CHAPTER 1

Introduction

1.1 THE GROWTH OF THEORIES

A very naïve view of science might go something like this: Scientists encounter a range of observable phenomena for which they have no explanatory account. Hypotheses are generated from the imaginations of the scientists who seek to explain the phenomena in question. These hypotheses are tested against the experimental results. If they fail to successfully account for those results, the hypotheses are rejected as unsatisfactory. But if they succeed in predicting and explaining that which is observed, they are accepted into the corpus of scientific belief. Then scientific attention is turned to some new domain of, as yet, unexplained phenomena.

This simple-minded picture of science has been challenged for a variety of reasons. Some are skeptical regarding the possibility of characterizing theory-independent realms of observational data against which hypotheses are to be tested. Others have noted the way in which the testing of hypothesis by data is a subtle matter indeed. It has often been noted, for example, that even our best, most widely accepted fundamental theories often survive despite the existence of "anomalies," observational results that are seemingly incompatible with the predictions of the theories.

This book is meant to challenge the naïve view as well. But the failure of the naïve view to do justice to how science really works is worth considering from a perspective that has, perhaps, not yet received the full attention it deserves. The simple view of theories is one that fails to do adequate justice to the fact that a fundamental theory can play its part in received science over a long period of time. Without making too much of a metaphor, it is useful to compare the life of a theory over time with the life of a living being. Theories have their "fetal" stage, playing a role in science even before they exist as fully formed hypotheses. One might speak of a "pre-theory" stage in the life of a theory. When theories are first fully hypothesized and first accepted into the body of scientifically accepted belief, they exist in their first "formative" state. But, just as a living being matures, and in maturing changes its aspect in deep and important ways, theories too may, over time, develop and change. Indeed, some years, decades, or even centuries, after a theory has first been accepted into the scientific corpus, it may still be playing a fundamental explanatory role in science. But what that theory looks like, indeed, what that theory *is*, in the later stages of its existence, may be something very different from what it was taken to be when it first appeared in science.

In its later stages, the account a theory gives of the phenomena it explains may have a very different form from the account it gave in its earlier stages. In fact, there may be good reason to say that in its later guises the theory may even deal with quite different phenomena than those in the domain of the theory in its earlier incarnation, or it may treat what we might call the same phenomena, but in such a way our very characterization of those phenomena may have taken on a quite novel and distinct form. Yet it will still be appropriate to speak of this later theory as being "the same theory" as the one we had years before in the theory's infancy. It is not fair to say that the earlier theory has been refuted by its newer incarnations, nor that the newer version of the theory has replaced the old. It is as appropriate to speak of one and the same theory over time, despite the remarkable transitions the theory has undergone, as it is to speak of one and the same person in infancy, adolescence, maturity and old age. Once again pushing a metaphor rather far, it is even appropriate to speak of theoretical senescence and, further, to contemplate the remains of a theory even after its demise, its corpse as it were.

Theories can be narrow in their scope and "shallow" in their place in the overall hierarchy of our theoretical description of the world. On the other hand, they can be of very broad scope, indeed, and deeply entrenched at a fundamental place in the overall scientific scheme of things. Even a broad-scope and fundamentally placed theory can occupy the esteem of science for a very short period. We shall see an example of this when we look at Descartes' dynamics and cosmology. And a narrow and shallow theory can have a long life span. But most interesting for our purposes are theories whose scope is broad, whose place in science is fundamental, and whose life span is greatly extended. It is from theories of this kind that we will learn most about the life-history and development of a theory in science.

We will focus our attention on one such theory, the "mother of all theories" in fundamental physics. The theory that will be the object of our attention is sometimes called Newtonian dynamics and sometimes pre-relativistic classical dynamics, or, more briefly, just classical dynamics.

What is classical dynamics? It is a theory that encompasses concepts designed to allow us to describe matter in motion through space during intervals of time, namely the concepts of what is called kinematics, and the concepts needed to give an explanatory account of just why matter in motion moves as it does, that is to say the concepts of what is called dynamics. It is a theory that held the throne as the ruling explanatory account of theoretical physics from the time of Newton's *Principia* in the last third of the seventeenth century until the special relativistic, general relativistic and quantum revolutions changed dynamics forever in the first third of the twentieth century. Its evolution during its reign as dominant explanatory account for over two hundred years was one of enormous richness and complexity. The many aspects of the evolving nature of this theory provide wonderful paradigm examples of almost any thesis one would wish to illustrate in a theoretical account of how theories are born, mature and age. It is by constant reference to the history and to the nature of classical dynamics as it grows and evolves that we shall learn our methodological lessons about the life of a theory.

