Alma Mater Studiorum · University of Bologna

School of Science
Department of Physics and Astronomy
Master Degree in Physics

String Theory as a Lakatosian Research Programme

Supervisor:
Prof. Vincenzo Fano

Submitted by:
Vincenzo Nespeca

Academic Year 2021/2022

Abstract

Theoretical physics is experiencing one of the worst crises in its history. The main objective is to find a quantum theory of gravity, but the lack of experimental data at energies near the Planck scale makes theoreticians advancing blindly, without any guide in the theory building process .

In the last four decades one theory has been considered the best candidate for a theory of quantum gravity (and in general for a unifying theory of all forces): string theory, which is still dominating the theoretical physics research landscape.

Despite the efforts of a huge number of scientists, string theory has never been able to give any testable prediction, so it is not experimentally testable.

The domain of a theory which fails to give any testable prediction has divided the theoretical physics community: string theorists are still convinced their research programme is the only way to achieve the goal, while their opponents complain the monopoly of a research field which failed to meet the expectations. Furthermore, theoretical difficulties inside the string theory research programme led to a division in the very string theory's community.

All these issues led to a stalemate that has lasted for decades, and the consequences are not only scientific but philosophical, methodological and sociological as well.

In this situation, philosophy of science can be of fundamental importance as an instrument for a deeper understanding of the scientific process, suggesting a rational appraisal of the status of contemporary physics so that scientists can continue their research with greater awareness.

In this work I analyze the string theory research programme from a historical and philosophical perspective. It is organized as follows: in the introduction I am going to give an overview of the situation of theoretical physics, of its recent history and the problems it is facing; in chapter 1 I will summarize the most important philosophical paradigms and achievements in the philosophy of science, deepening the thought of Imre Lakatos; in chapter 2 I will give a detailed summary of string theory's history; in chapter 3 I will interpret string theory research programme in the light of Lakatos methodology; in chapter 4 I will present and discuss Richard Dawid's philosophy and its application to the string theory research programme; in the conclusion I will give the final results of the analysis.

This work is intended as an almost self-contained text, being my intention to provide all the tools for the development of a personal perspective.

Contents

Introduction

The Standard Model is the best theory we have to describe the world of fundamental particles and interactions. Its birth can be traced back to 1927, when the first formulation of Quantum ElectroDynamics (QED) was presented, while its last experimental confirmation was in 2012, with the discovery of the Higgs boson. The name 'Standard Model' is in use in the scientific community since 1979, and its theoretical systematization was completed in 1973 with the introduction of a third generation of quarks[63]. In those fifty years a large number of new particles and phenomena was discovered, thanks to the huge amount of experimental data supplied by the accelerators and other technologies developed for the purpose. The Standard Model is without any doubt the most fruitful research programme of the last century, maybe of the whole history of fundamental physics. To date it has been passed all the tests, and still continues to be confirmed.
The last experimentally confirmed theoretical prediction was the existence of quark and bottom quarks in 1973. Their existence is the last theoretical prediction that was verified not only in the framework of the Standard Model, but in the whole research in theoretical particle physics, either because novel predictions failed experimental tests or because they are not verifiable by using contemporary technologies. I am writing fifty years later. It is natural to wonder how it is possible that such an incredibly successful period of research has been followed by such a terrible period of theoretical infertility.

The Standard Model is our best description of the microscopic world. On the other hand, General Relativity is our best theory explaining macroscopic phenomena. It was presented by Albert Einstein in 1916, and up to now it has received large empirical confirmation. The last experimental corroboration was the detection of gravitational waves in 2015. General Relativity, soon after its creation, was studied by cosmologists to investigate the birth, evolution and structure of the Universe. All the efforts led to the Big Bang hypothesis and the Lambda-CDM model, the standard model of Big Bang cosmology (or concisely the Standard Cosmological Model). This is the cosmological model commonly accepted, and it explains and successfully predicted many phenomena, but it is not free of problems. Some of them were solved adding to the model the concept of inflation, shared by the most of the scientific community.
The 'CDM' in the standard cosmological model's name stands for 'Cold Dark Matter'.

Dark matter is a new kind of matter whose existence is conjectured in order to solve some experimental anomalies. In particular, we know from Newton's gravitational law and dynamics that measuring the velocities of stars orbiting around the centre of a galaxy, we can predict the mass distribution of the galaxy. Now, if we compare the calculated mass of a galaxy with the experimentally measured mass, estimated by considering all the matter we can see, we found a discrepancy. This comparison was made for a large number of galaxies, and it has always been found that the measured mass is much lower than the calculated one, by an order of ten magnitudes. The best explanation of this result is that the calculated mass is the correct one, so there is a certain amount of mass that we are not able to see. Now, all the matter we know radiates, so if it is a known kind of matter we should be able to see it. The solution is to conjecture a new kind of matter which does not radiate or absorb light. For this reason, we call it 'cold dark matter', or briefly 'dark matter'. Observing the data, it represents the most of the matter present in galaxies.
But there is more. Observations revealed that the expansion of the Universe is accelerating, while General Relativity, considering the estimated energy in the Universe, predicts it should be decelerating. Being General Relativity well corroborated by experiments, we are not willing to reject it, for none of the above two reasons. One more time, the solution commonly accepted by the scientific community is to conjecture the existence of a new kind of energy, the so called 'dark energy', responsible for the expansion of the Universe.
The amount of dark matter and dark energy can be estimated, and it seems that the Universe is composed by about the 70% of dark energy, 26% of dark matter and only 4% of ordinary matter. It seems we understand only the 4% of the Universe we live in, and not even perfectly [52].
One of the best candidates identifiable with the dark energy is the cosmological constant λ, a term firstly introduces by Einstein in its field equations to keep the Universe static. However, after Hubble's discovery of the expansion of the Universe, λ was abandoned. In 1998, through observations of type 1A Supernovae, it was discovered that the expansion of the Universe is accelerating instead of decelerating, so λ was reintroduced to be identified with the dark energy responsible for this acceleration.
In the framework of quantum field theory, on the other hand, dark energy is identified with the zero-point energy, that is the vacuum energy. Quantum fluctuations make the vacuum energy be non-zero, because of the continuous creation and annihilation of virtual particles predicted by the theory.
Experimentally speaking, today the best measured value of the cosmological constant is about 10^{-120} times smaller than the calculated value from quantum field theory. It is often called "the worst prediction ever made". The value of the cosmological constant seems to be a very little positive number, close zero but not yet zero.

