1

Early string theory as a challenging case study for philosophers

Elena Castellani
Dipartimento di Filosofia, Università di Firenze

1.1 Introduction

The history of the origins and first developments of string theory, from Veneziano's formulation of his famous scattering amplitude in 1968 to the so-called first string revolution in 1984, provides invaluable material for philosophical reflection. The reasons why this episode in the history of modern physics – one still largely unknown to the philosophy of science community despite its centrality to theoretical physics – represents a particularly interesting case study are several and of various nature. It is the aim of the present Chapter to illustrate some of them.

In general, the story of the construction of a new scientific theory has an evident interest in itself, as a concrete example of how a particular theory has been discovered and developed by a given community and over a certain period of time. On the other hand, case studies taken from the history of science are commonly used, by those philosophers of science who pay attention to actual scientific practice, to provide some evidence for or against given positions on traditional epistemological or methodological issues. In other words, historical case studies are attributed, with respect to philosophical 'theories' on given aspects of the scientific enterprise, a role analogous to that of the data of experience in scientific theories. These aspects can be of a very general character, such as those regarding the methodology, aim and evaluation of scientific theories; or of a more specific kind, such as the significance of a certain principle, argument or concept.

The case study presented by early string theory, as narrated by the contributors to this Volume, is fruitful from both a general and a specific perspective. It provides, first of all, an illustration of the first steps of a scientific theory which has been dominating a significant part of theoretical physics research over the last decades, thus feeding the philosophical reflection with

valuable data on how scientific theories are constructed and selected. It sheds light, at the same time, on the original meaning of ideas – such as duality, supersymmetry and extra spacetime dimensions – and mathematical techniques that are basic ingredients in today's theoretical physics, as is also discussed in the Introduction to Part I.

The philosophical interest of early string theory is examined here on the basis of the historical outline emerging from the collection of reminiscences and surveys contained in the book. In this history, both 'internal' factors (such as the form, content and logic of the theory) and 'external' factors (such as the psychological or sociological aspects influencing the scientific work), to use a traditional distinction, are well represented. It is usually considered controversial whether an historical reconstruction mainly based on personal records, viewpoints and experiences, can indeed be objective. In this case, however, we are confident that the multifaceted account resulting from the gathering of the independent recollections of almost all the scientists involved, checked on the grounds of the original papers, can provide an accurate and balanced historical picture.

1.2 The case study

The discovery by Veneziano of his 'dual' amplitude for the scattering of four mesons is widely acknowledged as the starting point for the developments leading to string theory. In fact, as illustrated in the first Parts of the Volume, it immediately gave rise to the very intense theoretical activity that is known, in general, as the 'dual theory of strong interactions': from the first two models proposed – the Dual Resonance Model and the Shapiro-Virasoro Model, respectively – to all the subsequent endeavours to extend, complete and refine the theory, including its string interpretation and the addition of fermions.

This was the first phase of early string theory and was motivated by the aim of finding a viable theory of hadrons in the framework of the S-matrix theory as developed in the early Sixties. Chronologically, it extends from summer 1968, when Veneziano presented his formula at the Vienna Conference on High Energy Physics with a strong impact on the theoretical physics community, to the time when the interest in dual models for describing hadronic physics began to decline. This happened towards the end of 1973, primarily for the following two reasons: firstly, the presence of unphysical features in the theory (such as extra spacetime dimensions and an unrealistic particle spectrum); and secondly, the evidence of point-like constituents at short distance inside hadrons (subsequently to be identified

with the quarks) that was obtained in deep inelastic scattering experiments at the Stanford Linear Accelerator Center (SLAC)†. This feature of strong interactions could not be explained in terms of the dual theory, while the competing gauge field theory, Quantum Chromodynamics (QCD), was able to account for it.

This first phase has some remarkable and unusual characteristics and can be considered as a case study in itself. It provides a valuable example of the rising, blooming and apparent subsiding of a revolutionary theoretical project, in a rather short and well delimited period of time and with a rare confluence of people, intents and places. As testified by most of its exponents, the atmosphere in which the project was first developed was particularly enthusiastic and cooperative. The young age of the majority of the physicists involved – many were either graduate students or postdoctoral fellows – no doubt had something to do with it; but what counted most was the shared conviction that something new and exciting was being created.

Another 'external' factor was conductive to the creation of such an atmosphere, namely, the concentration of research in the field in just a few institutions across the USA and Europe. One of these in particular, the CERN Theory Division in Geneva, played a central role, mostly thanks to the charismatic presence of Amati (CERN staff member at the time). He was able to gather a strong group of research fellows and visitors and almost all those who were working on dual theory spent some time there during this period. Under Amati's leadership, the 'dual group' formed a sort of theoretical laboratory, with regular seminars and several collaborations going on, in an atmosphere that was not only stimulating but also very friendly. The outcome was impressive: in a few years, the theory advanced remarkably fast and many decisive results were obtained, such as the proof of the No-Ghost Theorem and the quantization of the string action.