The pre-history of classical dynamics goes back to ancient Greek astronomy and mechanics. The Greek cosmological models of the classical and Hellenistic period, culminating in the great work of Ptolemy, and the attempts to give a systematic account of change in general, and motion in particular, of Aristotle, as well as the brilliant insights into statics of Archimedes, provided the basis on which all further scientific understanding of motion and its causes was ultimately built. Underlying this early science, and also fundamental to the later development of dynamics, were the deep insights into mathematics, and especially into geometry, of the Greek mathematicians.

Deeply insightful critical comments on problems with the Aristotelian account of motion and its causes date back as early as the sixth century. Further profound illumination comes from the work of the Islamic scientists on the nature of projectile motion, and later from the deep insights of the impetus theorists of Latin Europe in the later middle ages. The true ripening of the pre-history of classical dynamics, though, begins with the Copernican revolution in astronomy and reaches its height in the dynamical insights of early modern dynamics, especially at the hands of Galileo and Huyghens.

The birth of modern classical dynamics, if one wants a single birth date for it, would likely be taken to be the appearance of Newton's great work, the *Principia*. There is one sense in which in that work we do have many of the core elements of the theory "fully formed." But, and this will be a major theme of this book, there is another sense in which it is truly impossible to see even in Newton's brilliant systematization of dynamics the true appearance of what classical dynamics was to become in its maturity. The more than three hundred years since the publication of Newton's work have seen classical dynamics explored by many of the greatest minds of modern physical theory: the three Bernoullis, Maupertuis, d'Alembert, Euler, Lagrange, Laplace, Poisson, Hamilton, Jacobi, Mach, Hertz and Poincaré, just to name a few of the very greatest. It would be the gravest error to think of the work done by these scientists of genius applying themselves to the theory of classical mechanics as merely "adding footnote" to Newton, or "filling in the details" of the theory, or, perhaps, "reformulating the theory for the purposes of convenience and practicality of application," although they did do all of these things. Rather, their work in exploring and transforming the Newtonian theory gave to that theory its ongoing developmental growth.

It will be our purpose here to look into that growth to see what methodological lessons can be learned from it by the philosopher of science.

1.2 THE FORMULATION AND REFORMULATION OF THEORIES

Fundamental theories of physics are subjects of repeated programs of reformulation. A theory may originally be proposed in the guise of a particular formal structure. Some concepts are introduced as primitive. Other concepts defined in terms of the original conceptual basis are subsequently introduced. Some laws are proposed as fundamental. From these basic laws various consequences are deduced using the apparatus of logic and various branches of pure mathematics.

Later, however, it is seen that there may be other ways of presenting the basic concepts and laws of the theory. Sometimes a reformulation of the theory is offered that can be argued to be formally "equivalent" to the original presentation of the theory. In other cases, though, the reformulation may go beyond the original version of the theory in significant ways. New concepts not obviously definable from those of the original formulation may be introduced. New laws or structural constraints may be posited that do not merely express the content of the original laws in a different manner. For example, the new laws may add significant generality to the original version of the theory. Yet they may do this in such a way that we are inclined not to speak of them as presenting a new, more general, alternative to the original theory, but, instead, as somehow "filling out" that which was already implicit in the original theory or as "completing" the task the original version of the theory had set itself.

In many cases the reformulations of a theory are motivated by "practical" considerations. The theory is designed to solve particular problems. In the case of dynamics, it is designed to allow us to predict and explain the motions of bodies of various kinds subject to various forces and constraints. But, in its original version, it may be quite difficult to apply the theory in a fruitful way to some classes or other of problem situations that ought to come within the theory's scope. Perhaps if the theory were given a variant formulation, some of these problem cases would prove more tractable. In its new presentation the theory could, perhaps, easily be applied to the cases that proved impenetrable to the original version of the theory. But, even if a reformulation is motivated initially by practical considerations of this sort, it may turn out to be the case that, once the reformulation is in hand, it is seen to have far broader implications for our understanding of the theory. A theoretical redesign originally motivated as a matter of convenience or practicality can turn out to have unintended deep, theoretically fundamental, consequences.

