

Unifying Scientific Theories

Physical Concepts and Mathematical Structures

MARGARET MORRISON

University of Toronto



CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK <http://www.cup.cam.ac.uk>
40 West 20th Street, New York, NY 10011-4211, USA <http://www.cup.org>
10 Stamford Road, Oakleigh, Melbourne 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain

© Margaret Morrison 2000

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2000

Printed in the United States of America

Typeface Garamond Book 10/12 pt. *System* $\text{\TeX}$ 2 ϵ [TB]

A catalog record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data

Morrison, Margaret, 1954–

Unifying scientific theories : physical concepts and mathematical
structures / Margaret Morrison.

p. cm.

Includes index.

ISBN 0-521-65216-2 (hbk.)

1. Unified field theories. I. Title.

QC794.6.G7M67 1999

530.14'2 –dc21 99-20874

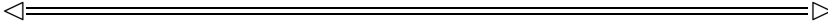
CIP

ISBN 0 521 65216 2 hardback

Contents

<i>Acknowledgements</i>	<i>page vii</i>
Introduction	1
1 The Many Faces of Unity	7
1.1 Kepler: Unity as Mathematical Metaphysics	7
1.2 Kant: Unity as a Heuristic and Logical Principle	12
1.3 Whewell: Unity as Consilience and Certainty	16
1.4 Logical Empiricism: Unity as Method and Integration	22
1.5 Unity as Explanation	25
2 Unification, Realism and Inference	35
2.1 The Friedman Model	38
2.2 The Importance of Conjunction	39
2.3 Reduction versus Representation	43
2.3.1 Is Reduction a Viable Approach?	43
2.3.2 The Problem of Many Models	47
2.4 Consilience and Unification	52
2.5 Unification as an Evidential or Epistemic Virtue	57
Appendix Derivation of the van der Waals Law: Historical Details	59
3 Maxwell's Unification of Electromagnetism and Optics	62
3.1 Development of Electromagnetic Theory: The Early Stages	64
3.1.1 Maxwell circa 1856: "On Faraday's Lines of Force"	64
3.1.2 "On Physical Lines of Force"	68
3.2 Unification and Realism: Some Problems for the Electromagnetic Theory	79
3.3 The Electromagnetic Theory: Later Developments	81
3.3.1 The Dynamical Theory	81
3.3.2 The Treatise on Electricity and Magnetism	90
3.4 Realism and Dynamical Explanation	99
3.4.1 Philosophical Conclusions	105
4 Gauges, Symmetries and Forces: The Electroweak Unification	109
4.1 Fermi and Beta Decay	110
4.2 Symmetries, Objects and Laws	114

4.3	The First Steps to Unity: Symmetry and Gauge Theory	118
4.4	Unity through Symmetry-Breaking	121
4.5	Renormalization: The Final Constraint	127
4.6	When Unity Isn't Enough	130
4.7	Unified Theories and Disparate Things	135
Appendix 4.1	Weinberg's Lepton Model	140
Appendix 4.2	Renormalization	143
5	Special Relativity and the Unity of Physics	147
5.1	Electrodynamics circa 1892–1904: Lorentz and the Origins of Relativity	149
5.1.1	1895: <i>Versuch</i> and the Theory of Corresponding States	156
5.2	Einstein circa 1905: Lorentz Transformed	162
5.2.1	Synchronizing Clocks: The Here and Now	165
5.2.2	Deriving the Transformation Equations: The Interconnection of Space and Time	169
5.2.3	Synthetic Unity: From Form to Content	176
5.2.4	The Most Famous Equation in the World	179
5.3	From Fields to Tensors: Minkowski Space-Time	183
6	Darwin and Natural Selection: Unification versus Explanation	192
6.1	On the Nature of Darwin's Theory	193
6.2	Kitcher's Argument Patterns: Explanation as Derivation and Systematization	196
6.3	Methods and Causes: History and Influence	202
6.4	Does Unity Produce Explanation?	206
7	Structural Unity and the Biological Synthesis	210
7.1	Synthesis and Unity	210
7.2	Mathematics and Theory	213
7.2.1	Mendelian Ratios	213
7.2.2	Fisher's Idealized Populations	214
7.2.3	Wright's Path Coefficients	224
7.3	Disunity and Explanation	227
	Conclusions	232
	<i>Notes</i>	238
	<i>References</i>	257
	<i>Index</i>	267



The Many Faces of Unity

1.1. Kepler: Unity as Mathematical Metaphysics

In the *Mysterium cosmographicum* Johannes Kepler claimed that it was his intention to show that the celestial “machine” was not a kind of divine living being,

but a kind of clockwork insofar as the multiplicity of motions depends on a single, quite simple magnetic and corporeal force, just as all the motions of a clock depend upon a simple weight. And I also show that this physical cause can be determined numerically and geometrically. (Kepler 1938, xv:232)

His research began with a specification of certain astronomical hypotheses based on observation; that was followed by a specification of geometrical hypotheses from which the astronomical ones would follow or could be calculated. Those geometrical hypotheses were grounded in the idea that God created the solar system according to a mathematical pattern. Given that assumption, Kepler attempted to correlate the distances of the planets from the sun with the radii of spherical shells that were inscribed within and circumscribed around a nest of solids. The goal was to find agreement between the observed ratios of the radii of the planets and the ratios calculated from the geometry of the nested solids. Although unsuccessful, Kepler remained convinced that there were underlying mathematical harmonies that could explain the discrepancies between his geometrical theory and ratios calculated from observations.

