1

Introduction and background

In recent years there has been a move to 'naturalize' the philosophy of science. This has meant basing work in the philosophy of science upon the actual historical record of real scientific practice and stressing (in varying degrees) the use of the methods of science in studying the scientific enterprise. This attention to actual scientific practice has been supported by traditional realists, an example being Ernan McMullin (1984) who early on (1976a) argued for a central role for the history of science in the philosophy of science; by philosophers of various anti-theoretical bents, such as Nancy Cartwright (1983) and Ian Hacking (1983); and by empiricists, like Bas van Fraassen (1980, 1985). In the early 1960s, it was attention to the historical record that led Thomas Kuhn (1970), in his *The Structure of Scientific Revolutions*, to stress the importance of social factors in the practice of real science. The spirit of the present work is that careful and detailed study must be made of the actual development of science *before* conclusions are drawn about the appropriateness of any particular methodology of science. Ours is mainly a story about theory, but not one uncoupled from its relation to experiment.

We claim that many of the philosophically interesting questions in science, especially in regard to possible changes in the methodology and goals of science, can be seen and appreciated only upon examination of the technical details of that practice. So, in this chapter we discuss some motivations for studying *current* scientific practice and then set the problem background out of which the *S*-matrix program arose. This sketch is somewhat ahistorical since we mention some field theory developments that occurred *after* 1943 (when Heisenberg introduced his *S*-matrix program). The formalism of classical particle mechanics and of wave phenomena have been, in this century, successively reinterpreted to yield (nonrelativistic) quantum mechanics, relativistic quantum field theory for electromagnetic phenomena and, finally, relativistic quantum field theory for strong (or 'nuclear') phenomena. The prototypical experimental arrangement (of either the real or *gedanken* type) for studying the fundamental interactions of nature has been a scattering process in which a target (such as a nucleus) is probed by a projectile (such as an electron or a proton).

However, the development of these quantum field theories has not always been a smooth one. In particular, in the 1930s and early 1940s, the quantum field theory program had run into considerable technical and experimental difficulties (Cassidy, 1981; Galison, 1983a). Mathematical inconsistencies, most notably divergences or infinities produced by calculations with the formalism, occurred for quantum electrodynamics (QED) at short distances (or, equivalently, at high energies) of the order of the size of the electron. Similar problems plagued Fermi's theory of β decay and Yukawa's meson theory of nuclear forces. Here we have examples of three of the four basic forces in nature: the electromagnetic, which is responsible for the atomic phenomena producing those features of the world we commonly encounter; the weak, which accounts for the spontaneous decay or conversion of a free neutron into a proton; and the strong, which predominates at very short distances for nuclear processes. The last is the gravitational force, which plays no essential role in our story here. During the same period of confusion in the arena of theory, experimental results from cosmic ray data seemed to contradict expectations based on quantum field theory (QFT). It appeared as though cosmic ray showers, or 'explosive' events, occurred (in contrast to the cascades built up from many essentially pairwise events, which could be readily accounted for by Dirac's hole theory). Heisenberg took the existence of these multiple processes to signal a breakdown of conventional quantum field theory and to require the introduction of a fundamental length into the theory. The field theory situation was further complicated by the confusion (in the 1930s) caused when the mesons observed in cosmic-ray interactions were at first identified with Yukawa's nuclear-interaction π meson (pion), before they were finally identified as μ mesons (muons), which are essentially 'heavy' electrons. These difficulties encountered by the quantum field theory program provided a significant part of the motivation for Heisenberg's proposing his *S*-matrix theory (SMT).

By the late 1940s, a mathematical technique (renormalization) had been formulated which allowed one to circumvent the divergences of QED and to make accurate predictions confirmed by experiment. This is the first cycle of the oscillation of theory between QFT and SMT. Others occurred when QFT was stymied by the strong interactions, from which it subsequently recovered with gauge field theories. This back and forth between formalisms, with their corresponding paradigms, is an important feature of the episodes we present. It will be especially relevant for our evaluation of methodology in science (in Chapter 10).