In order to understand the difficulties that contemporary physics is facing, we need to go deeper inside the Standard Model and summarize its properties.

The Standard Model puts together the electroweak theory (EWT) - unifying QED and weak interactions - and Quantum Chromodynamics (QCD). All these theories are based on the framework of quantum field theory, where quantum fields are seen as the elementary objects and particles are zero-dimensional quantized fluctuations of the corresponding fields. QED is the relativistic and quantum version of electromagnetism, describing the interactions between light and charged particles. The weak interactions involves subnuclear particles and explains the radioactive decay of atoms. QCD describes the strong interaction that binds together the quarks in the nucleus. The Standard Model describes three of the four fundamental interactions, being unable to give a microscopical description of gravity. The model needs the concept of spontaneous symmetry breaking to be consistent. This phenomenon gives the masses to all the fundamental particles and breaks the symmetry of the electroweak interaction down to the two distinct forces I have just presented.
Mathematically speaking, the Standard Model is based on the gauge group $SU(3)_{QCD} \times SU(2)_{weak} \times U(1)_{hypercharge}$. The $SU(2)_{weak} \times U(1)_{hypercharge}$ is the Yang-Mills theory describing the electroweak sector, and it is spontaneously broken to $U(1)_{EM}$ by the vacuum expectation value of the Higgs boson at $\sim 246Gev$, giving back electromagnetism [51]. Sometimes we find the group $SU(2)_{weak}$ denoted by a subscript "L", that is $SU(2)_L$, which stands for "left" and is due to the chiral character of the interaction, acting only on left-handed particles. $SU(3)$ is also a Yang-Mills theory and describes the QCD sector.
All the particles can be divided in two categories, fermions - spin 1/2 particles, constituting matter - and bosons - integer spin particles, mediating the interactions; fermions can further be divided into leptons and quarks. The left-handed leptons $(e, \nu_e, \mu, \nu_\mu, \tau, \mu_\tau$ and left-handed quarks (d, u, s, c, b, t) are charged under the weak interactions, so they can be organized in three generations of $SU(2)$ doublet pairs:

$$L^i = \begin{pmatrix} \nu_{e_L} \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_{\mu_L} \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_{\tau_L} \\ \tau_L \end{pmatrix}$$

The right-handed fermions are not charged under the weak interaction, so they are SU(2) singlets:

$$e^i_R = \{e_R, \mu_R, \tau_R\}\,, \quad \nu^i_R = \{\nu_{eR}, \nu_{\mu R}, \nu_{\tau R}\}\,, u^i_R = \{u_R, c_R, t_R\}\,, \quad d^i_R = \{d_R, s_R, b_R\}\,.$$

where the 'i' indexes the three generations in both cases. The right-handed neutrinos are never been observed, but they are supposed to exist in the model (ibid.). The generations differ only for the masses of the particles. We know there must be almost three generations of fermions, but we do not know if there are more than three; the number of generations is not predicted by the model, so they are an experimental input.
The bosonic particles include photons, mediating the electromagnetic force, W and Z

bosons, responsible for the weak interaction, gluons, mediating the strong interactions, and the Higgs boson, a complex scalar field which gives masses to the particles and acquires a vacuum expectation value that spontaneously breaks the $SU(2) \times U(1)$ group to $U(1)$.
The Standard Model, with its twelve particles and four interactions, is able to give finite results for every experiment for which we can neglect the gravitational force. In any case, it has 19 free parameters, representing the masses of particles, the intensity of the interactions and the mixing angles. Their values are not predicted by the theory, so we need to insert them by hand, that is measuring them through experiments.
I have just given a qualitative introduction of the fundamental particles and forces constituting the Universe, but which are the numerical values describing them? First of all, we know that the weak interaction is about 10^{25} times stronger than the gravitational force, the electromagnetic one 10^{36} times and the strong interaction 10^{38} times.
The vacuum expectation value of the Higgs boson is $\sim 246Gev$ and corresponds to the energy scale at which the electroweak interaction is broken. The energy scale at which all the interactions are supposed to unify is about $\sim 10^{16}Gev$, three orders smaller than the Planck energy ($10^{19}Gev$).