Part of Kepler’s unfaltering reliance on mathematical harmonies or hypotheses was based on their direct relationship to physical bodies. He considered a mathematical hypothesis to be physically true when it corresponded directly to physically real bodies. What “corresponding directly” meant was that it described their motions in the simplest way possible. Hence, according to Kepler, physical reality and simplicity implied one another; and it was because nature loves simplicity and unity that such agreement could exist. (Here unity was thought to be simply a manifestation of nature’s ultimate simplicity.) Perhaps his most concise statement of the relationship between truth and simplicity or between the mathematical and the physical can be found in the *Apologia*, where Kepler distinguished between “astronomical” and “geometrical” hypotheses:

If an astronomer says that the path of the moon is an oval, it is an astronomical hypothesis; when he shows by what combination of circular movements such an oval orbit may be brought about, he is using geometrical hypotheses. . . . In sum, there are three things in astronomy: geometrical hypotheses, astronomical hypotheses, and the apparent motions of the stars themselves; and, consequently, the astronomer has two distinct functions, the first, truly astronomical, to set up such astronomical hypotheses as will yield as consequences the apparent motions; second, geometrical, to set up geometrical hypotheses of whatsoever form (for in geometry there may often be many) such that from them the former astronomical hypotheses, that is, the true motions of the planets, uncorrupted by the variability of the appearances, both follow and can be calculated.¹

One was able to discover the true motions of the planets by determining their linear distances and using simplicity as the guiding principle in interpreting the observations.

Much of his early work in constructing physical theories (before the development of his laws of planetary motion) was dominated by the desire to provide a unified explanation of the causes of planetary motion. The Neoplatonic sun, to which he added a force that pushed the planets along in their orbits, served as the primary model for his solar hypothesis. But the foundation for that hypothesis was the metaphysical principle that one ought to reduce several explanatory devices to a single source. That principle, in turn, was based on Kepler's ideas about the Trinity. The sun served as the principle that unified and illuminated matter in the way that the Trinity symbolized the indivisible, creative God. Kepler then transformed the theological analogy into a mathematical relation in which solar force, like the light in a plane, was assumed to vary inversely with distance. The idea was that there existed one soul at the centre of all the planetary orbits that was responsible for their motions. God the Father created spirit in the same way that the sun dispersed spirit, and the sun emitted a moving force in the ecliptic in accordance with the same mathematical function as light propagating in a plane.

The important relation here, of course, was between mathematical simplicity and unity and the way in which those notions were used to both construct and justify astronomical hypotheses. As mentioned earlier, there was a direct relation between the symmetry of the mathematical relations used to describe physical bodies and the metaphysical underpinnings of those relations found in the Trinity. In his account of the interspacing of solid figures between planetary spheres, Kepler claimed that it ought to follow perfectly the proportionality of geometrical inscriptions and circumscriptions, and "thereby the conditions of the ratio of the inscribed to the circumscribed spheres. For nothing is more reasonable than that the physical inscription ought exactly to represent the geometrical, as a work of art its pattern" (1938, vi:354). And in analogy with the Trinity, he remarked that

there exists everywhere between point and surface the most absolute equality, the closest unity, the most beautiful harmony, connection, relation, proportion and commensurability.

And, although Centre, Surface and the Interval are manifestly Three, yet they are One, so that no one of them could be even imagined to be absent without destroying the whole. (Kepler 1938, vi:19)

Here we see an explicit statement of how unity and simplicity could be, in some cases, manifestations of the same thing. The unifying axiom that the planets were united by a single force, rather than a multiplicity of planetary “souls” acting in isolation, was, of course, also the simplest hypothesis. Hence, simplicity and unity were represented as oneness. In other contexts, however, unity and simplicity were related to each other via a kind of interconnectedness, the one as a manifestation of the many. For Kepler, the latter was apparent in the notion of the Trinity, but we can perhaps see it more clearly in the idea of a nation-state that embodies many people and perhaps many cultures, all of which are united in one identity – citizens of that state. It was that combination of unity and simplicity as a form of *interconnectedness* that provided the empirical basis on which Kepler’s astronomical hypotheses were justified.

Although Kepler saw the truth of a physical or astronomical hypothesis as metaphysically grounded in its simplicity or unity, the latter also had to be revealed empirically. Not only did the phenomena have to be describable using mathematically simple relations, but the interconnectedness among those descriptions had to be manifest at the empirical level in order for the hypothesis to be justified. Such was the case in Kepler’s famous argument for the elliptical orbit of Mars. Indeed, it was his belief that “physical” hypotheses regarding the quantifiable forces exerted by the sun on the motions of the planets could, in fact, be proved or demonstrated. And it was the idea that “one thing is frequently the cause of many effects” that served as the criterion for the truth or probability of a hypothesis, particularly in the *Astronomia nova*. The key to the argument in Kepler’s famous “war on Mars” was the geometrical relation that facilitated the combination of two quantifiable influences of the sun on the planet, the first being the planet’s orbit around the sun, and the second its libratory approach to and recession from the sun. Once those two were combined, Kepler could *justify* not only the elliptical orbit of Mars but also the fact that its motion was in accordance with the area law. The synthesis consisted in showing (1) that although libratory motion obeyed a law of its own, it was exactly because of the motion of libration that the planet described an elliptical orbit, and (2) that the second law or area law was valid only for an elliptical orbit. Kepler saw his argument as producing an integrated unity founded on mathematical simplicity. Let us look briefly at the physical details to see how they fit together.