In other cases the desire for deeper theoretical insight may be the explicit motivator of a theoretical reformulation. Exploring the nature of the theory already in hand, the theoretical community may discover hidden within that theory structural features that remain disguised in the original formulation of the theory. A search may then ensue for a way of reformulating the theory that brings the deep, hidden structural aspects to the forefront. Or, in exploring the theory, the theoreticians may discover fundamental analogies between structures of the theory in question and structures known from other important theoretical disciplines. Such cross-theoretical structural analogies may, once again, reveal important, theoretically interesting, aspects of the theory that might not be fully explicit when the theory is expressed in its existing form. So, again, a reformulation of the theory is sought that will make these inner structural features explicit in the surface presentation of the theory. Once these structures have been brought to the surface, they may then be employed fruitfully in further developing the theory or in applying it to still more general ranges of difficult cases.

Even purely "philosophical" motivations can lie behind the desire to reformulate a theory in fundamental ways. A philosophically minded scientist may object to the common understandings of the sort of world that a theory seems to demand in order that the theory give a correct account of nature. But the scientist may very well believe, not that the theory is incorrect, but that the inferences drawn from the theory about what the world must be like can be disputed. The scientist, that is, objects not to the theory but to its "interpretation." Perhaps if the theory were reformulated in a more satisfactory way, it may then be argued, one would no longer be misled into inferring bad "metaphysical" conclusions from the surface appearance of the theory. Perhaps, indeed, we can reformulate the theory in such a way that its better, more philosophically acceptable, interpretation can now be read off from its surface features.

Just as a theory can evolve over long periods of time, remaining one and the same theory while displaying radically different guises, the interpretive issues that plague a theory show the same ability to evolve and mutate. An interpretive issue that arises when the theory is at one stage of its development might be resolved by some later reformulation of the theory. But it might, instead, reappear as a problem area for the new reformulation of the theory, mutating in its form as the theory changes. Three of the great, classic interpretive problems of classical dynamics show this co-evolution of an interpretive problem with evolution of the theory. How should we understand the role of "force" in the theory? What understandings of the nature of space and time are necessary as underpinnings for dynamical theory? What modes of explanation need be taken as fundamental for dynamical theory? These are three interpretive questions that just will not go away. As the theory changes in its formulations and reformulations, these interpretive questions change as well. But just as we can see one and the same dynamical theory throughout its manifold reformulations, so also are we vexed by one and the same set of interpretive perplexities over the centuries, even if those perplexities evolve in their formulation as the theory itself changes.

Significant reformulations of a fundamental theory often reveal to science new perspectives on the theory and the world it describes that are of great importance. What are some of the possible consequences that can result from formulating a theory in a novel manner?

For one thing, a reformulation can supply radically new insights into the kinds of explanations of the phenomena a theory can offer. It is not simply that in the reformulated version of the theory our explanations will, of course, look somewhat different from those offered by the theory in its earlier versions. Rather, the new formulation of the theory may provide us with wholly distinct *kinds* of explanation, explanations of a sort entirely unexpected if one merely looked at the theory as it was construed in its older versions. Indeed, when examined in its new guise, the theory may lead us to reconsider, philosophically and methodologically, our very ideas of what sorts of structures may appropriately be called scientific explanations. As we shall see in some detail when we look at the sorts of explanations offered by several versions of classical dynamics, deep methodological controversy can be initiated by some claim to the effect that, once the theory had been reformulated, wholly novel sorts of explanations could be seen to receive scientific legitimacy.

A radical reformulation of a theory may lead to new insights into how the "metaphysics" of the theory is to be understood. What seemed to be the necessary ontological interpretation required in order that the theory, in its earlier versions, adequately describe the world, may seem, once the theory has received its novel reformulation, to be only an optional "interpretation" of the theory. If the earlier theory seemed to demand a world of a nature that one found philosophically objectionable, the existence of the reformulated theory may be used as part of an argument to the effect that nothing of the earlier objectionable metaphysical view of the world need be taken as imposed upon us by our desire to hold to the legitimacy and adequacy of the theory's scientific account of nature. Here again classical dynamics will provide an exemplary case of how just such uses of the possibility of theoretical reformulations to support alternative metaphysical interpretations of the theory function in practice.