In the course of the year 1974, however, the atmosphere changed radically. The fascination with the theory remained the same but, on account of the problems it encountered as a description of strong interactions, the interest of the high-energy community focussed on other developments, such as QCD and Standard Model physics. The dual string theory of strong interactions seemed to be a failed program, with the consequence that many of the young theorists involved felt forced to leave the field to retain any hope of pursuing their academic career.

---

† It is worth noting that, although the SLAC deep inelastic scattering experiments were contemporary to the discovery of the Veneziano amplitude, it took some time to formulate their appropriate theoretical description – a description which eventually provided essential support to QCD. In this time lapse the dual models could be developed.

Some people, however, resisted the general trend and went on addressing the unresolved problems of the theory, thus preparing its later renaissance in the form of modern string theory. This second phase, illustrated in the last two Parts of the Volume, extends from 1974 to when, towards the end of 1984, the interest in string theory exploded again thanks to crucial results obtained in the context of supersymmetric string theory. Following this 'first superstring revolution', as it is now usually referred to, string theory rapidly became a mainstream activity and a new phase began, one that continues right up to the present.

From a sociological point of view, the second period starting with the general change of attitude in 1974 is less homogeneous than the first. The motivation for pursuing string theory was a shared one: the theory was regarded as so beautiful and had such a compelling mathematical structure, obtained in agreement with consistency conditions and deep physical principles, that it was expected to be in some way related to the physical world. This was the attitude adopted, in particular, by Scherk and Schwarz and, independently, Yoneya. They took the view that those very features that were considered drawbacks in describing hadronic physics, such as the presence of spin-one and spin-two massless particles and extra dimensions, could instead reveal the true nature of string theory. This view led them to investigate in depth the connection between dual models and field theories, that turned out to be the most relevant ones: Yang-Mills (non-Abelian) gauge theories that are employed in the Standard Model of particle physics and general relativity describing gravity. The resulting remarkable proposal, advanced by Scherk and Schwarz, was that string theory should be considered as a unified quantum theory of all the fundamental interactions. It was a 'big conceptual leap', as Schwarz describes it in his Chapter (Part I). In particular, it implied a huge rescaling (by 19 orders of magnitude) of the theory's characteristic scale, in order to relate it to the Planck scale, that is, the scale associated to quantum gravity.

However remarkable and promising, such a change of the domain and goal of the theory was not really appreciated, outside of a limited circle, for almost a decade. String theory remained a side issue for several years, notwithstanding the two major results obtained in 1976: the formulation of the supersymmetric action for string theory and the construction, by means of the so-called GSO projection, of the first totally consistent supersymmetric string theory (i.e., superstring) in 10 dimensions.

The case study considered in this Chapter is focussed on the developments of the theory until the 'exile' period before its 1984 renaissance. In this last stage of dual string theory, covered in the final Part of the Volume, only a

few adepts remained, with the consequence that the theory's progress slowed down considerably. In contrast with the flourishing early phase, the work on string theory was no more the central activity of a consistent and intensively collaborating group, but rather the independent research of small groups or isolated individuals who were also actively engaged in the more popular area of supersymmetric extension of field theory and general relativity.

As discussed in Part V, supersymmetry – the symmetry relating bosonic and fermionic particles – was discovered in the context of early string theory. It was implicitly present in Ramond's construction of the fermionic string in 1971, and was explicitly remarked as a property of the string action by Gervais and Sakita in the same year. Supersymmetry was then extended to quantum field theory by Wess and Zumino, in 1974, as a global symmetry acting on four-dimensional spacetime. Following their work, supersymmetric field theories in four and higher dimensions were quickly developed. Meanwhile, theories with local supersymmetry were constructed and analyzed independently of string theory. Since these theories generalized Einstein's theory of general relativity they were referred to as 'supergravity theories'.

These developments were also motivated by the growing interest, toward the end of the Seventies, in the program of unification of electro-weak and strong interactions above the Standard Model energy scale and in that of unification with gravity. For several reasons that are illustrated, in particular, in the Introductions to Parts VI and VII (see also Chapter **??**, Section **??**), supersymmetric field theories and supergravity appeared to be promising candidates for realizing this program. Despite these shifts toward quantum field theory and supergravity as the primary interest of many theoretical physicists, research in string theory continued during these years and led to important developments. As described in Part VII, a new Lorentz covariant quantization of string theory, based on path-integral methods, was obtained by Polyakov in 1981; furthermore, superstring theories, offering a consistent unified quantum theory in ten spacetime dimensions, were classified and their properties were analyzed in the quest for a convincing unification framework.

In fact, the research activities in superstring and supergravity were deeply intertwined, and many ideas and techniques first motivated by the string theoretical context, such as Kaluza-Klein compactification of extra dimensions, were developed and applied in the supergravity context. The string theory workshop on 'Physics in Higher Dimensions' at the Aspen Center for Physics in August 1984 was particularly emblematic of the above situation: as recalled by Schwarz in his Chapter, many of the participants were working in supergravity theories.