1.1 Internal history of recent science

This work is largely, but not exclusively, an *internal* history of an extended episode in modern high-energy physics. That is, the published physics literature is a major source for the technical developments we present. Nevertheless, interviews and correspondence are also employed. The primary interest in and motivation for doing the research necessary for this case study are philosophical. Some obvious questions, that arise about the value and wisdom of doing an internal history of a *current* (and hence not completed) episode in a *highly technical* (or specialized) subject area, must be addressed.

Schweber (1984, p. 41), in his history of the early developments of quantum field theory, has stated one of the problems of internal history as follows:

> [I]nternal history faces the problem common to all good history: how to avoid the pitfalls of Whiggish history, that is, the writing of history with the final, culminating event or set of events in focus, with all prior events selected and polarized so as to lead to that climax.

So, while the philosophy of science must be based on history (i.e., events as they actually occurred), it can be important not to focus exclusively on the form and content of 'successful' scientific theories alone. The arguments and contingent events that inform the course of development and selection of theories are essential. That is, how things might have gone a very different way at certain crucial junctures and why they did not may be as important as the reasons for the 'right' choices that science has made. The present case study focuses on a 'failed' program that has never been proven to be incorrect. It is evident that one cannot explain its rejection just in terms of falsification. Perhaps there is something to be learned from the history of such a dead-end theory. The relevance of such ('sociological') factors as the previous interests and expertise of the participants becomes apparent enough.

Historians may extol the virtues (in fact, necessity) of doing the history of an episode only long after the clamor of the day has settled. They can argue that once time and events have produced a stable picture of the past, one feels some confidence that one may be able to find 'the objective truth' in those long-dead events (or 'corpse') (Burckhardt, 1963, pp. 74–76).[1] There is an old tradition of this attitude in the history of science. Thus, in Whewell's *History of the Inductive Sciences* (1857, Vol. II, p. 434) we find: 'It is only at an interval of time after such events have taken place that their history and character can be fully understood, so as to suggest lessons in the Philosophy of Science.' Even if one accepts that thesis (and it *can* be debated), he can still feel that something (perhaps important, Burckhardt and Whewell to the contrary notwithstanding) has slipped away. The detailed dynamics of the events and the motivations of the protagonists have been lost (in large measure, at least) behind the veil of time. Now if one believes that all final scientific positions are reached ultimately through rational judgments *alone*, then there is probably even virtue in waiting until the flotsam has been swept away by time to leave a residue of objective truth. But it is not clear that science operates (*even* in the long run) quite as objectively as we might like to think. An examination of the record of actual scientific practice may shed some light on that question. The goal is not to clear up the rules and mechanisms that regulate the eternal ups and downs of fashions and fads in theoretical physics. It is not certain that there is *a* set of rules and mechanisms, but we can learn what some of them might plausibly be.

There are inherent dangers in studying fairly recent episodes in physics (or in anything else) (Brinkley, 1984). But, since the final 'verdict' is not yet in and since many of the participants are still alive, there are opportunities here that are not available in more traditional 'corpse dissections'. Most obviously, one can ask the major figures involved what their motivations were, how they saw events at the time, and what they recall about the interactions of other scientists. An obvious danger in gathering such recollections is that people sometimes feel (rightly at times, but often not) that their own contributions have been slighted. This 'interview' approach can be taken too far, as when sociologists of science monitor the day-to-day routine activities of scientists. An additional useful dimension can be added by examining a recent episode in science.

Many of the most interesting questions in the philosophy of science come from studying actual scientific practice, rather than from armchair *a priori* reasoning that philosophers sometimes engage in. I have chosen the dispersion-theory and S-matrix theory program of theoretical physics because I had some familiarity with the technical literature of that program in the 1960s and 1970s, because the major activity in that area was confined to a time period of several years and that activity was reasonably localized (around relatively few central theorists), and because several philosophically significant issues, such as the origin, development and selection of theories, can be illuminated with specific instances from a history of that program.