Kepler’s dynamical account of libration was modelled on magnetic attraction and repulsion. In *Astronomia nova*, planetary motion was explained by the joint action of the sun and the planets themselves, whereas in his later work, the *Epitome*, the entire action was attributed to the sun. The motive radii of the sun’s species not

only led the planets around it but also repelled and attracted them depending on whether a planet displayed its “friendly” or “hostile” side toward the sun, that is, depending on which magnetic side was facing the sun. Kepler hypothesized that the source of that magnetism lay in magnetic fibres that passed through the planets. However, the planets themselves did not “exert” any force; rather, the action of the sun communicated a certain “inclination” to the fibres of the planetary body such that its entire libration derived from the sun. Hence, libration was not the result of any action or motion of the planet itself.² In order to give an exact account of the mechanism of orbital motion it would be necessary to determine the variations in the propelling and attracting forces throughout the entire path. That would require that one calculate the angles that the radius vectors of the sun made with the magnetic fibres of the planet. The sines of the complements of those angles (the cosines) would provide a measure of those portions of the forces that acted on the planet.³

Once Kepler developed the mechanism responsible for libration and proved that it was measured by the versed sine of the arc traversed by the planet, he was able to formally establish that an elliptical orbit resulted as a consequence of libration.⁴ From there Kepler went on to prove his second law, which describes the relationship between the time taken by a planet to travel a particular distance on its orbit and the area swept out by the radius vector. Again, the key to the synthesis was that the law, as Kepler formulated it, was valid *only* for an elliptical orbit, thereby establishing an interconnectedness between the dynamics of libratory motion and the geometry governing the motions of the planets. The relationships were all confirmed empirically, making the argument one that was not based solely on formal geometrical constraints, but one that united the physical and mathematical components of celestial phenomena in a simple coherent way. That unity created a kind of justification that not only applied to Kepler’s laws themselves but also extended to the metaphysical thesis regarding the relation between mathematical simplicity and truth. In other words, it was a justification determinable through the agreement of results. His laws governing planetary orbits described the simplest possible paths consistent with libratory motion, and the convergence of those results provided further evidence that his physical hypothesis was true. Kepler began with the belief that nature was founded on or determined by mathematical harmonies, and the correspondence of empirical observations with laws based on those harmonies further reinforced the idea of nature’s unity and mathematical simplicity.

It is exactly this kind of context, where one sees a particular law or principle yielding different kinds of interconnected results (i.e., the dynamics of libratory motion facilitating the derivation of Kepler’s first and second laws in a way that made each interdependent on the other), that we typically take as exemplifying at least some of the qualities of a unified theory.⁵ But in the assessment of such a theory, in the determination of its truth or confirmation value, one must be cautious in locating the truth component in the proper place. In other words, if we look at the

unity displayed by Kepler's account of planetary motion, it is tempting to describe its explanatory power as being grounded in the dynamics of libratory motion. And when Kepler succeeded in deducing an elliptical orbit from his physics, instead of just arriving at it through a process of observation coupled with manipulation of mathematical hypotheses designed to fit the facts, one is tempted to say that the physical basis of the theory must be true. Tempting as this may be, the question is whether or not we can, or even should, infer, as Kepler did, on the basis of this type of interconnectedness, the truth of the physical hypothesis.

With hindsight we know that Kepler's physics was mistaken, despite the fact that it yielded laws of planetary motion that were retained as empirically true approximations within Newtonian mechanics. The example, however, raises a number of issues that are important for understanding how one ought to think about unification. First, it seems that given what we know about the history of physics, it becomes obvious that one should not, as a general principle, attribute truth to a unifying hypothesis such as Kepler's theory of libratory motion simply because it yields a convergence of quantitative results. Second, and equally important, is the question of whether unification can be said to consist simply in a mathematical or quantitative convergence of different results or whether there needs to be an appropriate dynamical or causal explanation from which these results issue as consequences. The answer one gives to this latter question is significant not only for the link between explanatory power and unification but also for the connection between unified theories and the broader metaphysical thesis about unity in nature. That is, if unity is typically accompanied by an underlying physical dynamics, then it becomes necessary to determine whether or not the unifying power provides evidence for the physical hypothesis from which it emerges. Finally, it certainly is not an uncommon feature of scientific theories that they display, at least to some degree, the kind of interconnectedness present in Kepler's account of planetary motion. Yet his account is not typically thought to be a truly "unified" theory, in that it does not bring together *different* kinds of phenomena. For example, electromagnetic phenomena were thought to be radically different from optical phenomena until they were unified by Maxwell's theory and shown to obey the same laws. In fact, most instances of what we term "unification" are of this sort. Is it necessary, then, to specify particular conditions that must be satisfied in order for a theory to be truly unified, or is the notion of unification simply one that admits of degrees? In other words, do all theories unify to a greater or lesser extent? Each of these issues will be discussed in later chapters, but by way of contrast to Kepler's metaphysical/mathematical picture of unity and simplicity let us turn to Kant's account, which accords to unity a strictly heuristic role by characterizing it as a regulative ideal that guides our thinking and investigation about experience in general and scientific investigation in particular. What is interesting about Kant's notion of unity is that it carries with it no metaphysical commitments; yet it is indispensable for scientific research and the more general quest for knowledge.