This workshop can be considered the turning point in leading to the string theory renaissance, as it was on that occasion that Green and Schwarz obtained a crucial result in their study of chiral anomalies in type I superstring theory. Chiral anomaly cancellation was a necessary condition for a unifying theory to be 'realistic', that is, to incorporate the Standard Model spectrum of particles and interactions. As explained in Part VII, the weak-interacting fermions are chiral, i.e. occur in specific combinations of spin and momentum: the associated chiral symmetry may be violated at the quantum level (developing an anomaly); the endeavour was to check the absence of such a fatal violation.

Superstring theory could also provide a consistent quantum theory of gravity, in contrast with the supergravity theories that had so far been considered. The result of Green and Schwarz thus opened a concrete path toward string theory unification of the Standard Model physics with gravity, with the effect of producing a radical change of attitude in the theoretical physics community. This achievement marks the end of early string theory: a history initiated with the presentation of Veneziano's formula in the late summer of 1968 and covering a period of exactly 16 years.

1.3 Theory progress: generalizations, analogies and conjectures

As stressed by all the contributors to the Volume, Veneziano's formula represented a turning point in the physics of strong interactions developed in the Sixties in the context of the so-called 'analytic S-matrix' or 'S-matrix theory'. This approach, described in Part II, belongs to the prehistory of the case study considered here. The S-matrix program, pursued by Chew and his collaborators, has been thoroughly investigated from a historical and philosophical point of view in the 1990 book *Theory Construction and Selection in Modern Physics: The S-Matrix* by Cushing [Cus90]. For the purposes of this Chapter, it will be sufficient to recall the approach's general strategy, setting the agenda for the developments leading to the birth of string theory.

Motivated by the difficulties arising in a field theoretic description of strong interactions and inspired by earlier work of Heisenberg, the aim of S-matrix theory was to determine the relevant observable physical quantities, namely, the scattering amplitudes (which formed the elements of the S-matrix) on the basis of general principles such as unitarity, analyticity and crossing symmetry, and a minimal number of additional assumptions. In this context, the problem to which Veneziano's result provided a first,

brilliant solution was the following: to find a scattering amplitude that, in the framework of the S-matrix approach, obeyed also the 'duality principle' known as Dolen-Horn-Schmid duality (DHS duality) or dual bootstrap. As illustrated in detail in the Chapters by Veneziano, Schwarz and Ademollo, this was the assumption, suggested by the experimental data, that the contributions from resonance intermediate states and from particle exchange each formed a complete representation of the scattering process (so that they should not be added to one another in order to obtain the total amplitude). In terms of Mandelstam's variables and using the framework of the so-called Regge theory, the duality principle (as initially stated) established direct relations between a low-energy and a high-energy description of the hadronic scattering amplitude $A(s,t)$: namely, the low-energy description in terms of direct-channel (s-channel) resonance poles, and the high-energy description in terms of the exchange of Regge poles in the crossed-channel (t-channel), could each be obtained from the other by analytic continuation. In this sense, the duality principle represented an explicit and 'cheaper' formulation of the general bootstrap idea dominating the S-matrix program†: that is, the idea of a self-consistent hadronic structure in which the entire ensemble of hadrons provided the forces (by hadron exchange) making their own existence (as intermediate states) possible.

The task of finding a formula for a scattering amplitude $A(s,t)$ that could be expanded as an asymptotic series of poles in either the s-channel or the t-channel, and thus embody the duality principle, initially seemed very daunting. This explains the impact of Veneziano's result and the excitement it generated: his formula realized, in a simple and clear way, the DHS duality in the case of the four-meson scattering process $\pi\pi \to \pi\omega$ (a detailed description of the properties of Veneziano's amplitude is given in Appendix B). It is worth noting, in this regard, that the reconstruction offered by Veneziano and Ademollo, in their respective Chapters, of the steps leading to this achievement in around one year – from the collective works of Ademollo, Rubinstein, Veneziano and Virasoro on superconvergence sum rules to the discovery of the Veneziano formula – provides an illuminating example of the rationale of a scientific progress, one that is characterized by the close interplay of mathematically driven creativity and physical constraints (both theoretical and experimental). In fact, this modality of theory building is the predominant one in the history of early string theory, as the other examples we shall look at will also show.

Veneziano's four-particle amplitude represented a particular solution to

† See Veneziano's Chapter, Section **??**, for details on the distinction between Chew's 'expensive' bootstrap and the 'cheap' dual bootstrap.

the problem of constructing an S-matrix with the required properties for describing hadrons and their interactions. In this sense, it was a 'model' of the general S-matrix 'theory'. Indeed, the whole research activity to which the Veneziano formula gave rise, going under the label of 'dual models', can be seen as a theoretical process evolving from initially specific models (in the above sense) towards a general and consistent theory (of strong interactions, first; of all four fundamental interactions, later)†.

