isomorphic to those of the secondary theory, but nonetheless contain parameters such that, in the limit as these parameters approach certain values, the equations for the laws of the primary theory do become isomorphic to those of the secondary theory. The explanatory function of such limiting procedures is evident: the secondary theory is exhibited as an approximation which literally holds only as a limiting case. That the secondary theory is generally false is explained by the fact that the various parameters do not usually—and sometimes never—have their limiting values.[3] These two devices, limiting procedures and special assumptions, are not mutually exclusive; the theorist who sets about explaining one theory from another will generally use them both in combination with deductive inference. Nor do I claim that these are all of the syntactic devices he will use; it is only that I have not found any others.[4]

One's view of the nature of the "isomorphism" between the sentences generated from the primary theory and the laws of the secondary theory will depend on his view of what theories are. If one views a scientific theory as at least a collection of sentences in principle expressible within some formalized language—and I do so view them—then the obvious characterization of this isomorphism is in terms of a definitional extension. Let us suppose for generality that the vocabularies of the primary theory, T, and the secondary theory, T', are disjoint, and let us suppose that in attempting to explain T' a collection of sentences, T^*, is generated from T. T^* is supposed to have some structural similarity with T'. The structural similarity, I suggest, consists in the possibility of introducing a collection of definitions of the vocabulary in T' in terms of the vocabulary of T, such that T^*, together with the definitions, entails T'. It should be emphasized that this is a purely syntactic relation, although for first-order languages it has a semantic correlate via Beth's theorem. I will not argue for this account of the structural connection, primarily because I would not know how to, short of attempting a formalization of several theories.

[2] For further observations on the connections between reduction and *ceteris paribus* clauses, see L. Sklar, [14].

[3] Leonard Nash ([10], pp. 59–60), has noted the role of limiting procedures in inter-theoretical explanation.

[4] But see the conclusion to this essay.

In sum, the formal aspect of inter-theoretical explanation consists first, in the connection of terms in the secondary theory with those in the primary theory by means of syntactic definitions; and second, in the generation from the primary theory, by means of deduction, special assumptions, limiting procedures, and possibly other devices, of a collection of sentences which, together with the definitions, entails the secondary theory.

There is not usually a unique explanation of a given secondary theory or law from a given primary theory. Any of several special assumptions, or limiting procedures, will often suffice to provide conditions under which the secondary theory would hold. Generally, this plurality of explanations does no harm, but there is one case where it is troublesome. Sometimes we can explain the secondary theory in more than one way by changing the definitional extension used. Insofar as reduction has any ontological significance, this latter kind of plurality is inadmissible. If we think that the properties characterized by the secondary theory are caused by the properties postulated by the primary theory, or if we think that both theories are describing the relations among some of the very same properties and things (but the secondary theory describes them incorrectly), then there should be a unique admissible correlation of the vocabularies of the two theories. Later in this essay I will suggest some semantic and methodological considerations which I think play a significant role in eliminating such explanatory ambiguities. Primarily, however, the rest of this essay is concerned with illustrating both the conterfactual character, and the syntactic patterns, of the explanation of laws and theories.

2. Explaining Laws: Some Examples. Rather than attempting to formalize laws and theories in order to exhibit syntactical connections which arise when limits are taken or when special assumptions are made, I will content myself with presenting informal evidence for such connections. Thus instead of presenting van der Waals' and the ideal gas laws as sentences in a first-order language and showing that under some conditions one is derivable from a definitional extension of the other, it seems satisfactory enough to point out that the equation for the van der Waals' gas law

$$P = \frac{RT}{\overline{V} - b} - \frac{a}{\overline{V}^2}$$

becomes identical in form with the equation for the ideal gas law

$$P = \frac{RT}{\overline{V}}$$

when $\overline{V} \rightarrow \infty$, a and b constant. This "identity in form" simply consists in an identification of combinations of symbols in one law or theory with symbols, or combinations thereof, in another law or theory, the identification being made in such a way that mathematical operations among symbols are preserved, as is the equality relation. Part of the reason for this informality is that I do not want the account of reduction I am proposing to depend on any particular attempt at a

formalization of laws in a first-order language—or even on the assumption that a first-order language is sufficient for expressing scientific laws.