1.2. Kant: Unity as a Heuristic and Logical Principle

Within the Kantian framework it is the faculty of reason that is responsible for synthesizing knowledge of individual objects into systems. An example is Kant's notion of the "order of nature", an entire system of phenomena united under laws that are themselves unified under higher-order laws. This systematic arrangement of knowledge is guided by reason to the extent that the latter directs the search for the ultimate conditions for all experience – conditions that are not, however, to be found within the domain of experience itself. That is, we could never unify all our knowledge, because such a grand unification could never be found in experience. Hence, the quest for unity is one that, by definition, is never fulfilled; it remains simply an ideal or a goal – in Kant's terms, a "problem" for which there is no solution. What reason does, then, is introduce as an ideal or an uncompletable task a set of rational conditions that must be satisfied for all of our knowledge to constitute a unified system. Examples of such conditions are (1) that we act as though nature constitutes a unified whole and (2) that we act as if it is the product of an intelligent designer. Consequently, this ideal *regulates* our search for knowledge and directs us toward a unified end. The fact that we can never achieve this complete unity should not and cannot be an obstacle to our constant striving toward it, for it is only in that striving that we can achieve any scientific knowledge.

To the extent that complete unity is not attainable, reason is said to function in a "hypothetical" way; the conditions referred to earlier take on the role of hypotheses that function as methodological precepts. Consequently, the systematic unity that reason prescribes has a *logical* status designed to secure a measure of coherence in the domain of empirical investigation. Kant specifically remarks that we would have no coherent employment of the understanding – no systematic classifications or scientific knowledge – were it not for this presupposition of systematic unity. But how can something that is in principle unrealizable, that is merely and always hypothetical, function in such a powerful way to determine the structure of empirical knowledge? Part of the answer lies in the fact that the search for unity is an essential logical feature of experience.

The notion of a *logical* principle serves an important function in the Kantian architectonic. Principles of reason are dependent on thought alone. The logical employment of reason involves the attempt to reduce the knowledge obtained through the understanding "to the smallest number of principles (universal conditions) and thereby achieve the highest possible unity" (Kant 1933, A305). Although we are required to bring about this unity in as complete a form as possible, there is nothing about a logical principle that guarantees that nature must subscribe to it. In that sense the logical employment and hypothetical employment of reason describe the same function. The logical aspect refers to the desire for systematic coherence, and the hypothetical component is a reminder that this ultimate unity as it applies to nature always has the status of a hypothesis. The principle that bids us to seek unity is necessary insofar as it is definitive of the role of reason in cognition; without it we would have no intervention on the part of reason and, as a result,

no coherent systematization of empirical knowledge. In other words, it is a necessary presupposition for all inquiry. And as a logical principle it specifies an ideal structure for knowledge in the way that first-order logic is thought to provide the structure for natural language.

One of the interesting things about the requirement to seek systematic unity is that it not only encompasses a demand for a unified picture of experience but also involves what Kant classifies as “subjective or logical maxims” – rules that demand that we seek not just homogeneity but also variety and affinity in our scientific investigations and classifications. These maxims are the principles of genera (homogeneity), specification (species) and continuity of form (affinity). Homogeneity requires us to search for unity among different original genera; specification imposes a check on this tendency to unify by requiring us to distinguish certain subspecies; and continuity, the affinity of all concepts, is a combination of the previous two insofar as it demands that we proceed from each species to every other by a gradual increase in diversity. Kant expands on this point in the *Jäsche Logic* (sec. 11), where he discusses the concepts “iron”, “metal”, “body”, “substance” and “thing”. In this example we can obtain ever higher genera, because every species can always be considered a genus with respect to a lower concept, in the way iron is a species of the genus metal. We can continue this process until we come to a genus that cannot be considered a species. Kant claims that we must be able to arrive at such a genus because there must be, in the end, a highest concept from which no further abstraction can be made. In contrast, there can be no lowest concept or species in the series, because such a concept would be impossible to determine. Even in the case of concepts applied directly to individuals, there may be differences that we either disregard or fail to notice. Only relative to *use* are there “lowest” concepts; they are determined by convention insofar as one has agreed to limit differentiation.

These logical maxims, which rest entirely on the hypothetical interests of reason, *regulate* scientific activity by dictating *particular* methodological practices. Again, this connection between logic and methodology is a crucial one for Kant. At the core of his view of science as a systematic body of knowledge lies the belief that science must constitute a logical system, a hierarchy of deductively related propositions in ascending order of generality. The act of systematizing the knowledge gained through experience enables us to discover certain logical relations that hold between particular laws of nature. This in turn enables us to unify these laws under more general principles of reason.