After this brief summary of the two main research programmes dominating the last century, we can present the problems the physics community is called to solve. To this aim we will follow the accounts given by Lee Smolin [52] and Peter Woit [63] in their celebrated books. We are going to combine them, reserving a deeper discussion in the next sections and in the appropriate thematic moments. I start with questions put forward by Woit in the eight chapter of his book, concerning the Standard Model alone:

- *Why $U(1) \times SU(2)_L \times SU(3)$?*
  The gauge group the Standard Model is based on is not predicted by any fundamental principle, it is constructed by hand for a simple reason: it works. Historically speaking, scientists firstly found that the $U(1)$ gauge group could account for the electromagnetic force, than $SU(2)$ was constructed to explain the weak interactions and finally $SU(3)$ was discovered as a suitable group to describe the strong interaction. They are not a result of some physical reasoning, but rather mathematical instruments able to give a good description of the world. Obviously, we would know why the gauge group is that one, so we wish that a unifying theory will give an answer to this question.

- *Why do the quarks and leptons of each generation come in a certain pattern?*
  The fact is that they are three copies of the representations [53]:

  $(1, 2, -\frac{1}{2}) \oplus (1, 1, +1) \oplus (3, 2, +\frac{1}{6}) \oplus (\overline{3}, 1, -\frac{1}{2}) \oplus (\overline{3}, 1, +\frac{1}{3}),$

  so we ask why these representations and not others.

- *Why is the electroweak interaction chiral?*
  The chirality of the weak interaction is an experimental outcome. In 1956-57 it was detected a violation of parity in the β-decay of the cobalt-60. This result was not predicted by the model, so it was inserted after the discovery. We would like to know why the weak interaction violates it, and why it does it in that way.

- *Why three generations?*
  We have already seen that the number of generations is not a prediction of the model, but it is an experimental fact, so we need to know if there are other generations and why their number is what it is.

- *Why is the Θ-parameter zero or very close to zero?*
  The θ-parameter represents the QCD vacuum angle, a parameter in QCD whose null value implies that strong interactions conserve parity. Recent experiments suggest a soft-violation of parity by the strong interaction, which means the θ has likely a very small value, so we want to know if this is true and why it is so small.

- *What determines the masses and mixing angles of the quark and leptons in the theory?*
  As we have just seen, the Standard Model contains 19 free parameters. We are not able to detect any specific pattern, nor a mathematical or physical reason for their values. This problem is related to the next one, but we want to keep them separated for reasons we will present in the fifth chapter.

- *Fine-tuning, hierarchy problem and naturalness.*
  These do not represent problems strictly speaking, because they do not imply any inconsistency. These are mainly problems concerning how we expect a theory should be.
  Fine-tuning corresponds to a sensitivity of the values of physical observables to the variation of parameters in the theory. We have fine-tuning when we must adjust very precisely a parameter in order to get the correct measured value. The most important case of fine-tuning in the Standard Model is represented by the value of the Higgs mass. This mass is UV sensitive,that is it undergoes quantum corrections which are quadratically divergent at low distance scale, so not knowing the UV completion of the Standard Model we take the energy cut-off Λ to be close to the Planck scale. We have the following expression for the bare mass: $m^2 \approx \Lambda^2 + m_P^2$. Now, being $m_P \approx 125Gev$ and $\Lambda \sim M_{PI} \sim 10^{19}Gev$ we would need $m^2 = (1 + 10^{-34})\Lambda^2$. We see that in order to get the measured value of the Higgs mass we need an incredible precise cancellation between the bare mass and quantum corrections[1]. Most of the time the bare parameters and the measured values are

[1] Cf. ([51], p.410-411).

very close, but in cases like the one above they are very different, needing a fine-tuning of parameters in order to give the correct experimental result. The Higgs mass is not protected by any symmetry, so it could have a value much larger than the weak scale; nevertheless, it has a small value compared to the Planck energy scale. This huge difference in the energy scale of the bare mass and renormalized mass suggests that some mechanism is working to cancel precisely the quantum corrections, a mechanism concerning the UV completion we do not understand. The problem of the large discrepancy between the Higgs mass and the Planck scale is referred to as the hierarchy problem, also leading to wonder which physics lives in this huge range of energy. This one is a problem about 'technical naturalness' not to be confused with the more general concept of naturalness. Technical naturalness demands for an explanation of the fact that a certain phenomenon occurs at an energy scale much different with respect to the maximum energy scale at which the theory is still valid. On the other hand, naturalness is a general principle in the light of which a theory is considered to be 'natural' if there are no highly fine-tuned parameter values. In order to avoid fine-tuning, naturalness demands that the parameter values should be of order ~ 1 in the appropriate energy scale. A natural theory is more appealing than a non-natural one both because we do not need to fine tune parameters with too much precision, and because fine-tuning means that we are missing something, that the theory is not complete.
Another concept related to naturalness is 'hierarchical naturalness', demanding an explanation of the discrepancy and hierarchical order of fundamental interaction's energy scales.
An important parameter to be mentioned talking about fine-tuning and naturalness is the cosmological constant λ. Its estimated observable value is $\sim 10^{-9}\frac{J}{m^3}$, while the value predicted by quantum field theory is $4.6330910^{113}\frac{J}{m^3}$. This is the so called "vacuum catastrophe", and as we have already mentioned it is one of the biggest problems in contemporary physics. The problem is that we measure a very small value of the vacuum energy respect to the calculated one, so - in the QFT framework, where λ is identified with the vacuum energy - there should exist some physical mechanism that lowers down the value obtained considering quantum fluctuations in order to give the correct one. On the other hand, if we consider λ as a fundamental constant not arising from any physical process, its very small value seems to demand for a justification, because a very small deviation from it would have made our existence impossible, inhibiting the formation of galaxies. This point will be crucial when discussing the anthropic principle.
Naturalness represents an important principle widely used in contemporary physics, also providing a tool for the appraisal of theories, but its methodological status is debated[2].