The Veneziano amplitude, which started the above process, itself contained the germs for its successive evolution. It was natural to immediately try to extend it, in order to overcome its limitations, such as, first of all, the limited number of particles considered and their specific type. Moreover, the model violated unitarity because of the narrow-resonance assumption (the approximation corresponding to the fact that only single-particle and stable intermediate states are allowed; see Ademollo's Chapter, Section **??**). It was natural, as well, to search for other models that could include neglected but important physical features, and, more generally, to try to reach a better understanding of the physical theory underlying the models that were being constructed: in other words, to search for a satisfying physical interpretation of the mathematical structures obtained.

Methodologically, the theoretical work initiated with the Veneziano formula mainly advanced by the concurrent action of generalizations, analogies, and conjectures. The following part of the Section is devoted to highlighting illustrative significant threads in this process. As is stressed by the majority of those who participated in the building of the theory, this was essentially a bottom-up activity that had more the character of a patchwork than of an organic construction. Results were obtained by following various alternative paths and (apparently) side issues, and their convergence often seemed almost miraculous. In fact, with hindsight, the story is much less surprising or serendipitous: the cohesiveness of the description obtained can be understood as a consequence of the strong constraints put on the theory by the underlying symmetry, i.e. the infinite dimensional conformal symmetry, as we will discuss later on.

1.3.1 Generalizations

Following Veneziano's paper, generalizations in various directions immediately led to progress both with respect to obtaining a complete S-matrix

---

† On the issue of the relation between models and theories, see later on in this Section, and Section 1.4

and to constructing more realistic models. The most salient of these developments were the following.

- The generalization of the amplitude to the scattering of more than four particles (initially five and subsequently an arbitrary number N of particles), in order to implement the S-matrix consistency condition of 'factorization' and, by means of it, analyze the full spectrum of the physical states. (The factorization condition is that to which unitarity reduces in the narrow-resonance approximation.) Identifying all the physical states was necessary for ensuring the absence of negative-norm or ghosts states (leading to unphysical negative probabilities) and hence maintaining the consistency of the theory. This is the line of research that led to the construction of the Dual Resonance Model, (DRM) (also called the generalized or multiparticle Veneziano model), later understood as the bosonic open string theory and illustrated, in particular, in the contributions to the Volume by Veneziano, Schwarz, Di Vecchia†, Goddard, Mandelstam and Brower.

- The generalization toward more realistic models. These efforts extended from the independent attempts of Lovelace and Shapiro to generalize the amplitude to describe the more accessible scattering of four pions, recalled in their respective Chapters, to the construction of dual models for the scattering of particles with internal symmetry and with spin. By these means were obtained the first dual model including fermions, by Ramond, as well as the immediately following model proposed by Neveu and Schwarz for extending the Lovelace-Shapiro amplitude to an arbitrary number of pions. The Ramond and Neveu-Schwarz models were soon recognized as the two sectors, the fermionic and the bosonic, of the same model, called Ramond-Neveu-Schwarz (RNS) model. The steps leading to these results are reconstructed in detail in the Chapters by Ramond, Neveu and Schwarz.

- The generalization of the N-particle amplitudes, considered as the lowest order or 'tree' Feynman diagrams of a perturbative expansion, to include 'loops' in order to go beyond the narrow-resonance approximation and thus fulfil the S-matrix unitarity condition. This 'unitarization program' was based on the analogy between the narrow-resonance approximation in dual theory and the Born approximation (involving tree diagrams only) in conventional quantum field theory.

† Chapter 11.

Once the general dual amplitudes and the couplings were known, the theoretical activity could focus on the construction of loop amplitudes. From the first attempt in 1969 by Kikkawa, Sakita and Virasoro, this program was actively pursued in the flourishing period of early string theory.

- Another parallel development implying generalization started immediately after Veneziano's formula with the introduction by Virasoro of a different representation of the four-particle amplitude. The Virasoro amplitude was then generalized by Shapiro to the case of N particles by using the technique of the electrostatic 'analogue model' of Fairlie and Nielsen, as illustrated in the Chapters by Shapiro and Fairlie (on the analogue model see also below, next Subsection). This line of research resulted in the alternative dual model known as the Shapiro-Virasoro model that was later understood as describing the scattering of closed strings. In fact, the generalized Veneziano model (interpreted as a theory of open relativistic strings) and the Shapiro-Virasoro model (interpreted as a theory of closed strings) were themselves parts of the same theory (as shown by considering loop corrections)†.

The above general 'guide lines' provided the research framework into which the dual theory of hadrons was initially investigated and extended. At this point, one could be tempted to interpret the first phase of early string theory in the sense of what philosophers of science call 'normal science activity', after the terminology and scheme introduced by the historian and philosopher of physics Thomas Kuhn in his 1962 book *The structure of scientific revolutions*: that is, a problem solving activity under a well established paradigm. The above framework, however, was far from being fixed: the first generalizing steps immediately gave rise, in their turn, to other generalizations and different issues as well as to the introduction of new ideas, methods and formalisms. The history of dual models offers plenty of examples in this sense. Here, we focus on a choice of illustrative cases and analyze some of the most significant analogies, conjectures and discoveries that played a important role in the building of early string theory.