This classification process, which includes the unification of dissimilar laws and diversification of various species, exemplifies Kant’s *logical* employment of reason. A properly unified system exhibits the characteristics of a logical system displaying coherence as well as deductive relationships among its members. Scientific theories are themselves logical systems that consist of classificatory schemes that unify our knowledge of empirical phenomena. Kant recognizes, however, that reason cannot, simply by means of a logical principle, command us to treat diversity as disguised unity if it does not presuppose that nature is itself unified. Yet he claims that

the only conclusion which we are justified in drawing from these considerations is that the systematic unity of the manifold of knowledge of understanding, as prescribed by reason, is a *logical* principle. (Kant 1933, A648/B676)

This leaves us in the rather puzzling position of having logical or subjective maxims whose use is contextually determined, while at the same time upholding an overriding principle of unity in nature as prescribed by reason. In other words, Kant seems to sanction the idea of disunity while at the same time requiring that we seek unity. At A649 he discusses the search for fundamental powers that will enable us to unify seemingly diverse substances. Again the idea of such a power is set as a problem; he does not assert that such a power must actually be met with, but only that we must seek it in the interest of reason. As Kant remarks at A650/B678, “this unity of reason is purely hypothetical”. Yet in the discussion of logical maxims the principle of unity seems to take on a more prominent role. His example concerns a chemist who reduces all salts to two main genera: acids and alkalies. Dissatisfied with that classification, the chemist attempts to show that even the difference between these two main genera involves merely a variety or diverse manifestations of one and the same fundamental material; and so the chemist seeks a common principle for earths and salts, thereby reducing them to one genus. Kant goes on to point out that it might be supposed that this kind of unification is merely an economical contrivance, a hypothetical attempt that will impart probability to the unifying principle if the endeavour is successful. However, such a “selfish purpose” can very easily be distinguished from the *idea* that requires us to seek unity. In other words, we don’t simply postulate unity in nature and then when we find it claim that our hypothesis is true.

For in conformity with the idea everyone presupposes that this unity of reason accords with nature itself, and that reason – although indeed unable to determine the limits of this unity – does not here beg but command. (Kant 1933, A653/B681)

Put differently, the overall demand of reason to seek unity is the primary goal of all cognition in the attempt to reconstruct nature as a logical system. The mere fact that we engage in cognitive goals implicitly commits us to the search for unity. Within that context there are several different methodological approaches that can be employed for achieving systematic classification of empirical knowledge. Reason presupposes this systematic unity on the ground that we can conjoin certain natural laws under a more general law in the way that we reduce all salts to two main genera. Hence, the logical maxim of parsimony in principles not only is an economical requirement of reason but also is necessary in the sense that it plays a role in defining experience or nature as a systematically organized whole. Hence, what appear to be conflicting research strategies, as outlined by the subjective maxims, are simply different ways that reason can attain its end. For example, the logical principle of genera responsible for postulating identity is balanced by the principle of species, which calls for diversity; the latter may be important in biology,

whereas the former is more important for physics. But Kant is no reductionist; the idea of a “unified knowledge” is one that may consist of several different ways of systematizing empirical facts.

The logical maxims are not derived from any empirical considerations, nor are they put forward as merely tentative suggestions. However, when these maxims are confirmed empirically, they yield strong evidence in support of the view that the projected unity postulated by reason is indeed well-grounded. But in contrast to the strategy described earlier, the motivation behind the unifying methodology is not based on utilitarian considerations; it is not employed because we think it will be successful. Nevertheless, when we do employ a particular maxim in view of a desired end and are successful in achieving our goal, be it unity, specification or continuity, we *assume* that nature itself acts in accordance with the maxim we have chosen. On that basis we claim that the principles prescribing parsimony of causes, manifoldness of effects and affinity of the parts of nature accord with both reason and nature itself.

We must keep in mind, however, that although these principles are said to “accord with” nature, what Kant means is that although we must *think* in this way in order to acquire knowledge, there is also some evidence that this way of thinking is correct. The latter, however, can never be known with certainty, because we can never know that nature itself is constituted in this way. From the discussion of the logical employment of reason we know that in order to achieve the systematic unity of knowledge that we call science it is necessary that this unity display the properties of a logical system. In other words, if one agrees with Kant that science is founded on *projected* systematization and that this system is ultimately reducible to logical form (non-contradiction, identity and deductive closure over classification systems), then the principles that best cohere with the demand of systematic unity recommend themselves. Parsimony, manifoldness and affinity are not only methodological principles for organizing nature according to our interests; they are also the most efficient way of realizing the one interest of reason – the systematic unity of all knowledge. Because we empirically verify the extent to which this unity has been achieved, we are thereby supplied with the means to judge the success of the maxims in furthering our ends (Kant 1933, A692/B720), but ends that we, admittedly, never attain. We employ a particular maxim based on what we think will be the most successful approach in achieving systematic unity given the context at hand.

As mentioned earlier, the motivating idea for Kant is the construction of a logical system rather than the realization of a metaphysical ideal regarding the unity of nature. Kant is silent on the question of whether or not this notion of systematization constitutes the basis for scientific explanation. Although it seems clear that classification of phenomena does serve some explanatory function, there is nothing in the Kantian account of unity to suggest that it is in any way coincident with explaining or understanding the nature of phenomena. In essence, the Kantian account of unity constitutes a methodological approach that is grounded in the basic

principles of human reason and cognition. The unity has a hypothetical and pre-suppositional status; it is an *assumption* that the world is a unified whole, rather than a metaphysical principle stating how the world is *actually* structured. In that sense it is simply an idealization that is necessary for scientific inquiry.

Kant's views about the role of ideas in producing unity both in and for science were taken up in the nineteenth century by William Whewell. His views about unity as a logical system were also adopted, albeit in a different form, in the twentieth century by Rudolph Carnap. Unlike Kant, Whewell took a more substantive approach by linking his notion of unity (termed the consilience of inductions) to explanation by way of a set of fundamental ideas: Each member in the set of ideas would ground a particular science. Consequently, Whewell also adopted a much stronger epistemological position by claiming that unified or consilient theories would have the mark of certainty and truth.