[2] A popular example is given by [24].

Now I present the problems listed by Lee Smolin in the first chapter of his book, concerning the most fundamental and general problems in theoretical physics:

- *Combine General Relativity and quantum field theory in a single theory pretending to be the complete theory of nature. This is called the quantum gravity problem.* This is the main objective of contemporary theoretical physics. General Relativity describes very well large scale physics, where we can neglect quantum effects, while quantum field theory explains microscopic physics up to the Planck scale, above which we can neglect gravitational effects. Unfortunately, the unification of the two frameworks in a quantum theory of General Relativity gives a non-renormalizable theory, meaning that the perturbation series diverge for energies at the Planck scale. So, if we want to understand phenomena at such a scale, we need a UV completion, a finite (or renormalizable) quantum theory of gravity. We need such a theory because there are events in the space-time where both gravitational and quantum effects become important, so we cannot use neither General Relativity nor Quantum Field Theory. In order to understand black hole's singularities, the Big Bang and other phenomena, a quantum theory of gravity is necessary. Huge efforts had been made in this directions, but with very little results. The most important attempt is without any doubt string theory, by many considered the best candidate for a quantum theory of gravity. This argument, being the most important one justifying string theory research, will be largely discussed in what fallows.

- *Solve the problems relating the fundamentals of quantum mechanics either understanding the meaning of the theory in its actual formulation or inventing a theory that makes sense.* Quantum mechanics is a very counter-intuitive theory, and someone might say that this problematic is related to our being macroscopical observers, not accustomed to quantum effects. This is certainly true, but even if the problem seems to be metaphysical rather than physical, many people feel unsatisfied with its description of the microscopical world. Wave function, superimposition, entanglement and the asymmetry in the roles of experiments and observers are all concepts which suggest the lack of some fundamental element in our understanding. The debate concerns the intrinsic probability of nature and the realism of physical states, and a lot of questions remain unanswered in this incredibly successful framework. There are scientists not minded with these problems, accepting the strange nature of microscopic objects without asking too many questions, while others think resolving these problems is of fundamental importance, feeling that a complete understanding of the quantum world could help in the search for a quantum theory of gravity, and in general to get the right framework capable to answer all the questions we are asking here. A lots of different interpretations and answers has been given, but we need further explanations to shed light on this mysteries.

- *Determine if the various particles and forces can be unified or not in a theory explaining all of them as a manifestation of a single fundamental entity.*
  Nature seems to be quite complex, with all its particles and forces, so we ask if they can be unified all together in a unique element explaining the complexity we experience.

- *Explain dark matter and dark energy. Or, if they do not exist, determine how and why gravity modifies at large scales. More generally, explain why the parameters in the standard cosmological model, dark energy included, have those specific values.*
  I have already argued that in order to explain some anomalies in the large scale behaviour of the Universe we need to introduce two new concepts, dark energy and dark matter, which seems to constitute the greatest part of the Universe we can observe. We do not know anything about the nature of these objects, and their understanding is one of the main questions to be answered by the physics beyond the Standard Model. We have many theories with many candidates corresponding to them, and String Theory proposes its own too.

These are the main problems the scientific community is called on to solve, and we will encounter a part of them when dealing with string theory. Finding a theory capable to solve all of them is a big challenge, which may need help from various professions. The way forward is far from being clear, so we need a methodological evaluation at each step in order not to get lost in the way. The manner we do this (and the manner we continue on the track) is a matter of choices, and like any wrong choice it has a number of consequences. In this context the consequences are mainly methodological and sociological, and they are of great importance for what concerns the development of theoretical physics and the scientific process in general.
Philosophy of science investigates the methods and rationality of science. This subject has always had a marginal role in the pure science, not being taught in scientific universities, so that most of scientists know nothing about it, and they do not seem to be interested as well. It is certainly understandable, because philosophy of science often gives a retrospective view on scientific facts, and does not participate to the technical development of science. But things are changing. As I have already mentioned, we are witnessing a great crisis in theoretical physics. I think it is important to understand how it happened, and what the consequences might be. Methodological debates between scientists are already on the scene, this underlining the importance of a deeper reflection about the scientific methodology in contemporary theoretical physics. In particular, both the lack of experimental support and the development of string theory's research programme might lead to a drastic modification of scientific methods and philosophical thinking, so we must be able to give our rational evaluations and act accordingly. To do so, we are going to explore the tools philosophy of science provides us with.

Chapter 1

Philosophical apparatus

In this chapter I am going to present a brief summary of modern philosophy of science, focusing on Lakatos' methodology of scientific research programmes. To this aim, I strictly follow [33], a collection of Lakatos' papers.

1.1 Summary of modern philosophy

1.1.1 Justificationism

Justificationism represented the dominant tradition in the history of philosophy of science. It claimed that only proven propositions deserved to be considered as scientific knowledge, so knowledge was identified only with proven knowledge.
Justificationists can be divided into two currents of thought: classical intellectualists (or classical rationalists), admitting not only experimental demonstrations but also proofs by intellectual intuition and deductive logic, and classical empiricists, accepting axioms only in the form of 'factual propositions' expressing the 'hard facts', whose truth value was to be established by experience. In their mind, these propositions constituted objective knowledge, so even if they were axioms they pretended to be able to assess to them an objective truth value. These factual proposition constituted the so called "empirical basis" of science, the basis of reliable knowledge from which we can construct our theories. In any case, to start only from the empirical basis and build up theories, they needed a logic much more powerful than deductive logic, so their theory building process was based on 'inductive logic', admitting the generalization of singular facts to universal propositions thus allowing to construct theories starting from factual proposition, that is from the empirical basis.
Not all justificationists accepted inductive logic, but all of them thought that an hard fact could disprove a universal theory. For them *scientific honesty* meant asserting only

proven statements. In any case, both rationalists and empiricists were defeated, because it was demonstrated that no theory can be proven with certainty. Some justificationists tried to save the possibility to evaluate a theory, switching from the possibility to prove the truth value of a theory to the possibility of assessing at least a probability for its truth value, so they reformulated scientific honesty stating that one has to propose only those theories which appear highly probable in the light of some specific evidence. In any case, they were defeated too, because it was shown that all theories have zero probability, so no theory can be considered more probable than another.
We have just introduced above an important concept, that is scientific honesty, representing the declared attitude of scientists towards theories.