1.3.2 Analogies

As well as generalizations, analogies were also extensively used in the development of dual theory from its very beginning. They were generally inspired

---

† As discussed, for example, by Di Vecchia in his 2007 review [DiV08], Section 8.

by properties and progresses of the mathematical formalism, such as (limiting the examples to those directly relevant to the cases examined below): the Koba-Nielsen integral representation of dual amplitudes, suggesting their relation with two-dimensional surfaces and conformal invariance; the operator formalism (introduced as a way of exhibiting the factorization properties of the multiparticle Veneziano amplitude and thus simplifying the study of the spectrum of states), suggesting the string analogy; and the so-called Virasoro conditions (involving the Virasoro generators of conformal coordinate transformations), suggesting the connection with gauge conditions. A significant part of analogical reasoning, on the other hand, was based on pictorial components, as in the case of the (planar and nonplanar) duality diagrams (discussed, in particular, in the Chapter by Freund), as well as on conceptual similarities. In this last respect, let us emphasize the leading and continuous influence exercised by quantum field theory, its methods and language.

The role of analogical reasoning in extending scientific knowledge, providing explanations and generating new predictions is a traditional issue in the philosophy of science. In particular, analogy is discussed in connection with scientific models and such issues as the relation between a theory and its models or between a model and the portion of physical world it is intended to represent. With respect to models, analogy is first of all the means by which one constructs (what philosophers call) 'analogical models', i.e. representations based on similarity of properties, structures or functions. In the most interesting cases, analogies transform into real identifications: what was initially taken to be merely an analogy and used accordingly is later understood as indicating an underlying essential aspect of the theory. In other words, the analogy provides an 'interpretation' and thus plays a decisive role in the transformation process from initial incomplete descriptions (the 'models') into a full-fledged theory. This is exactly what happened in the case of the 'string analogy', that is, the analogy leading to the crucial conjecture that the underlying structure of the dual resonance model was that of a quantum-relativistic string. In fact, early string theory offers relevant illustrations both of purely analogical models and deeper analogies. Here, we focus on three especially representative examples of the role of analogy in the theory's construction: the (emblematically named) 'analogue model' proposed by Fairlie and Nielsen in 1970, and the already mentioned string analogy and gauge analogies.

- *The analogue model.* This model, the idea and motivations of which are described in the Chapters by Fairlie and Nielsen, is a paradigmatic instance

of the concurrence of formal, figurative and conceptual elements in analogical reasoning. The analogy on which it is based is with two-dimensional electrostatics. The dual amplitude, in its Koba-Nielsen integral form, is described by means of the picture of external currents (the analogues, in the model, of the momenta of the external particles) fed into a conducting disk. The amplitude integrand is then interpreted and calculated in terms of the heat generated by the currents inside the disk, with the important property that the result does not depend on the shape of the conducting surface, owing to conformal invariance. The use of this electrostatic representation was purely analogical but it led to important developments, especially by suggesting appropriate mathematical techniques for loop calculations. As underlined by Di Vecchia and Schwimmer in their 2007 historical review [DS08], it represented 'the first appearance of the two-dimensional worldsheet in a mathematical role rather than just as a picture in the duality diagram'.

Although this is a typical example of a purely analogical model (in the philosophers' sense), its original motivations were deeper. For Nielsen, in particular, the electric analogue was connected with his 1969 view of hadrons as 'threads' or 'strings', the propagation of which was described by two-dimensional surfaces or 'fishnet diagrams' (a fishnet diagram being approximated, in the analogy, by a planar homogeneous conductor). This view resulted from his idea, motivated by the search for a physical interpretation of the N-particle Veneziano amplitude, that strong interactions should be treated by very high-order Feynman diagrams (given the strength of the coupling constants) in field theory and his consequent attempt to visualize such diagrams with many lines.

• *The string analogy.* In the course of 1969, Nambu, Nielsen and Susskind each arrived at the conjecture that the dynamics of the dual resonance model could be represented by that of an oscillating string, though each in an independent way as underlined in their contributions to the Volume. A feature that undoubtedly bore great responsibility for the influence of analogical reasoning in this interpretation of dual theory was the similarity that could be established between the DRM spectrum and that of a vibrating string. The analogy, based on the harmonic oscillator, was clear: on the one hand, the DRM states had been described in terms of an infinite number of creation and annihilation operators of the harmonic oscillator; on the other hand, a vibrating string could also be described by harmonic oscillators whose frequencies, i.e. harmonics, are multiples of a fundamental frequency.

The analogy with the string was, indeed, very deep and came to reveal the

very nature of the theory. In this respect, it is worth emphasizing that the string conjecture originated in an attempt to arrive at a deeper understanding of the physics described by dual amplitudes. In other words, although the dominant framework was that of the S-matrix theory based on the observable scattering amplitudes, a year after the formulation of Veneziano amplitude some people were already searching for a physical interpretation of the dual amplitudes in terms of an underlying dynamics and an appropriate Lagrangian. This is just one of the many examples in the history of early string theory that speak against the appropriateness of interpreting a theory's construction according to rigid schemes.