The van der Waals'-ideal-gas-law example illustrates the explanatory role of limiting procedures in an especially simple way. If van der Waals' is true, then the ideal gas law would only hold were a system to become arbitrarily dilute; the ideal gas law fails for all actual systems *because* such systems are not composed of molecules which are point masses and *because* there are forces of attraction among the molecules of a gas. The ideal gas law, according to van der Waals, should work rather well when the ratios V/b and V/a are very large, and rather poorly when one or both of the ratios is small. Part of the theoretical role of the constants a and b, which disappear on taking limits, is to help explain why the ideal gas law works where it works and fails where it fails.

One of the most important empirical relations of classical thermodynamics is the law of corresponding states. Given a gas system in which the temperature, pressure, and volume can be varied, it is found that for various combinations of these parameters the gas will be in equilibrium with its liquid; for other combinations the liquid disappears and the system becomes homogeneous. For each gas it is found that there is a unique maximum pressure, P_c; and maximum temperature, T_c, at which the liquid and vapor can coexist. Corresponding to these values is a unique value of the volume V_c. Thus for each gas there is a set of values, P_c, T_c, V_c, of relevant variables. A system which has these values is said to be in a critical state. We can define pseudo-variables, called reduced variables, by dividing the gas variables by their critical values. Thus

$$\pi = P/P_c; \qquad \tau = T/T_c; \qquad \chi = \overline{V}/\overline{V}_c$$

Of course P_c, T_c, and $\overline{V}_c$ will be different for different gases. The law of corresponding states asserts that if any two gases have the same values for any two reduced variables, then they will have the same value for the third reduced variable.

In other words the law claims that for all gases

$$\pi = f(\tau, \chi); \qquad \tau = g(\pi, \chi); \qquad \chi = h(\pi, \tau)$$

where f, g, h are the same functions no matter what gas is considered.

The law of corresponding states provides a challenge to any attempt at reducing gases to other entities. If gases are collections of molecules, then a molecular theory ought to explain why the law of corresponding states works so well, and where its limitations are.

A formal analogue of the law of corresponding states can be derived from statistical mechanics *if it is assumed* that the potential energy of interaction of the N molecules of the system can be expressed as a sum of pair potentials and the pair potential $\mu(r)$ can be expressed as a universal function ϕ, containing two scale factors ϵ and σ, dependent on the substance in question. We have

$$\mu(r) = \epsilon\phi\left(\frac{r}{\sigma}\right)$$

where r is the intermolecular distance. Now this assumption is reasonable if the

molecules in question are either spherical, or are effectively spherical because of rotation, and if the molecules are not polar. From other evidence we can determine which substances have these properties, and thus can give a detailed account of just where and why the law of corresponding states fails or succeeds ([3], pp. 333–337).

Even in the explanation of relatively simple empirical laws, there arise cases which require the use of both special assumptions added to the principles of the explaining theory and the use of limiting procedures. The Newtonian explanation of Galileo's law of falling bodies and the classical thermodynamic explanation of of the law of mass action both require such a combination.

Let us take Galileo's law as the claim that the acceleration of a body falling near the surface of the earth is constant. To Newton's law we shall add the postulates that the only force acting on a falling body is the Earth's gravitational attraction and that the Earth is a sphere. Then, with these extra assumptions, we have for a falling body at height h above the earth:

$$F = ma = G\frac{mM}{(R+h)^2}$$

where M is the mass of the earth and R is the radius of the earth. Hence in the limit as $h \to 0$ we have

$$a = G\frac{M}{R^2}$$

and the expression on the right is a constant.

Thus Galileo's law is an approximation which *would* approach the Newtonian truth as a falling body comes arbitrarily close to the surface of the earth, *if* all forces other than the gravitational attraction of the earth were negligible and if the earth were spherical. Galileo's law fails in fact *because* the earth is not spherical and because forces other than the gravity of the earth are not zero and because the gravitational force is a function of distance. In the explanation of why Galileo's law fails one is not simply committing the fallacy of denying the antecedent. Rather, one is implicitly contrasting a contrary-to-fact situation in which Galileo's law would hold with the real situation, in which Newton's laws entail the denial of Galileo's law—or at least the denial of a formal analogue of that law.