1.1.2 Dogmatic falsificationism

Dogmatic falsificationists admitted the existence of an empirical basis whose truth value can be objectively established, but they deny the possibility of using inductive logic to construct theories starting from it, that is to transfer the truth value of the empirical basis to theories. Falsificationists believed in fallibilism: they recognized the fallibility of all scientific theories. This attitude is very different from classical rationalists and empiricists, who thought to be able to prove once and for all the truth value of a theory. The crucial difference between justificationists and falsificationists was that the latter considered counterevidence the only arbiter able to judge a theory. This means that they did not consider proofs definite demonstrations, but they recognized that all theories are fallible, so only disproofs are considered able to express a judgement. Dogmatic falsificationists recognized that no theory can be proven, and all theories are conjectural. In any case, they thought that even if theories cannot be proved, there exists an empirical basis of facts that can be used to disprove them. In their point of view, scientific honesty consists of specifying an experiment that, if its results contradicts the theory, the theory has to be abandoned. As Lakatos points out, in their viewpoint a factual proposition was considered falsifiable 'if there are experimental and mathematical techniques *available at the time* which designate certain statements as potential falsifiers'([33], p.12-13). Unfalsifiable propositions are considered metaphysical, so they are not scientific statements. This represents the so called *demarcation criterion*, which enables to recognize scientific theories from pseudo-scientific theories: if one does not enable an experiment to definitely overthrow his theory, it cannot be considered scientific knowledge.
Dogmatic falsificationists give to scientists definite roles: the theoretician make conjectures, while the experimenter tries to falsify them. Conjectures are fundamental because falsificationists deny the possibility of using the empirical basis and inductive logic to build up theories with certain truth value, so science must proceed through conjectures. Also, those conjectures cannot be proved by the experimenter: they can only try to

disprove them. If a single or universal statement passes the test, it is considered 'not falsified' up to a new experiment, and so on until it is finally falsified and replaced by a new theory. For dogmatic falsificationists, the process of science can be represented by a succession of theories each one disproved by hard facts. It is important to underline that a new theory should both account for all the successes of the replaced one and give new predictions.

Up to now I have mentioned that justificationism is based on wrong beliefes, because all theories are unprovable and improbable; dogmatic falsificationism is also untenable, because its demarcation criterion is too narrow. Lakatos, for this purpose, claims that it is based on two false assumptions.

The first one is that there is an evident difference between theoretical and observational propositions. This is a false assumption because observational propositions are necessarily theoretical as well, they cannot be free from theory. This is because observations are based on theories too, that is experimental technologies are based on a so called 'observational theory'; furthermore, data need to be interpreted. The point is that pure observations, made only by human senses, do not exist; they necessarily rely on the reliability of the instrument, the observational theory and the interpretation of data. Lakatos gives as an example Galileo's observations: they were made using a telescope, their reliability must be addressed to the reliability of the optical instrument. In this sense there does not exist a natural distinction between observational propositions and theoretical statements.

The second false assumption is that an observational proposition is true, because it is proved by facts. This assumption is false because experiments are not able to prove anything, they do not allow us to assign an indubitably truth-value to any observational statement. Proposition cannot be derived directly from facts, but only from other propositions, so one cannot prove anything directly from experiments. For this reason, there is no distinction between the hard facts of the empirical basis and the unproven theoretical statements, but all propositions are theoretical and, as a consequence, fallible.

As we have already mentioned, we can recognize a *demarcation criterion* stating that a theory can be considered scientific only if it allows some observational facts to disprove it, the so called 'potential falsifiers' constituting the so called 'empirical basis'. In other words, dogmatic falsificationists thought there always exist observations able to disprove a scientific theory. This additional assumption is, once again, a false one. We know that also the best theories we have do not forbid any observation at all.

This is a major point, because a retrospective look on history of science may fail to recognize the real methodology followed by past successful research programmes. If we look deeper in history, we can see that none of the most successful theories followed falsification method: every time a theory seems to be falsified, scientists can save it throughout auxiliary hypothesis. An illuminating example is the deviation of a planet from a calculated path in the framework of newtonian research programme. In the falsificationists' viewpoint, such a counterevidence should overthrow the whole newtonian