Although the breakthrough came in 1969, the process leading to the full acceptance of the string interpretation took some time. The correct string action – formulated in terms of the area of the surface swept out by a one-dimensional extended object moving in spacetime (the string world-sheet), in analogy with the formulation of the action of a point particle in terms of the length of its trajectory – was proposed first by Nambu in 1970 and then by Goto. But this interpretation became effectively applied only after the quantization of the string action obtained by Goddard, Goldstone, Rebbi and Thorn (GGRT) in 1973. With this result it became possible to derive, in a clear and unified way, all that had previously been discovered regarding the DRM spectrum by proceeding along various paths and according to a bottom-up approach. It is worth stressing in this connection the following methodological point: the above process shows clearly that neither the framework of the S-matrix (implemented with DHS duality), nor the string Lagrangian framework, if taken separately, could work as an appropriate dominating paradigm for an accurate reconstruction of the evolution of early string theory in its first, flourishing phase.

● *The gauge analogies.* The analogies with gauge field theories were initially investigated and used in the attempt to overcome a problematic feature in the dual theory's spectrum of states, namely, the presence of unphysical negative-norm states, that were called 'ghosts' at the time. The ghost elimination program was implemented on the basis of an analogy suggested by a similar situation encountered in the quantization of electrodynamics, where the unphysical negative-norm states were removed by means of a condition (the 'Fermi condition') following from gauge invariance of the theory, as described in full detail by Di Vecchia in Chapter 11.

This route to ghost elimination in turn generated new developments and also new problems. The analogue, in DRM, of the Fermi condition for quantum electrodynamics was found to be given by the so-called Virasoro con-

ditions, providing the necessary infinite number of gauge conditions. These were associated with the infinite-dimensional symmetry corresponding to the conformal transformations of the two-dimensional world-sheet, that is, the complex plane of the Koba-Nielsen variables. However, for the Virasoro conditions to obtain, the intercept of the leading Regge trajectory had to be taken to equal unity ($\alpha_0 = 1$), and this implied, in particular, that the lowest state of the DRM spectrum had negative mass squared – that is, it was an unphysical particle called tachyon. In fact, beside this problem (solved only with the work by Gliozzi, Scherk and Olive in 1976), there was another puzzling aspect in the spectrum, related to this intercept value†: it contained a massless spin-one particle, not observed to occur amongst the hadrons.

In this case, however, what initially appeared to be a drawback turned out to be a decisive factor in the theory's progress, as already mentioned in Section 1.2. The presence of massless particles with spin in the spectrum (a massless spin-one particle in the above case of open string theory; a massless spin-two particle in closed string theory), related to the conformal symmetry, suggested that these particles could be identified with gauge bosons and gravitons, respectively. This is underlined by Yoneya in particular, who devotes special attention to the issue of the relation between dual models and field theory in his Chapter.

In fact, the correspondence with gauge field theories – a correspondence in which the massless particles played the role of gauge bosons and the gauge symmetry proper of string theory was given by reparameterization invariance of the string world-sheet (implying the conformal invariance of the amplitudes) – revealed itself as much more than a simple analogy. That the analogy had a deeper meaning was indeed suggested by the connection that could be established between dual string theory considered in the low-energy limit and quantum field theory. In the pioneering work by Scherk in 1971, the Dual Resonance Model was studied in the limit given by the slope α' of the Regge trajectories going to zero, which is equivalent to the low-energy limit (this corresponds, in string terminology, to the limit of infinite string tension, see the Introduction to Part VI). Then Scherk and Neveu could show, in 1972, how the massless spin-one states interacted in agreement with gauge theory in that limit. Thus, dual models, which were 'originally very close to the S-matrix approach', had gone 'closer and closer towards field theory', to quote from the review article on dual models and

---

† The other condition that had to be assumed for the elimination of ghosts in the dual resonance model, namely, that the number of spacetime dimensions d had to be equal to 26, is discussed later on in this Chapter, Section 1.3.3.

strings written (in the academic year 1973-74) by Scherk [Sch75], whose crucial scientific contributions to early string theory are illustrated in detail by Schwarz in Chapter 41. It is also worth recalling how, at the same time, in his introduction to the 1974 Physics Reports Volume on *Dual Theory* [Jac74] Fubini, another leading exponent of early string theory, pointed out that the very striking analogy with conventional field theory led one 'to think about a 'strong photon' and a 'strong graviton' in the framework of 'strong gauge theories'.