A difficulty in doing a case study of a major episode in modern theoretical physics is that one of the traditional sources of corroboration – an extensive personal correspondence among the major creators of the theories – is by and large no longer available. That is, historians of science are wary of taking at face value and relying solely upon the personal recollections of individuals. While such recollections are an invaluable source for leads about what actually went on behind the 'story' as reconstructed from the published physics literature, those recollections must be checked for support against other documents, usually the published literature and the private correspondence among key theorists. While the leading theorists of an earlier era (e.g., Einstein, Bohr, Schrödinger, Heisenberg, Pauli) did correspond frequently and extensively with one another (and much of that correspondence survives), markedly increased use of the telephone and relatively easily-available travel to many topical conferences have obviated the need for such correspondence among already busy individuals. The situation for the history of recent experimental physics is not so bad since laboratory notebooks and, more often, research proposals to funding agencies and the internal memoranda of large groups give details of what was going on in the major experiments (cf., Galison, 1987). However, research proposals for theoretical work provide a less reliable guide to what a theorist actually ends up doing.

This problem of a missing record of corrspondence among theorists is especially bad with those generations of theorists who have begun working since the end of the Second World War. Lacking a large body of such correspondence, the only resource available appears to be getting as many independent recollections as possible of key episodes in the development of the disperison-theory and S-matrix programs and then looking for the common overlap among these.

One can also question the value of studying a frontier area involving the creation of new physical theories, since this may be a singular exercise in science. Rigden (1987) has characterized such creative developments as follows.

> When first-rate minds are engaged in the intellectual activity called physics, as was the case in February 1927 when Heisenberg was struggling with the 'pq–qp swindle', it is an activity with no equivalent in any other natural science. In fact, there is no equivalent in any intellectual arena except, possibly, first-rate theological thinking. These special times in physics do not come often, but when they do, physicists must often create new constructs for which neither previous experience nor previous thought patterns provide guidance. New words representing entirely new concepts must be created, words whose meaning cannot be rendered even by the most deliberate use of older words. The new meaning takes form slowly, but with a groping awkwardness. Soon the new ideas become the basis for empirical predictions and, in the process, a 'sense of understanding' emerges. However, in the end, the basic concepts of physics are aloof, they remain outside our ability to convey their meaning.

Rather than taking this to mean that such activities in theoretical physics are largely irrelevant to the philosophy of science, we can see in these episodes a unique opportunity to examine how foundational theories are created – perhaps at a time of singular flexibility and underdetermination of the outcome.

1.2 Philosophical issues and the Forman thesis

Since a primary interest of ours here is certain philosophical questions, references are not given to *every* technical development in *S*-matrix and dispersion theory. By examining in detail a major episode in contemporary physics, we hope to illuminate somewhat the processes by which theories are generated and selected by the scientific community. A question of central interest for us is the relative importance of internal versus external factors in the development of a scientific theory. Forman (1971) initially raised this issue with regard to the origin and acceptance of the concept of acausality in physics in Germany after the First World War when modern quantum mechanics was being formulated. We shall often use the expression 'Forman thesis' to refer more generally to the role of social and sociological influences in the development and acceptance of a scientific concept or theory. It does seem evident that, once we 'buy' into a set of starting assumptions, then the 'internal' logic of a formalism can largely take over (Raine and Heller, 1981). However, the origin of hypotheses central to a theory often lies in very specific and technical developments, having little, if anything, to do with overarching philosophical schemes. For that reason some fairly extended discussion of technical details is necessary. (The reader can get an overview of the philosophical issues and conclusions from Chapters 1 and 10 alone, aided perhaps by the introductions and brief summaries at the end of each intervening chapter.) Retrospectively, the central tenets of a theory may be put into or associated with a particular philosophical world view. Furthermore, the acceptance (or the effective infectivity) of a theory can be greatly influenced by the social environment and by generally accepted overarching principles. This case study does not support the radical Forman (1971) thesis that the social milieu plays a central role in the *creation* of scientific theories, but it is consonant with the more modest Forman-type thesis (Forman, 1979; Hendry, 1980) that social factors are relevant for the *acceptance* of a theory. There does remain an important distinction for science between internal factors (such as formalism, logic and experiment) and external ones (such as group interests and social influences).