theory, but obviously scientists do not follow this way of reasoning. We are not supposed to abandon a successful theory because of one single counterevidence. Rather, scientists propose a further hypothesis to explain the observed counterevidence, for example a new planet influencing the path. This reasoning may be used every time a scientist faces a counterevidence, and usually this is the way things go.
This example shows that in similar cases one can conjecture some other phenomena influencing the one at issue, so that scientists can avoid to reject the theory. An event can be considered a counterevidence only if we are sure that no other events are influencing it. A certain theory never contradicts a certain statement alone, but at most it contradicts that statements together with another implicit statement assuming that no other causes are at work. In these cases it is the theory together with this second statement, also called a *ceteris paribus* clause, that may be refuted. In any case, the theory can always be saved by replacing this clause with another one making the theory to fit the experimental results.
The ceteris paribus clause cannot belong to the empirical basis, because it cannot be observed, so such theories have no empirical basis. Consequently, in the dogmatic falsificationists' viewpoint, these theories are not falsifiable and so not scientific. The final result is that, following this line of reasoning, they will consider all the most successful theories in the history of science as pseudo-scientific. Dogmatic falsificationism is then unable to account for the success of science, because its reconstruction of scientific method and development is wrong, resulting in a metaphysical character of great physical theories [1].
Not only can theories not be verified or considered to be probable, they cannot be disproved either. This is a very strong statement, because thinking that neither theoretical nor observational statements can be proved may lead to scepticism, to the idea that there does not exist any knowledge at all and so the scientific process would be an irrational succeeding of miracle theories. In order to give a rational account of scientific process we need to find a loophole from fallibilistic arguments. To this aim, we are going to discuss a new form of falsificationism, the so called *methodological falsificationism.*

1.1.3 Methodological falsificationism

One may wonder how scientific criticism can be saved from scepticism. If all theories are unprovable, improbable, and undisprovable, how can science work? In which grounds scientists decide to retain or to reject a theory? If all propositions are fallible, how can one evaluate scientific propositions?
In order to answer this questions, I start by discussing *conventionalism.* We can rec-

[1]Falsificationists' method does not resemble the real method used by scientists: as we will investigate further talking about Lakatos methodology, research programmes develop 'in an ocean of anomalies'.

ognize two schools of conventionalism, *conservative conventionalism* and *revolutionary conventionalism.* The first one states that once a theory has demonstrated to be successful, passing empirical tests, scientists can make a 'methodological decision', deciding conventionally to not allow anomalies to disprove the theory. This can be done throughout auxiliary hypothesis or 'conventionalist stratagems'. By the way, as we have already seen, saving a theory is not so difficult. The main point is that conventionalists are aware of their arbitrary choice, recognizing the conventional value of their judgement. The problem with this brand of conventionalism is that being any choice arbitrary, one has no evident reason to eliminate a theory. Once a well established theory is chosen to be defended from anomalies, empirical evidence loses its power, because one can always find a way to account for it.
This criticism led to revolutionary conventionalism, which can be divided into to other schools of thought, that are *Duhem's simplicism* and *Popper's methodological falsificationism.* Duhem shared with conventionalists the idea that theories are not abandoned because of anomalies, but also argues that adding auxiliary hypothesis and modifying theories in order to save them can lead to the loss of *simplicity* and the necessity of replacing them. The falsification, in any case, remains arbitrary, a matter of taste. Duhem claimed to protect a theory deserving the qualities of simplicity and beauty, and to abandon it when it loses them. Popper, on the other hand, also recognized the conventional character of all propositions and falsification, but wanted to return to experiments their power also against well established theories. In his viewpoint, even if all choices are conventional, we can make some singular statements unfalsifiable if "there exists at the time a 'relevant technique' such that 'anyone who as learned it' will be able to *decide* that the statement is 'acceptable' " [47].
The methodological falsificationist is aware that measurements involves fallible theories, but he *conventionally* decides to relegate them to *'unproblematic background knowledge'*, he *"uses our most successful theories as extensions of our senses"* [33]. The conventional character of this methodology is located in the assignment of the status of 'observational theories' to the one used to perform and interpret experiments, seen as uncriticized background knowledge. An additional element of convention is in the truth-value assigned to experimental results. Being any proposition not undoubtedly provable, the better we can do is to repeat an experiment - that is a potential falsifier - a conventional number of times in order to minimize the risk of errors, so that we can consider it quite reliable. In this way we can establish a sort of 'empirical basis', not meaning some provable set of propositions (proposition can never be proved) but a set of potential falsifiers that when *conventionally* considered in contradiction with a theory they can 'falsify' it, and consequently eliminate it *conclusively.* Methodological falsificationists are well aware that they risk to eliminate a theory even if it is 'true', but this is the risk one must assume in order to make science proceeding. Scientific methodology must be severe in order to save only the best theories and reject the 'falsified' one. The awareness of the fallible character of this elimination is a major property of this methodology. Methodological