1.3. Whewell: Unity as Consilience and Certainty

In the *Novum Organon Renovatum* William Whewell discusses various tests of hypotheses that fall into three distinct but seemingly related categories. The first involves the prediction of untried instances; the second concerns what Whewell refers to as the consilience of inductions; the third features the convergence of a theory toward unity and simplicity. Predictive success is relatively straightforward and encompasses facts of a kind previously observed but predicted to occur in new cases. Consilience, on the other hand, involves the explanation and prediction of facts of a kind different from those that were contemplated in the formation of the hypothesis or law in question. What makes consilience so significant is the finding that classes of facts that were thought to be completely different are revealed as belonging to the same group. This “jumping together” of different facts, as Whewell calls it, is thought to belong to only the best-established theories in the history of science, the prime example being Newton's account of universal gravitation. But Whewell wants to claim more than that for consilience; he specifically states that the instances where this “jumping together” has occurred

impress us with a conviction that the truth of our hypothesis is certain. . . . No false suppositions could, after being adjusted to one class of phenomena, exactly represent a different class, where the agreement was unforeseen and un contemplated. That rules springing from remote and unconnected quarters should thus leap to the same point, can only arise from that being the point where truth resides.⁶

Finally, such a consilience contributes to unity insofar as it demonstrates that facts that once appeared to be of different kinds are in fact the same. This in turn results in simpler theories by reducing the number of hypotheses and laws required to account for natural phenomena. Hence, unity is a step in the direction of the goal of ultimate simplicity in which all knowledge within a particular branch of science will follow from one basic principle.

Part of what is involved in a consilience of inductions is what Whewell refers to as “the colligation of ascertained facts into general propositions”. This also takes place through a three-step process that involves (1) selection of the idea, (2) construction of the conception and (3) determination of the magnitudes. These steps have analogues in mathematical investigations that consist in determining the independent variable, the formula and the coefficients. It was Whewell’s contention that each science had its own fundamental idea; the study of mathematics was based on the ideas of number and space, whereas mechanics relied on the idea of force to make it intelligible. Similarly, the idea of polarity was predominant in the study of chemical phenomena, and ideas of resemblance and difference were crucial to the study of natural history and the classificatory sciences.

Once the requisite idea was chosen, one could then proceed to the construction of a conception, which was a more precise specification of the idea. For example, a circle or a square is a kind of spatial configuration, and a uniform force is a particular manifestation of the general notion of a force. So if we have a phenomenon like the weather and we are trying to establish some order that will assist in predictions, we must decide whether we wish to select (1) the idea of time, and introduce the conception of a cycle, or (2) the idea of force, accompanied by the conception of the moon’s action. One selects the appropriate conception by comparing it with observed facts, that is, determining whether or not the weather really is in fact cyclical by comparing the supposed cycle with a register of the seasons. The idea is the core concept that grounds each field of inquiry. When we achieve a consilience of inductions, the result is that two different conceptions governing different classes of facts are seen in a new way, either as belonging to a totally new conception or as manifestations of one or the other of the original conceptions. The important point is that consilience does not involve a “jumping together” of two distinct ideas from different branches of science. The classes of facts usually are drawn from within an individual science.

There is no definitive method for selecting the right idea, nor the right conception for that matter. The only requirement or rule is that the idea must be tested by the facts. This is done by applying the various conceptions derived from the idea to the facts until one succeeds in uncovering what Whewell refers to as the “law of the phenomena”.

Although my intention is not to provide a complete analysis of Whewell’s account of induction and unification, I do think it important to discuss briefly the mathematical representation of this procedure, in an attempt clarify how unification takes place. The interesting question is whether the convergence of numerical results that occurs in a consilience of inductions is what ultimately constitutes unity or whether there are further implications for the supposed connection between explanation and unification.

In a section entitled “General Rules for the Construction of a Conception” Whewell describes the process as the construction of a mathematical formula that coincides with the numerical expression of the facts. Although the construction of

the formula and the determination of its coefficients have been separated into two steps, Whewell claims that in practice they are almost necessarily simultaneous. Once one selects the independent variable and the formula that connects the observations to form laws, there are particular technical processes whereby the values of the coefficients can be determined, thereby making the formula more accurate. These include the methods of curves, of means, of least squares and of residues. In the case of the method of curves, we have a specific quantity that is undergoing changes that depend on another quantity. This dependence is expressed by a curve. The method enables us to detect regularities and formulate laws based not only on good observations but also on those that are imperfect, because drawing a line among the points given by observations allows us to maintain a regular curve by cutting off the small and irregular sinuosities. When we remove the errors of actual observation by making the curve smooth and regular, we are left with separate facts corrected by what Whewell calls their “general tendency”; hence, we obtain data that are “more true than the individual facts themselves” (Whewell 1847, vol. 2). The obstacles that prove problematic for the method are ignorance of the nature of the quantity on which the changes depend and the presence of several different laws interacting with one another.

The method of curves assumes that errors in observation will balance one another, because we select quantities that are equally distant from the extremes that observation provides. In cases where we have a number of unequal quantities and we choose one equally distant from the greater and smaller, we use the method of means rather than the method of curves. The implicit assumption, again, is that the deviations will balance one another. In fact, the method of means is really just an arithmetical procedure analogous to the method of curves, with one significant difference: In the method of curves, observation usually enables us to detect the law of recurrence in the sinuosities, but when we have a collection of numbers we must divide them into classes using whatever selection procedures we think relevant.