Another set of issues to be discussed in the context of this episode in physics is the interplay between the discovery and the justification of a scientific theory (both initially and later in the program) and the symbiosis between theory and experiment in the development of a program. This is just *one* case study and its conclusions may or may not have any general applicability to the way other scientific theories have developed. It is by no means clear that there is *a* (i.e., one) scientific rationality that applies usefully to all science in all eras (in spite of some claims made, for example, by Popper, with his emphasis on falsification, and Lakatos, with his representation of the dynamics of science in terms of progressive and degenerating research programs).

1.3 The purview of this case study

Because this case study is intended mainly for historians and philosophers of science who have an interest in the modern scientific enterprise, I have attempted to give an essentially accurate representation on technical matters, but have usually avoided telling the *whole* truth (i.e., giving all the technical details). Rather, the central concepts and techniques are often illustrated with simple mathematical examples. Although I do not want to reconstruct past developments from the biased vantage of today's state of knowledge, I have nevertheless employed a unified notation in these mathematical examples in order to make the line of argument more accessible to a wider audience. Along the way I do point out important notational and conceptual differences between my illustrative examples and the original presentations found in the physics literature.

The last introductory comment concerns my use of the expression '*S*-matrix program'. I do not mean to equate the dispersion-theory program and the *S*-matrix theory (SMT) program[2] nor do I wish to obfuscate the distinction between the 'bootstrap condition' as a uniqueness criterion and the much broader implications that term has in the program associated with Geoffrey Chew and his collaborators. This case study should make it clear (1) that at any given time the term 'bootstrap' has not had a unique, universally accepted meaning among theoretical physicists (if, indeed, it has any specific meaning at all) and (2) that within a given group or school of theorists the meaning of that term has evolved over the years. To respect this caveat, I shall use the designation 'autonomous *S*-matrix program' to distinguish the radical or fundamentally revisionary conjecture from the more general *S*-matrix and dispersion-theory program[3].

Much of the *S*-matrix program discussed in this study will appear as a largely American project and this may give the entire project too much of an American, even 'Chewian', flavor. It is true that major developments in several areas took place in Europe or have been made by Europeans. I attempt to point this out in the narration that follows. Nevertheless, it does remain that much of the major activity of the *S*-matrix program was centered around Geoffrey Chew and his collaborators in Berkeley.

1.4 Quantum field theory (QFT) background

A brief sketch of the development of quantum field theory (QFT) is necessary in order to place *S*-matrix theory (SMT) in some historical context for the reader. More complete discussions of the history of QFT can be found in Cushing (1982), Darrigol (1984, 1986) and Schweber (1984, 1990). Philosophical problems associated with QFT have been addressed by Stöckler (1984). In fact, the following is largely a summary from my (1982) paper on high-energy theoretical physics. Detailed references to the relevant physics literature can be found there, so that we shall not repeat those references here. This condensed outline is essentially ahistorical in that it presents only the central ideas of QFT without any pretense at maintaining a strict historical sequence. Let us begin by reproducing (Cushing, 1987a) a chronological outline of the sequence of developments (Schweber, 1987) we shall take as background for the following chapters in this case study. Some of these topics will be elaborated here, others in subsequent chapters.