A reformulation of a theory may lead to the realization that the original theory had whole realms of phenomena to which it could be applied, but which escaped notice as being within the theory's purview when the only versions of the theory available were those in its earlier formal incarnations. A theory once developed to deal with some domain of observable phenomena in the world may prove to have within its scope the possibility of explanatory accounts of many things that go on in the world that were never suspected, at the time the theory was originally devised, to be treatable by the theory's methods. We shall see just such an extension of a theory's applicability, in part suggested and motivated by the opportune systematic reformulation of the theory, once more in an example chosen from the history of the development of classical dynamics.

A theory may have as a derivative consequence of its basic assumptions some results whose generality and importance goes far beyond anything apparent from the place of these consequences in the original version of the theory. That is, there may be something that follows from a theory, but seems initially to be but one consequence of the theory among many, and whose importance may be masked by the way in which the original theory is formulated and the way in which the consequence functions within that original formalization. A reformulation of the theory may serve to reveal, finally, the true generality and importance of the consequence of the theory, the importance that had remained hidden. Indeed, once reformulated the theory may even reveal to us that these consequences of hidden importance had a generality and profoundness of such scope and depth that they would prove applicable far outside the limits of the concerns of the original theory. And the fundamental nature of these consequences might prove to be such that even when the original theory, in all its versions and guises, became rejected as no longer a true account of the world, the old theory's consequences, now in their true representation as fundamental principles, might even survive the wreckage of the theory from which they were first derived. We shall also see an example of just this process in our exploration of classical dynamics.

Accepted scientific theories are usually only transient place-holders in our scientific esteem. Even the most widely accepted and most fundamental theories must always contend with the prospect that at some future date they will be replaced by an incompatible successor theory. But scientific theory change, at least at the level of foundational physics, is rarely a kind of change in which the newer theory is wholly unrelated to the theory whose place it is usurping. In even the most revolutionary changes in foundational physics, the successor theories borrow greatly, in terms of concepts and formalism, from their predecessors. How should we react if a theory is found wanting? It might be found to be unsuccessful in correctly predicting the observational and experimental facts. Or it might be found to suffer from some internal incoherence. Or it may be the case that we must reject the theory because of its incompatibility with some other accepted theories of science that we cannot think of rejecting at the time in question.

When a theory is found to suffer from one or more of these faults, something must be done. But what is to be done may itself be suggested to us by the structure of the theory found wanting. We may, for example, be inclined to localize the failure of the existing theory in a particular one of its many components, suggesting that we might leave the other components alone in making our change to a newer theory. But, once again, how to evaluate the impact of some failure or another of a theory, and how to move ahead with changing our theory when failure is encountered, may be matters highly dependent on just how the theory found to be at fault is being formulated. Alternative reformulations of the theory may suggest quite different programs for modification and revision in our search for an improved successor to our, now to be rejected, current theory.

When successor theories are created, they often are discovered by taking apart the components of the theories they are to replace and making systematic changes in those parts of the existing theory that are to be replaced. But, then, how the existing theory is formulated, how its basic concepts and assumptions are characterized, will be highly influential in suggesting just what parts of the older theory are to be used, in modified form, to construct its new replacement, and just how these aspects of the older theory are to be changed to construct the new one. Classical dynamics was used in just such a way to construct its three famous successor theories: the special theory of relativity, the general theory of relativity and quantum mechanics. The existence of multiple reformulations of classical dynamics was invaluable in providing a rich source of suggested novel theoretical elements with which to build the theories that would replace classical dynamics itself.

1.3 THE STRUCTURE OF THIS BOOK

This book blends together a sketch of one thread in the history of science, an informal exposition of a number of aspects of one branch of theoretical science, and an attempt to derive a number of methodological conclusions in the philosophy of science from the historical and scientific material.

Let me first make some disclaimers. The history of science presented here is all quite derivative. Although I have tried to make use of original sources (at least in translation), most of the history outlined here is familiar from the established secondary sources – few as they are in the history of dynamics. Let me add that the kind of history of science with which we will be concerned is purely of the "internal" variety, and even then far from what would be expected in a work primarily devoted to the history itself. We shall look at how one idea led to another within physical science. Very rarely indeed will we touch on anything having to do with the more general historical, social or cultural context in which the science appeared. Nor will we be concerned with biographical or psychological aspects of the scientists involved. Even from the purely "internal" perspective, our focus will not be on the details of origin and influence. The history will be present, rather, as a source of illustrative material from which we might hope to extract some methodological insights.