The law of mass action, whether for gases or for liquid solutions, is not a logical consequence of classical thermodynamics. In the case of gases, for example, one can nonetheless obtain a formal analogue of the law of mass action by first adding the assumption that the pressure is constant and then taking the limit as the system becomes arbitrarily dilute, so that fugacities can be replaced by partial pressures.[5] Examples could be multiplied, but perhaps enough has been said to make the case that many explanations of empirical laws are of a counterfactual kind, and do illustrate the pattern I have outlined.

3. Explaining Theories. The syntactical aspect of explaining one theory from

[5] For details see any text on physical chemistry, e.g., [2], Chap. 11.

another generally consists in generating from the primary theory, by means of special assumptions, limiting procedures, and deductive inference, a collection of sentences which has a definitional extension which entails the secondary theory. The only consistency condition which it seems necessary to impose on this process is that if A and B are theorems of the secondary theory, and A entails B, then the same procedures which generate an analogue of A from the primary theory be used to generate the analogue of B. It is simply a matter of keeping our counterfactuals consistent. There is, in general, nothing to prevent us from explaining one law of the secondary theory from one set of special assumptions and limits and another, logically independent, law from quite different—conceivably even inconsistent—special assumptions and limits. Thus to a large extent, the syntactical aspect of explaining one theory from another can be analyzed in terms of explanations of the laws of the secondary theory. Of course, however the sentences analogous to the laws of the secondary theory happen to be generated, a unique correspondence must be established between quantities in the secondary theory and quantities, or combinations of quantities, in the primary theory.

Any number of theories might be used to exhibit the roles of special assumptions and limiting procedures in explaining one theory from another. There is, for example, the standard textbook remark that Newtonian mechanics is a limiting case of the special theory. The example I should like to consider, however, is the one which seems to have been most discussed by philosophers, namely the statistical mechanical explanation of classical thermodynamics. Since the most rigorous presentation of classical thermodynamics of which I am apprised, Carathéodory's [1], is far from an axiomatization in the modern sense, there can be no question of presenting or examining a formal explanation of the whole of thermodynamics. Rather, I will content myself with an informal consideration of the connection between statistical mechanics and particular phenomenological laws, a procedure which I hope is at least partially justified by my preceding remarks. The bone of contention among philosophers has not been the first law or the so-called zeroth law or the third law (which can't be obtained from classical statistical mechanics in any case), but the statistical mechanical explanation of the phenomenological second law.

One of the earliest and clearest modern attempts at a mechanical explanation of thermodynamics is contained in J. Willard Gibbs' *Elementary Principles in Statistical Mechanics*. Gibbs' book is an attempt not only to develop the statistical mechanical ideas of Boltzman, Maxwell, and himself, in a unified way, but also "... to find in rational mechanics an *a priori* foundation for the principles of thermodynamics" ([5], p. 165). The Kantian notions aside, there can be no question but that Gibbs was attempting a statistical mechanical explanation of classical thermodynamics. Let us consider what he does.

Gibbs deals with mechanical systems with $2n$ degrees of freedom, subject to conservative external forces which are functions of external coordinates a_i. The potential energy is assumed to be a function of the internal coordinates $q_1, \ldots, q_n$ and the external coordinates. He constructs a $2n$-dimensional phase space in the usual way and considers the distribution D, which he calls the "density in-phase,"

of a space ensemble in the phase space. The systems in the ensemble are assumed to be subject to the same force laws, and at any given instant the values of the external coordinates a_i are the same for all systems in the ensemble.[6]

Gibbs defines the probability of finding a system in a given volume of phase space as

$$\int \cdots \int P\, dp_1 \ldots dq_{2n}$$

where the p_i are generalized moments, $P = D/N$ and N is the number of systems in the ensemble. The problem of choosing an explicit function for P still remains. Since he has already shown that for a system in statistical equilibrium—and that is the case in which he is interested—P is a constant of the motion of the ensemble, Gibbs chooses P to be a function of the energy, and writes for the canonical ensemble

$$P = \exp\left(\frac{\psi - \epsilon}{\theta}\right)$$

or

$$n = \ln P = \frac{\psi - \epsilon}{\theta},$$

where ϵ is the energy, ψ and θ are constants.