1925–1927: formulation of nonrelativistic quantum mechanics
1927–1947: formulation of relativistic quantum field theory (QFT)
1947–1950: renormalization program for quantum electrodynamics (QED)
late 1950s–1970: a period of serious problems for perturbative QFT, with various alternative avenues pursued
(*a*) axiomatic QFT (Wightman school)
(*b*) local, asymptotic QFT (LSZ formalism)
(*c*) dispersion relations and *S*-matrix theory (SMT)
↓
duality program

1970s–present: gauge QFT ↘ strings ↙ duality program

The transition from classical mechanics to (nonrelativistic) quantum mechanics in the period 1925-27 can be seen (at least now, retrospectively) as a reinterpretation of the equations and of the formalism of classical mechanics to represent phenomena in the atomic domain. In the Hamiltonian formulation of classical mechanics (for a single particle here), the time evolution of the canonical variables $q(t)$ ('position') and $p(t)$ ('momentum') is governed by Hamilton's equations of motion

$$\dot{q}=\frac{\partial H}{\partial p}, \tag{1.1a}$$

$$\dot{p}=\frac{-\partial H}{\partial q}. \tag{1.1b}$$

Here $H(q,p)$ is the Hamiltonian and is (in our case) just the total mechanical energy of the system. For example, a particle of mass m moving in a conservative force field $F(q)=-\partial V/\partial q$ has the Hamiltonian

$$H(q,p)=\frac{p^2}{2m}+V(q) \tag{1.2}$$

in terms of the potential energy function $V(q)$. For a classical system, $q(t)$ and $p(t)$ are simply ordinary functions of the independent time variable t. They are solutions to the coupled set of differential equations (1.1) subject to the initial conditions $q_0=q(t_0)$, $p_0=p(t_0)$ at some (arbitrary but definite) initial time $t = t_0$. In this simple example, Eqs. (1.1) are nothing more or less than (equivalent to) Newton's second law of motion

$$m\ddot{q} \equiv ma = \frac{-\partial V}{\partial q} = F(q). \tag{1.3}$$

In the early part of the present century, it became evident that for atomic systems not all of the solutions (or 'orbits' for particle motion) are in fact allowed or realized in nature. For example, only certain orbits, or energy levels, for a bound electron in a hydrogen atom are permitted (as evidenced by the discrete spectrum of the light emitted or absorbed by a hydrogen atom). The program of the old quantum theory (say, 1913–1925) was to find a set of rules that would allow one to select from the (continuous) infinity of classically-allowed solutions (or 'orbits' or energy levels) those actually realized in nature. Bohr's classic 1913 paper gave one such rule in terms of the quantization of the orbital angular momentum l,

$$l = n\hbar, \quad \mathrm{n} = 0, 1, 2, \ldots \tag{1.4}$$

Here $\hbar$ is $h/2\pi$, where h is Planck's constant. The old quantum 'theory' amounted in essence to a set of quantization rules, that were generalizations of Eq. (1.4). It consisted of a set of *ad hoc* guesses guided by Bohr's correspondence principle, which was initially a requirement that certain quantities derived in the (old) quantum theory should pass over into their classical counterparts in a suitable limit.

Heisenberg's 1925 paper laid the foundations of a systematic quantum mechanics by reinterpreting the classical q and p variables as quantities satisfying the commutator relation (in units with $\hbar = 1$)

$$qp - pq \equiv [q, p] = \mathrm{i}. \tag{1.5}$$

(We make *no* claim that Heisenberg, Schrödinger or Dirac originally presented their ideas in the form we represent them here. This is 'Whiggish' history, which we avoid in our study proper.) Hamilton's equations (1.1) and the Hamiltonian (1.2) were to be retained, but the (operators) q and p were now required to satisfy the (commutator) condition of Eq. (1.5). That is, one must seek solutions to the eigenvalue problem

$$\mathrm{H}(q, p)\Psi = E\Psi \tag{1.6}$$

for the allowed eigenvalues E. Here (to make a long story short) Ψ is the (Schrödinger) eigenfunction (or eigenvector). One typically finds a