The science presented here is familiar material, and it has been dealt with abundantly in many sources at many levels of technical sophistication. Once again, our attention to the scientific material will be selective, picking and choosing those elements in the history of classical dynamics that our useful for our methodological inquiry. Formal exposition of the science, especially the details of how results are derived and proved, will be kept to a minimum. I will delve into details only when they illuminate methodological points. The focus of the book at all times will be neither the history of classical dynamics, nor the contents of the theory itself, but, rather, how the history and internal structure of this theory can illustrate the philosophical and methodological themes outlined above.

Although this book is not a work in the history of science, I shall use a chronological history of classical dynamics as the framework on which to hang the methodological points. Overall, the book is structured temporally, starting with the pre-history of classical dynamics, and following its evolution through the Scientific Revolution, then through the period of its great development in the eighteenth and nineteenth centuries, and on through some developments in the theory, and in the theory's relation to its successor theories, in the twentieth century. Strict chronology, however, will not be adhered to, since in some cases following a single topic over a long time span will provide a more coherent organizational structure than would coming back to that topic again and again as its development occurred in time.

We will begin, then, with a sketch of the pre-history of classical dynamics. First there will be an outline of ancient Greek astronomy and ancient Greek work on the theory of motion. This will be followed by an even briefer look at the contributions made to dynamics in the period following Aristotle and preceding the great discoveries in astronomy and mechanics of the Scientific Revolution. This will include a look at an early critique of Aristotle, a glance at the work on motion of the Islamic school and a quick look at the impetus theorists of the later middle ages in Latin Europe.

Next there will be somewhat more detailed attention paid to the revolutionary results in astronomy from Copernicus to Kepler, and to the explosive developments that led to the modern science of motion in the work of Galileo, Descartes, Huyghens, Leibniz and others whose work directly impacted on that of Newton. After that we shall outline the most crucial aspects of the great Newtonian synthesis that set the stage for all further work in classical dynamics.

The work of Newton by itself led to a number of fundamental philosophical and methodological debates. The theory gave rise to issues of metaphysics: What must the world be like in order that the Newtonian theory could describe it? It stirred up issues of epistemology: What sort of theory of knowledge and inference could justify accepting the Newtonian claims? And it restored vigor to arguments about the nature of scientific explanation: What does the Newtonian theory tell us about what a scientific explanation should be like? Indeed, the philosophical controversy aroused by reflection on the Newtonian theory provided the paradigm for all future philosophical debates rising out of major revolutionary changes in foundational science. These debates will next take up our attention.

The core focus of the book, however, is on the development of classical dynamics following Newton's great synthetic work. We will be looking at how classical dynamics grew and evolved in the three centuries following the publication of the *Principia*. Here we will be concentrating for the most part on the many formulations and reformulations of the fundamentals of the theory, and how each reformulation brought with it its own issues of interest to the philosopher and the methodologist.

One of the topics to be explored will be the role of extremal principles in offering alternative fundamental dynamical laws for the theory and how such principles themselves evolved over the years. Another topic will be the difficulties encountered in applying the dynamical rules to ever more general classes of material bodies, going from the dynamics of point particles to that of rigid bodies, and finally making the theory applicable to fluids as well. Another topic to be explored is the way in which dynamics was reformulated to deal with the motion of objects subject to constraints. Whereas modifying the theory to allow it to effectively deal with the dynamics of bodies whose motion is restricted by specified constraints seems initially to be a practical problem of little fundamental theoretical interest, we shall see how the means devised to deal with the problem of constrained motion eventually opened up new conceptual vistas to the theory.

The issue of the origin, development and perfection of conservation principles is another topic we will look at. Principles of the "conservation of motion" were invoked very early in the development of the theory prior to Newton. But their role in the theory was one of constant refinement and constant re-evaluation. We shall see how they went from initial limited basic postulates, to derived consequences of other fundamental posits, to, ultimately, principles whose foundation was to be sought in symmetry considerations and whose scope outran that of the classical dynamical theory itself. After that we shall look at Hamilton's derivation of a novel set of foundational dynamical equations for the theory and how these invoked new concepts, generalized momentum and phase space, which, once again, recast the theory in such a novel framework as to open up wholly new insights into the theory and the world it described.

Hamilton was also responsible for reinvoking analogies between optical phenomena and dynamical phenomena in a way that put classical dynamics into an additional novel framework and perspective. We shall look at this.