falsificationists consider experiments fundamental, because they play the crucial role of guides in theory building, avoiding scientists to lose the right track in the process.
They consider a theory as 'scientific' only if it contains an 'empirical basis'. The difference from dogmatic falsificationism is in the inverted commas: as we have already argued, the 'empirical basis' is a conventional set of potential falsifiers tested by a conventionally chosen observational theory. This difference make methodological falsificationism much more liberal and flexible than the dogmatic one, because there is no provable empirical basis, so the falsification and consequent elimination is a matter of convention.
In any case, we have already seen that we cannot falsify any theory without a ceteris paribus clause, and it remains true also in this latter context. In order to allow elimination of theories, methodological falsificationists states that when we test a theory with a ceteris paribus clause and we find disagreement, we can decide to reject the former or investigate the latter. This investigation can be carried on by assuming the existence of other factors that have to be specified and tested. Then, if they are not verified, the ceteris paribus clause can be considered as reliable.
This argument allows methodological falsificationism to recognize Newton's theory as scientific, because it added the possibility to investigate the ceteris paribus clause instead of rejecting a theory when faced with counterevidence. In this sense, this brand of conventionalism can account for a wider rational reconstruction of the opposite one, so it can be considered a step forward.
Another important point, in the methodological falsificationist's viewpoint, is that falsification is necessarily in order to control scientific honesty. If we deny the possibility of eliminating theories, that is the conclusive character of falsification, scientists may stick to their theories and claim that 'the discrepancies which are asserted to exist between the experimental results and the theory are only apparent and that they will disappear with the advance of our understanding' [47].
Another useful observation concerns theories that do not admit potential falsifiers because of their logical form, so that we cannot falsify them in the ways discussed above. These theories can be falsified when they clashes with another well established theory, which we can relegate to the unproblematic background knowledge. This is not so different from the observational theories' case, and it is coherent with the above line of reasoning.
Summarizing, methodological falsificationism allows to reconcile scientific criticism and fallibilism thanks to its conventional but strict character. A major point is the the concept of scientific honesty: a scientist must specify, an advance, an experiment which could contradict the theory, so that he must abandon the theory when it is verified. Also, we must remember that this methodology is a brand of conventionalism: decisions play a crucial role, especially in the elimination of theories. Methodological falsificationists think that we must have a pragmatic method allowing us to immediately decide if we should retain or reject a theory, and that the rejection must be conclusive when conventional agreement on falsification is reached. In this sense, methodological falsificationists

think that in those cases a scientist is faced with a forced choice: for the wellness of science, he must eliminate the theory, even if it may be still true. The only alternative, in their viewpoint, is irrationalism.
It is quite evident that this methodology, even if quite liberal in accounting for past scientific successes, has a too narrow normative character, which is contemporary too arbitrary in the crucial decisions. In this sense, it not only fails to describe the real process of science, but it is also impracticable.
Giving a deeper look at history of science, we realise that there are at least three discrepancies between the discussed methodology and the rational of scientific process.
First of all, there have been times when theoreticians challenged experimental results and demonstrated that observations were wrong[2]. This kind of event is forbidden by the above methodology, because once an experiment is shown to contradict a theory, this one must be rejected. Methodological falsificatonism does not allow such appeals.
A second discrepancy can be found in the idea that scientific process consists of a challenge between one theory and experiments only. This is an excessive simplification, because we are neglecting the importance of a third factor: the presence of rival theories.
Finally, methodological falsificationists think that "the only interesting outcome of such confrontation is (conclusive) falsification: 'discoveries are refutations of scientific hypothesis' " [1]. Lakatos thought instead that the real interesting outcome is confirmation rather than falsification. Scientists do not care of the impossibility to prove statements, they see at corroborations as real proofs. Furthermore, falsification has never had a crucial role in history of science, as methodological falsificationists claim.
Considering the situation, we are left with two possibilities: to abandon the goal of giving a rational explanation of the success of science, following sceptics viewpoint, or to modify this naive form of methodological falsificationism proposing a different method to eliminate theories, that is a different and less conventional concept of falsification. This new version is called *sophisticated methodological falsificationism.*

1.1.4 Sophisticated methodological falsificationism

The main difference between naive methodological falsificationism and the sophisticated one is that while the former considers 'scientific' or 'acceptable' only those theories able to propose in advance an experiment that could falsify it, the latter claims that "a theory is 'acceptable' or 'scientific' only if it has corroborated excess empirical content over its predecessor (or rival), that is, only it leads to the discovery of novel facts. This condition can be analyzed into two clauses: that the new theory has excess empirical content

[2]For example, when the royal astronomer Flamsteed presented observational data in disagreement with Newton's theory, Newton provided him with a better theory of the atmosphere's refractive power. This correction brought to new empirical data in agreement with Newton's theory.

(*acceptability*$_1$ and that some of this excess empirical content is verified (*acceptability*$_2$. The first clause can be checked instantly by *a priori* logical analysis; the second can be checked only empirically and *this may take an indefinite time* "[33]. In addition to this redefinition of acceptability, the sophisticated brand also modifies the falsification condition, imposing three different properties that a new theory should have in order to 'falsify' the old one. A theory can be falsified when a new theory: 1) has excess empirical content respect to the old one, that is predicts novel facts; 2) contains all the empirical content of the old theory; 3) some of its excess content - that at least one of the novel facts it predicts - is experimentally verified.
The necessity of a new methodology was due mainly to the fact that in the conventionalists' viewpoint every theory can be saved through some auxiliary hypothesis, so we would like to recognize scientific auxiliary hypothesis from ad-hoc adjustments, that is we want to recognize real science from pseudo-science.
This new kind of methodology is able to discern between ad-hoc theoretical adjustments and scientific auxiliary hypothesis, being this difference implicit in its redefinitions of 'scientific' and 'falsification' : we can say that an auxiliary hypothesis is not ad-hoc when it lead to scientific progress, that is when it lead to an excess of empirical content. This excess is a relational concept, because we can see a progressive theory as a series of theories with new auxiliary hypothesis and new empirical content at each step; we consider not isolated theories, but series of theories
We can rework the above discussion introducing some fundamental concepts.
I quote Lakatos' words, coincisely explaining these crucial concepts:

> let us take take a series of theories , $T_1, T_2, T_3, ...$ where each subsequent theory results from adding auxiliary clauses to (or from semantical reinterpretation of) the previous theory in order to accomodate some anomaly, each theory having at least as much content as the unrefuted content of its predecessor. Let us say that such a series of theories is *theoretically progressive (or "constitute a theoretically progressive problemshift")* if each new theory has some excess empirical content over its predecessor, that is, if it predicts some novel, hitherto unexpected fact. Let us say a theoretically progressive series of theories is also *empirically progressive (or "constitutes an empirically progressive problemshift")* if some of this excess empirical content is also corroborated, that is, if each new theory leads us to the actual discovery of some *new fact*. Finally, let us call a problemshift *progressive* if it is both theoretically and empirically progressive, and *degenerating* if it is not. We *"accept"* problemshifts as "scientific" only if they are at least theoretically progressive; if they are not, we *"reject"* them as "pseudoscientific. Progress is measured by the degree to which a problemshift is progressive, by the degree to which the series of theories leads us to the discovery of novel facts. We regard a theory in the series "falsified" when it is superseded by a theory with higher corroborated content ([33], p.33-34).