The above remarks by Scherk and Fubini were made just before the change of perspective according to which string theory was to be viewed no longer as a description of hadronic physics but 'as a quantum theory of gravity unified with the other forces' (see Schwarz, Chapter 3). In fact, as recalled by Schwarz and Yoneya in their Chapters, the decisive step toward this 'revolution' in the theory's interpretation was precisely the idea of extending the earlier analysis of Scherk and Neveu for the case of open strings to the case of closed strings, and in so doing exploring whether the interactions of the massless spin-two particle in string theory, considered in the low-energy limit, agree with those of the graviton in general relativity. The result was quite remarkable: both dual string theories (i.e. both open and closed) could be viewed as short-distance modifications of their field-theory analogues (Yang-Mills theory and general relativity, respectively), thus opening the possibility of interpreting string theory as a unified theory of all four fundamental interactions.

1.3.3 Discoveries, alternative ways and convergent results

As the above developments already illustrate, the history of early string theory provides many examples of how decisive conjectures or discoveries originated and of how important results were obtained and corroborated. From a philosophical point of view, it thus offers novel data for discussing traditional issues, such as, in particular:

• the nature of scientific discovery and the controversial distinction between 'discovery' and 'justification' as two separate moments of scientific activity, and

• the role and characteristics of the evidential support (empirical as well as extra-empirical) in the construction process of a scientific theory.

As regards the issue of discovery, the detailed reconstructions available of the rational steps leading to many of the impressive ideas and conjectures characterizing the development of dual string theory undoubtedly speak in favour of some 'rationality in scientific discovery'. However bold and

requiring of unusual physical intuition as some of the principal ideas may have been, they didn't come out of nowhere or emerge in a purely irrational way. A clear example is offered in the Chapters by Susskind, Nielsen and Nambu, where each of them retraces his own path to the conjecture that the underlying dynamics of dual theory was that of a string. Moreover, the very fact that they arrived at the same conjecture by proceeding in independent ways can be used as a further argument for the rationality of their discovery.

The same can be said for the case of the 'big conceptual leap' leading to a new interpretation of dual string theory as a theory of all fundamental interactions. This was a true change of paradigm, in that it compelled physicists to view the scope and the domain of the theory in a radically different way. It represented a discontinuity also from a sociological point of view, since high-energy physicists and general relativists formed two separate communities at the time. Nevertheless, this paradigm change was the result of theorists working through independent theoretical processes the steps of which are precisely described in the Chapters by Schwarz and Yoneya.

As well as being fertile ground for discussions of issues of discovery, this case also provides an illustrative example of a form of evidential support similar to that provided by (what philosophers call) an 'inference to the best explanation'. Here, the inference is more or less the following: interpreting string theory as a unified theory including gravity, the assumption that the fundamental physical entities are strings predicts the existence of gravity, which is an empirical fact (see Schwarz, Chapter 3, Section **??**). It is worth noting how this kind of inference, not rare in the history of physics, is very similar to that used by Dirac in support of his theory of magnetic monopoles. In that case, the experimental fact was that of the quantization of electric charge, to which the theory of monopoles provided the unique theoretical explanation (at the time), and the hypothesis in need of support was that of the existence of the unobserved magnetic charges [Dir48].

Another significant 'discovery case', illustrative of both the rationale leading to apparently bold guesses and the kind of evidential support motivating a theory's progress, is that of the so-called critical dimension: that is, the discovery that consistency conditions of the Dual Resonance Model required the value $d = 26$ for the spacetime dimension (reducing to $d = 10$ dimensions in the case of the Ramond-Schwarz-Neveu model).

The critical value $d = 26$ was originally obtained in two independent ways. The first one was by Lovelace who, in a work published in 1971, addressed a problematic singularity arising in the construction of the nonplanar one-loop amplitude in the framework of the unitarization program discussed in Section 1.3.2. As Lovelace describes in his Chapter, the problem posed by

the singularity was solved by turning it into a pole, that is, by interpreting it as being due to the propagation of a new intermediate particle state. This he conjectured to be the Pomeron – the particle that was later understood as the graviton†. This solution was arrived at through adjusting the theoretical description by considering the possibility that the spacetime dimension d might be different from 4 and treating it as a free parameter (together with leaving arbitrary the number d' of degrees of freedom circulating in the loop): the result was that the singularity became a pole only for $d = 26$ (and $d' = 24$).

Soon after, the same result for the critical dimension issued through another route: namely, from the examination of the DRM physical spectrum of states in the context of the ghost elimination program (discussed above in this Section. As described in detail in the Chapters by Goddard and Brower, by using the infinite set of positive-norm states found by Del Giudice, Di Vecchia and Fubini (the so-called DDF states) it was possible to prove that the DRM had no ghosts if the spacetime dimension d was less or equal to 26. This result, known as the 'no-ghost theorem', was obtained by Brower, and by Goddard and Thorn, in 1972. For $d = 26$, it was shown that the DDF states could span the entire Hilbert space of physical states, as illustrated in detail in Chapter 11 by Di Vecchia.