The method of least squares is also similar to the method of means. It allows us to discover the most probable law from a number of quantities obtained from observation. The method assumes that small errors are more probable than larger ones, and it defines the best mean as that which makes the sum of the squares of the errors the least probable sum. Finally, the method of residues involves an analysis of unexplained facts that have been left as residue after the formation of a law governing changes of a variable quantity. The residue is analysed in the same way as the original observations until a law is found that can account for it. This continues until all the facts are accounted for.⁷

The notable feature present in these methods, and what is important for our purposes here, is the level of generality that is introduced in order to assist in the formulation of laws of phenomena. Although it is true that induction is the process by which one formulates a general proposition from a number of particular instances, the difference in these cases is that the general conclusion is not simply the result of a juxtaposition or conjunction of the particulars. In each case a

“conception” is introduced that is not contained in the bare facts of observations. The conception is the new fact that has been arrived at through a reinterpretation of the data using the relevant methods. This new element or conception can then be superimposed on existing facts, combining them in a unique way. Such was the case with the ellipse law governing the orbit of Mars. What Whewell describes are methods for data reduction that facilitate the formulation of a conception; but one need not employ all of these methods in order to arrive at a conception. For example, after trying both circular and oval orbits and finding that they did not agree with observations of the observed longitudes of Mars (or the area law), Kepler was led to the ellipse, which, taken together with the area law, gives the best agreement with the available observations. Some methods of data reduction were employed, because the object of the exercise was to find a structure that would fit with the observations. As we saw earlier, there was a convergence of numerical results in establishing the ellipse law, which led Kepler to believe that he had hit on the right formulation. Although we don’t have predictions of different *kinds* of data or classes of facts, as in the case of a true consilience, we do have better predictions for not only Mars but also Mercury and the earth. In that sense, then, there is a colligation of facts made possible by the introduction of the conception (i.e., the ellipse) based on the idea of space.

So the induction does not consist in an enumerative process that establishes a general conclusion; rather, the inductive step refers to the suggestion of a general concept that can be applied to particular cases and can thereby unify different phenomena. According to Whewell, this “general conception” is supplied by the mind, rather than the phenomena; in other words, we don’t simply “read off” the conception from the data. Rather, it requires a process of conceptualization. The inference that the phenomena instantiate this general conception involves going beyond the particulars of the cases that are immediately present and instead seeing them as exemplifications of some ideal case that provides a standard against which the facts can be measured. Again, the important point is that the standard is constructed by us, rather than being supplied by nature. That the conception presents us with an “idealized” standard is not surprising, because the mathematical methods used to arrive at it embody a great deal of generality – generality that obscures the specific nature of the phenomena by focusing instead on a constructed feature that can be applied across a variety of cases.⁸ It is this issue of generality that I want to claim is crucial not only to the unifying process but also to the connection (or lack thereof) between unification and explanation. My focus is not so much the notions of data reduction, as described by Whewell, but more general mathematical techniques used to represent physical theories. The importance of calling attention to Whewell’s methods is to emphasize the role of mathematics generally in the formulation of specific hypotheses. The more general the hypothesis one begins with, the more instances or particulars it can, in principle, account for, thereby “unifying” the phenomena under one single law or concept. However, the more general the concept or law, the fewer the details that one can infer about the

phenomena. Hence, the less likely it will be able to “explain” how and why particular phenomena behave as they do. If even part of the practice of giving an explanation involves describing how and why particular processes occur – something that frequently requires that we know specific details about the phenomena in question – then the case for separating unification and explanation becomes not just desirable but imperative.

It has been claimed by Robert Butts, and more recently by William Harper, and even by Whewell himself, that in a consilience there is an *explanation* of one distinct class of facts by another class from a separate domain.⁹ However, it is important here to see just what that explanation consists in. As Butts has pointed out, we cannot simply think of the explanatory power of consilience in terms of entailment relations, because in most cases the deductive relationship between the consilient theory and the domains that it unifies is less than straightforward. The best-known example is Newton’s theory and its unification of Kepler’s and Galileo’s laws. That synthesis required changes in the characterization of the nature of the physical systems involved, as well as changes in the way that the mathematics was used and understood, all of which combined to produce nothing like a straightforward deduction of the laws for terrestrial and celestial phenomena from the inverse-square law. Given that consilience cannot be expressed in terms of entailment relations, is it possible to think of the connection between explanatory power and consilience in terms of the convergence of numerical results? Such seems the case with Maxwellian electrodynamics, in which calculation showed that the velocity of electromagnetic waves propagating through a material medium (supposedly an electromagnetic aether) had the same value as light waves propagating through the luminiferous aether. That coincidence of values suggested that light and electromagnetic waves were in fact different aspects of the same kind of process. However, as we shall see in Chapter 3, whether or not this kind of convergence constitutes an explanation depends on whether or not there is a well-established theoretical framework in place that can “account for” why and how the phenomena are unified. The latter component was in fact absent from Maxwell’s formulation of the theory. Yet the theory undoubtedly produced a remarkable degree of unity.