To motivate the assumption that θ is proportional to the thermodynamic temperature, Gibbs proves a statistical analogue of the classical principle that systems at the same temperature and at thermal equilibrium remain in thermal equilibrium when they are coupled thermally—provided there are no forces of interaction. That is, Gibbs shows that if two systems have the same modulus, θ, and are in statistical equilibrium then they remain in statistical equilibrium when coupled, provided again that the interaction forces are negligible.

To get to the second law, Gibbs defines the average value of a quantity, u, in the usual way:

$$\bar{u} = \iint_{\text{all phases}} u \exp\left[\frac{\psi - \epsilon}{\theta}\right] dp_1 \ldots dq_{2n}$$

and then argues that the following relation holds:

$$d\bar{\epsilon} = \theta\, d(-\bar{n}) - \bar{A}_1\, da_1 - \bar{A}_2\, da_2 - \cdots$$

He remarks that:

> This equation, if we neglect the sign of the averages, is identical in form with the thermodynamic equation . . .
>
> $$d\epsilon = T\, dn - A_1\, da_1 - A_2\, da_2 - \text{etc.}$$
>
> which expresses the relation between the energy, temperature, and entropy of a body in thermodynamic equilibrium, and the forces which it exerts on external bodies—a relation which is the mathematical expression of the second law of thermodynamics for reversible changes. The modulus in the statistical equation corresponds to temperature in the thermodynamic equation, and the average index of probability *with its sign reversed* corresponds to entropy ([5], p. 44).

[6] Understandably, in view of the ergodic problem, Gibbs is rather vague concerning the connection between space ensembles and time ensembles.

Gibbs finds the analogy unsatisfactory for semantic reasons—the average value of the energy is not the energy, and so on. Hence he attempts to find conditions under which ϵ will equal $\bar{\epsilon}$. To complete the connection, Gibbs argues that the "anomaly" of the energy, i.e. the mean square deviation, is inversely proportional to the number of degrees of freedom of the system, provided that the quantity $d\bar{\epsilon}_q/d\bar{\epsilon}_p$ (where ϵ_q is the potential energy due to the internal coordinates, ϵ_p the kinetic energy) does not increase without bound as the number of degrees of freedom increases.

In the limit, then, as the number of degrees of freedom of a system becomes arbitrarily large, Gibbs feels justified in equating, *within statistical mechanics*, the energy of a system with the average value of the energy in a canonical ensemble of such systems, and similarly for other physical quantities. For the limiting case, Gibbs then identifies the statistical mechanical energy and index of probability with the classical thermodynamic energy and entropy. The semantic considerations are important here, and I will say something about them subsequently, but for now I am concerned only with the restrictions they impose on the syntactical connection between statistical mechanics and the second law. Evidently, the connection, according to Gibbs, is that the classical second law is obtained as a limiting case from statistical mechanics. In his own words:

> a very little study of the statistical properties of conservative systems of a finite number of degrees of freedom is sufficient to make it appear, more or less distinctly, that the general laws of thermodynamics are the limit toward which the exact laws of such systems approximate, when their number of degrees of freedom is indefinitely increased ([5], p. 166).

One can criticize Gibbs' rigor, but there seems little question that what he was attempting to do corresponds with the explanatory schemes I have outlined. It is worth noting that in the most rigous presentation of classical statistical mechanics available, A. I. Khinchin comes to the same conclusion as Gibbs. Khinchin shows, by applying the central limit theorem, that in the limit as the number of partitions of a system becomes arbitrarily large, one obtains a precise analogue of the classical second law ([8], esp. Chap. 7).

Sometimes different explanations of the secondary theory can be generated from the same primary theory by making different identifications of quantities in the two theories and thus we are faced with the task of providing criteria for choosing among alternative explanations. Here, however, a simple and obvious principle comes to our aid. Let A and B be two explanations of a theory T from a theory T'; if every law of T explained by B is also explained by A, but there are laws of T explained by A and not by B, then we prefer explanation A.