In the critical dimension, consistency was thus satisfied - but at the high price of an extra 22 space dimensions. This was a rather unrealistic feature, especially for a theory that was intended to describe hadronic physics. Nonetheless, the extra dimensions became gradually accepted, owing to the fact that the critical dimension result received further evidential support from successive theoretical developments. In particular, a third decisive evidence was provided by the 1973 work of Goddard, Goldstone, Rebbi and Thorn where the $d = 26$ condition was obtained from the canonical quantization of the string in the light-cone gauge: together with the condition $\alpha_0 = 1$ for the intercept value, it resulted from the requirement of Lorentz invariance in the quantum theory, as thoroughly described by Goddard in his Chapter, at Section **??** (see also the Introduction to Part IV, Section **??**).

In fact, with hindsight, the critical dimension is a consequence of the conformal symmetry of string theory. More precisely, an 'anomaly', i.e. a violation of a symmetry, arises in the implementation of this symmetry in the quantum theory unless $d = 26$. Prior to a definitive understanding of the conformal anomaly – which owed to the 1981 work by Polyakov – the

† See, in regard, the Introduction to Part III, Section **??**. An interesting recollection of how Lovelace arrived at his result can be found in the Chapters by Lovelace and Olive.

critical dimension condition could be found only on the basis of 'side effects': that is, as a condition required by unitarity of the theory or by Lorentz invariance in the quantization of the string action†. Thus, what had appeared to be a surprising convergence of different calculational procedures to one and the same result could be seen as a natural consequence of the theory in its fully-fledged form. It is surely worth stressing that the convergence of results obtained in alternative ways and from different starting points provided important evidential support, that motivated persevering with the theory notwithstanding the presence of unrealistic features such as extra space dimensions.

1.4 Conclusion

To sum up, many fruitful lessons may be drawn from a closer look at the developments of dual string theory for philosophical reflections on both scientific practice and theoretical progress. The history of early string theory is, of course, far richer than what has been possible to highlight in this Chapter. Nevertheless, the hope is that some points of notable philosophical interest might come out clearly even from such a partial presentation. To finish, let us list them:

• Pluralistic scientific methodology. The case study provided by early string theory shows, once again, how traditional methodological schemes for describing scientific progress (inductivism, falsificationism, normal science activity intertwined with revolutionary changes of paradigms, competing research programmes, and so on) are both too rigid and limited to appropriately account for the actual dynamics of a theory building process, at least if taken separately.

• Models and theories. The development of string theory out of the original dual models offer precious examples to the current philosophical discussion on how to characterize a scientific model, a scientific theory, and the relation between models and theories – in particular, for what regards the so-called semantic conception of scientific theories, according to which a theory is defined in terms of the collection of its models.

• The role and nature of the evidential support in scientific progress. The phenomenological origin of dual models, initially developed in the context of the S-matrix ideology and on the basis of the experimental hadronic data available, as well as the falsification of dual string theory as a description of hadronic physics, both show the relevance of the experimental evidence in the evolution of early string theory. On the other hand, cases such as those

† For more details on this point, see for example[DS08], Section 5.

discussed in Section 1.3.3 clearly illustrate the decisive role of extra-empirical evidential support in the theory building process.

• The influence of analogical reasoning in theoretical developments and, in particular, in the discovery process, as described in Section 1.3.2.

• The nature of new ideas, such as that of the string, in the light of their evolution. In the case of duality and its role in string theory, for example, it is important to clear out the different meanings that are attributed to this notion. The historical reconstruction of early string theory allows one to understand how the original DHS duality or 'dual bootstrap', at the core of the dual models, is in fact a consequence of the conformal symmetry of the theory.

Acknowledgements

(To be completed)

I am very grateful to Andrea Cappelli, Filippo Colomo and Paolo Di Vecchia for their constant support. I would like to thank also Roberto Casalbuoni Leonardo Castellani, Kerry McKenzie, Antigone Nounou, Giulio Peruzzi and Chris Wüthrich for their very helpful comments and suggestions.

Note on the references

....

References

[Cus90] Cushing, J. T. (1990). *Theory Construction and Selection in Modern Physics: The S-Matrix* (Cambridge University Press, Cambridge).

[Dir48] Dirac, P. M. A. (1948). The theory of magnetic poles, *Phys. Rev.* **47**, 817–830.

[DiV08] Di Vecchia, P. (2008). The birth of string theory, in *String Theory and Fundamental Interactions*, ed. M. Gasperini and J. Maharana (Springer, Berlin), *Lect. Notes Phys.* **737**, 59–118.

[DS08] Di Vecchia, P. and Schwimmer, A. (2008). The beginning of string theory: a historical sketch, in *String Theory and Fundamental Interactions*, ed. M. Gasperini and J. Maharana (Springer, Berlin), *Lect. Notes Phys.* **737**, 119–136.

[Jac74] Jacob, M. ed. (1974). *Dual Theory* Physics Reports Reprints Book Series Vol. 1, (North Holland, Amsterdam).

[Sch75] Scherk, J. (1975). An introduction to the theory of dual models and strings, *Rev. Mod. Phys.* **47**, 123–163.