A similar problem exists in the Kepler case. Recall, for instance, the way in which Kepler’s first and second laws fit together in a coherent way, given the physics of libratory motion. Although there was an explanatory story embedded in Kepler’s physics, it was incorrect; hence, contrary to what Whewell would claim, a coincidence of results by no means guarantees the truth of the explanatory hypothesis. Although the convergence of coefficients may count as a unification of diverse phenomena, more is needed if one is to count this unification as explanatory. This is especially true given that phenomena are often unified by fitting them into a very general mathematical framework that can incorporate large bodies of diverse data within a single representational scheme (e.g., gauge theory, Lagrangian mechanics). And mathematical techniques of the sort described by Whewell are important for determining a general trait or tendency that is common to the data while ignoring other important characteristics.

But Whewell himself seems to have recognized that more was needed if consilience was to count as truly explanatory; specifically, one needed a *vera causa* to complete the picture. In the conclusion to the section on methods of induction he remarks that those methods applicable to quantity and resemblance usually lead only to laws of phenomena that represent common patterns, whereas inductions, based on the idea of cause and substance, tend to provide knowledge of the essential nature and real connections among things (Whewell 1967, p. 425). Laws of phenomena were simply formulae that expressed results in terms of ideas such as space and time (i.e., formal laws of motion). Causes, on the other hand, provided an account of that motion in terms of force.

Unfortunately, Whewell was somewhat ambiguous about the relations between causes and explanations and sometimes suggested that the inference to a true cause was the *result* of an explanation of two distinct phenomena; at other times he simply claimed that “when a convergence of two trains of induction point to the same spot, we can no longer suspect that we are wrong. Such an accumulation of proof really persuades us that we have a *vera causa*”.¹⁰ Although the force of universal gravitation functioned as just such a true cause by explaining why terrestrial and celestial phenomena obeyed the same laws, gravitation itself was not well understood. That is, there was no real explanatory mechanism that could account for the way that the force operated in nature; and in that sense, I want to claim that as a cause it failed to function in a truly explanatory way. Hence, even though a *why* question may be answered by citing a cause, if there is no accompanying answer to the question of *how* the cause operates, or what it is in itself, we fail to have a complete explanation.

With hindsight, of course, we know that Whewell’s notion of unity through consilience could not guarantee the kind of certainty that he claimed for it. Regardless of whether or not one sees Whewell’s account of consilience of inductions as a model for current science, it is certainly the case that Whewell’s history and philosophy of the inductive sciences provided a unity of method that at the same time respected the integrity and differences that existed within the distinct sciences. It provided not only a way of constructing unified theories but also a way of thinking about the broader issue of unity in science. Each science was grounded on its own fundamental idea; some shared inductive methods (e.g., means, least squares), but only if they seemed appropriate to the kind of inquiry pursued in that particular science. In that sense, Whewell was no champion of the kind of scientific reductionism that has become commonplace in much of the philosophical literature on unity. Consilience of inductions was a goal valued from within the boundaries of a specific domain, rather than a global methodology mistakenly used to try to incorporate the same kinds of forces operant in physics into chemistry (Whewell 1847, p. 99).

Now let us turn to another context, one in which the focus is not on unified theories specifically but more generally on unity in science defined in terms of unity of method. I am referring to the programme outlined in the *International Encyclopedia of Unified Science*, a collection of volumes written largely by the

proponents of logical empiricism and first published in 1938. Although there were similarities to Whewell's attempt to retain the independence of particular sciences, the proponents of that version of the unity of science (Neurath) claimed that the localized unity achieved within specific domains carried no obvious epistemic warrant for any metaphysical assumptions about unity in nature. Their desire to banish metaphysics also resembled the Kantian ideal of unity as a methodology. What is especially interesting about that movement, as characterized by each of the contributions to the *Encyclopedia*, is the diversity of ideas about what the unity of science consisted in. Although that may seem the appropriate sort of unity for an encyclopedia, more importantly it enables us to see, in concrete terms, how unity and disunity can coexist – evidence that the dichotomy is in fact a false one.

1.4. Logical Empiricism: Unity as Method and Integration

It has frequently been thought that the unity of science advocated by the logical empiricists had its roots in logical analysis and the development of a common language, a language that would in turn guarantee a kind of unity of method in the articulation of scientific knowledge. In his famous 1938 essay "Logical Foundations of the Unity of Science", published in the *International Encyclopedia of Unified Science* (Neurath et al. 1971), Rudolph Carnap remarks that the question of the unity of science is a problem in the logic of science, not one of ontology. We do not ask "Is the world one?", "Are all events fundamentally of the same kind?". Carnap thought it doubtful that these philosophical questions really had any theoretical content. Instead, when we ask whether or not there is a unity in science we are inquiring into the logical relationships between the terms and the laws of the various branches of science. The goal of the logical empiricists was to reduce all the terms used in particular sciences to a kind of universal language. That language would consist in the class of observable thing-predicates, which would serve as a sufficient reduction basis for the whole of the language of science. Despite the restriction to that very narrow and homogeneous class of terms, no extension to a unified system of laws could be produced; nevertheless, the unity of language was seen as the basis for the practical application of theoretical knowledge.

We can see, then, that the goal of scientific unity, at least as expressed by Carnap, is directly at odds with the notion of unity advocated by Whewell. The kind of reductionist programme suggested by the logical unity of science would, according to Whewell, stand in the way and indeed adversely affect the growth of knowledge in different branches of science. According to him, the diversity and disunity among the sciences were to be retained and even encouraged, while upholding a unity within the confines of the individual branches of science.

But, as with the problem of the unity of science itself, within the logical-empiricist movement there were various ways in which the notion of unity was understood, even among those who contributed to the *International Encyclopedia of*