The electrodynamic explanation of physical optics illustrates the preceding principle rather nicely. Polarization experiments show that the optical disturbance associated with a light wave is a vector quantity and the optical vector is perpendicular to the direction of propagation of the wave. Now in electrodynamics the electric intensity vector E and the magnetic intensity H are both vector quantities associated with an electromagnetic wave, both quantities satisfy a wave equation, and both are perpendicular to the direction of propagation of an electromagnetic wave. It would seem, therefore, that two different explanations of physical optics from

electrodynamics are possible: one which identifies the optical vector with E, and one which identifies the optical vector with H. Nonetheless, the optical vector is always identified with E, never with H. The reason is that there are laws within physical optics which cannot be explained if the optical vector is identified with H, but which receive a very neat explanation if the optical vector is identified with E.

When a beam of linearly polarized light meets a surface of separation of two transparent media, part of the beam is reflected and part refracted. It is found experimentally that the relative intensities of the reflected and refracted beams depends on the angle of incidence and on the angle between the plane of motion of the optical vector and the plane of incidence. When the optical vector is parallel to the plane of the surface, the intensity of the reflected beam is a maximum and the intensity of the refracted beam is a minimum; the reverse holds when the optical vector, the direction of propagation, and a normal to the surface all lie in the same plane. In the latter case, when the optical vector lies in the plane of incidence, there is an angle of incidence, called Brewster's angle or the polarization angle, for which the beam is totally refracted. The angle is given by

$$\theta = \frac{n_2}{n_1}$$

where n_1 and n_2 are the indices of refraction of the two media. This relation is called Brewster's law.

Now let θ be the angle of incidence, and θ' the angle of refraction of a linearly polarized electromagnetic wave. Then it follows from Maxwell's equations[7] that if E lies in the plane of incidence, then the amplitude of the *reflected* beam is given by

$$A'' = -\frac{\tan(\theta - \theta')}{\tan(\theta + \theta')} A$$

where A is the amplitude of the incident beam. When $\theta + \theta' = \pi/2$ the denominator goes to infinity and the amplitude of the reflected wave becomes zero. Hence when E lies in the plane of incidence there is an angle for which the incident wave is totally refracted, and, therefore, if E is identified with the optical vector we obtain an explanation of Brewster's law. No such explanation is possible if the optical vector is identified with H.

Ernest Nagel has claimed that identification of terms across theories are synthetic material claims. And, it would seem, one way of looking at Brewster's law is as empirical confirmation of the identification of E with the optical vector. The clear moral is that an adequate explanation of a theory should not only explain the fundamental principles of the theory, but should explain as well any and all "empirical" relations among the quantities, objects, and processes postulated by the theory.

4. Explanations and Analogies. Considerations of meaning are of obvious importance in inter-theoretical explanation; we have seen, for example, their importance

[7] [12], Chap. 8. I am indebted to a referee for suggesting the need to treat this case.

in Gibbs' attempt to explain thermodynamics. The sorts of syntactical connections I have attempted to describe are far from sufficient for explanation. After all, formal analogies abound in physics—for example in potential theory and between mechanical systems and electrical circuits—but no one thinks such analogies amount to an explanation of anything. I very much doubt that our understanding of the semantic aspects of explanation is furthered by grand debates about "meaning invariance" and the like. Instead of joining the lists on one side or the other, I want to suggest, rather tentatively, a consideration which seems to me of some help in understanding the difference between explanations and formal analogies.

Given a theory, there are various measurements we can make, or procedures we can carry out, to determine the value of a quantity used in the theory for some systems, or to determine whether or not parts or aspects of a system exhibit properties or stand in relations postulated by the theory. I will call any such procedure a measurement, even though that is a bit misleading. Now I want to distinguish two kinds of measurements associated with a theory: those which use some of the laws of the theory and those which do not. I do not know how to make this distinction clear except by example. Suppose we are dealing with a law, say a gas law such as the ideal gas law, which employs the notion of temperature. We can measure temperature by means of a thermometer, or by resistence measurements, or by a thermocouple, and so on. None of these measurement procedures uses the ideal gas law. By contrast we might measure the temperature of a system with a contraption which determines the volume of the system, by means of meter rods say, and determines the pressure of the system, perhaps with a manometer, and then *computes* the temperature from these data according to the ideal gas law. Such a procedure uses the ideal gas law in an essential way. Again, when doing calorimetric measurements of heats of reaction, one generally makes temperature measurements and then computes the heat of reaction by means of a thermodynamic law connecting the heat of reaction with the temperature values and empirical constants, such as the heat capacity of the calorimeter. As a measurement of heat of reaction such a procedure makes essential use of the thermodynamic law. A measurement which uses a law cannot be used to test that same law—it cannot, just because the measurement is performed in such a way as to *guarantee* that the results will accord with the law. Since I will often want to refer to measurement procedures which measure a quantity occurring in a theory without using any laws of that theory, I shall call such measurements *primary* for that theory.

The distinction which I am attempting to draw is not the same as the traditional distinction between fundamental and derived measurements or magnitudes. For one thing, fundamental magnitudes are often characterized as additive, or extensive, so that temperature or density cannot be fundamental magnitudes; but additivity has nothing to do with whether or not a quantity occurring in a theory admits of measurements which are primary for that theory. Further, the distinction between fundamental and derived measurements is usually drawn in an absolute way, whereas a measurement is primary or not only relative to a specific theory. Thus any measurement of a quantity occurring in thermodynamics, such as pressure, which uses only mechanical laws is primary for thermodynamics. Pressure,

temperature, volume, electrical work, all these are quantities which can occur in thermodynamics and which admit of measurements which are primary for thermodynamics. Entropy, by contrast, is a thermodynamic quantity which admits of no primary measurement—at least I know of no way of measuring the entropy of a system without using some law of thermodynamics. Again, in mechanics position and time are obviously quantities which admit measurements which are primary for mechanics. For all practical purposes the distinction can often be made historically: measurement procedures which are primary for a theory tend to be procedures which were used prior to the introduction of the theory and the use of which is taken over by the theory.

According to the account I have given, the explanation of one theory by another requires that quantities in the secondary theory be identified with quantities, or combinations of quantities, in the primary theory. In terms of the distinction just drawn, an obvious semantic restriction on such identifications is this: The identification must be made so that every primary measurement of a quantity in the secondary theory is also a measurement of the corresponding quantity, or combination of quantities, in the primary theory.

It should be evident why it is that only primary measurements of a quantity in the secondary theory need be measurements of the corresponding quantity in the primary theory. Measurements which are not primary for the secondary theory use some of the laws in the secondary theory; but the laws of the secondary theory are generally inconsistent with the laws of the primary theory, and hence measurements which use the laws of the secondary theory will not be satisfactory measurements of quantities in the primary theory. It is difficult to justify the preceding analysis by considering examples, for in scientific writings semantic relations are seldom stated explicitly. Perhaps, nonetheless, something can be done to show how the above restriction can be used to help rationalize the moves Gibbs makes in his attempt to explain thermodynamics.

I will assume that Gibbs *intended* that the modulus θ in the canonical distribution should be determined by the same primary measurements as determine the temperature, T, and, moreover, that he intended that primary measurements of the quantities of work, A_i, in thermodynamics should also be measurements of the quantities A_i in statistical mechanics. These intentions, according to the semantic criterion proposed, already severely limit the admissible identifications of quantities in the two theories.

Gibbs, it will be recalled, derives an equation from statistical mechanics

$$d\bar{\epsilon} = \theta\, d(-\bar{n}) - \bar{A}_1\, da_1 - \bar{A}_2\, da_2 - \cdots$$

which is formally analogous to a theorem of classical thermodynamics combining the first and second laws. Nonetheless he is unwilling to identify $\bar{\epsilon}$ with ϵ, and so on; he only allows the identification in the limit as the number of degrees of freedom of a system approaches infinity. The identification of $\bar{\epsilon}$ with ϵ and $\bar{n}$ with n is not permissible because it forces the identification of A_i in thermodynamics with $\bar{A}_i$ in statistical mechanics, and a primary measurement of work is *not* a measurement of the average value of the work for a canonical ensemble. In the limit, however, as

the number of degrees of freedom becomes arbitrarily large, the "anomaly" of the work, just like the anomaly of the energy, becomes arbitrarily small. In the limit, then, a measurement of work done by a statistical mechanical system will be a measurement of the average value of the work done by systems in a canonical ensemble, and vice versa. In this case, then, the identification of thermodynamic energy and entropy with statistical mechanical energy and index of probablility results in laws with the same structure.

In sum, we can rationalize Gibbs' procedure in this way: To explain thermodynamics from statistical mechanics, Gibbs had to show that, given certain identifications of the quantities in one theory with those in the other, there are specifiable conditions under which the laws of statistical mechanics entail the laws of thermodynamics. He intended that primary measurements of thermodynamic temperature and work should also be measurements of the statistical mechanical modulus and work, respectively. In accord with the proposed semantic criterion, this intention limited the admissible identifications of quantities in the two theories to those which identified T with a multiple of θ and A_i in one theory, with A_i in the other. *In particular this intention excluded the possibility of exhibiting thermodynamics as a deductive consequence of statistical mechanics—without any special conditions or limits—via an identification of thermodynamic quantities with the average values of their statistical mechanical analogues.* By contrast, the identification of the thermodynamic energy and entropy with the energy and index of probability in the canonical ensemble was compatible with Gibbs' intent and permitted—or so Gibbs thought—the derivation of thermodynamic laws from statistical mechanics in the limiting case.

Conclusion. Even if everything I have had to say here is true, which is not very likely, there still remain major aspects of inter-theoretical explanation which are not well understood. There is, of course, the question of the meaning relations among different theories. Here I can only urge that the counterfactual character of inter-theoretical explanation be kept in mind in subsequent analyses. An explanation of a theory or law does not generally proceed by showing why the theory or law is true, but rather by arguing that the theory or law is false *because* something or other is the case, or that the theory or law *would be true if* something or other were the case. Equally important in understanding explanation is the notion of approximation. A great many inter-theoretical arguments involve such familiar moves as the replacement of a sum by an integral, the deletion of terms in a power series, the separation of variables which are not separable. A detailed account of the warrants for such moves, and of how inconsistency is avoided in using them, is still needed.

REFERENCES

[1] Carathéodory, C., "Untersuchungen über die Grundlagen der Thermodynamik," *Math. Ann.*, vol. 67, 1909.
[2] Castellan, G., *Physical Chemistry*, Addison-Wesley, 1964.
[3] Davidson, N., *Statistical Mechanics*, McGraw-Hill, 1962.

[4] Feyerabend, P., "Explanation, Reduction and Empiricism," in *Minnesota Studies in Philosophy of Science* (eds. H. Feigl and G. Maxwell), vol. III, University of Minnesota, 1962.
[5] Gibbs, J., *Elementary Principles in Statistical Mechanics*, Dover, 1960.
[6] Hempel, C., *Aspects of Scientific Explanation*, The Free Press, 1965.
[7] Kemeny, J. and Oppenheim, P., "On Reduction," *Phil. Studies*, vol. VII, 1956 (Jan.).
[8] Khinchin, A., *Mathematical Foundations of Statistical Mechanics*, Dover, 1949.
[9] Nagel, E., "The Meaning of Reduction in the Natural Sciences," in *Philosophy of Science* (eds. A. Danto and S. Morgenbesser), World Publishing Co., 1960.
[10] Nash, L., *The Nature of the Natural Sciences*, Little Brown, 1963.
[11] Putnam, H., "How Not to Talk about Meaning," in *Boston Studies in the Philosophy of Science* (eds. R. Cohen and M. Watofsky), Humanities Press, 1965.
[12] Rossi, B., *Optics*, Addison-Wesley, 1957.
[13] Sellars, W., "The Language of Theories," in *Current Issues in the Philosophy of Science* (eds. H. Feigl and G. Maxwell), Holt Rinehart and Winston, 1961.
[14] Sklar, L., *Intertheoretical Reduction in Natural Science*, unpublished doctoral dissertation, Princeton, 1964.