We can recognize four fundamental new concepts. First of all, we cannot appraise a single scientific theory, but we can call 'scientific' or 'unscientific' only a series of theories. Second, we have a new kind of 'falsification': a theory is falsified when it is superseded

by a better theory. In this sense 'falsification' can be made only if there exists a better theory. This new concept changes the meaning of counterevidence: there cannot be counterevidence able to reject a theory; we can redefine a counterevidence as an experiment which contemporary corroborates a new theory. We can get a new look on crucial experiments: they are experiments not only falsifying one theory, but mainly corroborating a new one. If the new theory also contains the other theory's empirical content, it can supersed it.
Third, we can finally return to experiments their power, because of the concept of empirical progress. Conventionalists thought we can always accomodate any anomaly adding auxiliary hypothesis, decreasing the power of experiments. Now, having underlined the difference between a progressive and a degenerating problemshift, we can recognize when an auxiliary hypothesis is scientific or ad hoc, this thanks to the predicted and corroborated novel facts, that is thanks to experiments. The excess of corroborated content plays a crucial role, and only the anticipated evidence is considered relevant, 'true' evidence. In this perspective theoretical progress and empirical evidence are strictly connected, but in new ways.
Four, scientific process is finally understood as a relation between more theories and experiments, and not only as one theory only facing experiments. Furthermore, the above concepts underline a competitive mechanism between theories, shedding light on the relational character of scientific research. That shows the constructive and historical character of sophisticated falsificationism: we analyze competing series of theories, and there is no necessity to reject a theory because of a single counterevidence.
It is fundamental to understand the elements we now look at when analyzing a theory: for the acceptance of a new theory we are interested in its excess of empirical content respect to another theory, that is in progressive theoretical and empirical problemshift; for the falsification of a theory, we are interested in its degeneration, that is in its being superseded by a new progressive theory. These are complementar aspects: we reject a theory only if we have a better one substituting it. This is a main point because it encourages proliferation of theories in any moment, while naive falsificationism claimed for new theories only when a theory was rejected because of counterevidece. Naive falsificationists needed new theories because of their strict falsificational method, while sophisticated falsificationists ask for new theories at each moment, without needing them because of a deficiency.
Another concept needing to be underlined is that a theory can live 'in an ocean of anomalies' without being refuted. This is fundamental because it is historically accurate. We will see that in the history of science also the most successful theories developed among anomalies without being refuted, meaning that falsification in the naive sense never occurred in history. This methodology gives us a more rational explanation of scientific process than the naive falsificationists', even if it needs to be implemented by Lakatos methodology of scientific research programmes, as we are going to see in the next subsection. Additionally, while the naive methodology suggested an untenable normative

methodology, this sophisticated version appears as a rational and viable methodology.
This methodology also suggests a new conception of *scientific honesty*: scientists should propose new theories predicting some novel facts and should reject only those theories which appears to be superseded by others with much more empirical content.
This sophisticated version also has another virtue respect to the naive one: it reduces the conventional element. The naive version was characterized by a risky conventional decision: when faced with a counterevidence, one must decide which part of the theory to replace (for example, we could decide to reject the theory or to reject the ceteris paribus clause and investigate it further). Now we have no such decision to make: we can try to replace any part of the theory until we reach an explanation of the anomaly able to increase the empirical content, that is not only accounting for the anomaly but also predicting unexpected novel facts. In this way we not only reduce the conventional character, but also avoid the risky decisions characterising naive falsificationism.
Another important virtue of this methodology is that it also explain how we can decide to retain or reject a metaphysical theory. The line of reasoning is always the same: we retain it if we can add or change auxiliary hypothesis explaining the anomaly and resulting in an increase of empirical content. In this case it is a progressive metaphysical theory, otherwise it is degenerating. Naive falsificationists thought that we must eliminate a metaphysical theory when it clashes with a well established theory. Now, we are giving a much different criterion: we reject it if it is degenerating *and* if there is a better theory replacing it. The virtue of this argument is that the methodology of a scientific research programme based on some metaphysical statement - that is a statement for which we cannot define any empirical basis - does not differ substantially from the methodology of a scientific research programme with a falsifiable core. This is a great achievement because we have a general methodology allowing us to appraise any kind of research programme in a very pragmatic way, without the necessity of differentiate them from strictly falsifiable and metaphysical, but only looking at their ability to predict novel facts. The scientific character of any theory is not in its own characteristics, but in its efficiency, in its ability to achieve scientific success. This is at the same time a very flexible and very pragmatic methodology, able to reduce conventional and risky decisions too.
Anyway, we can still identify the presence of some conventional elements.
As we have already seen, there is no sharp separation between theoretical and observational statements, so the assignment of this quality is a matter of convention. We construct experimental devices and interpretative models based on well established theories in order to test new theories, but we could also use the new theory in order to test the observational one. We proceed in the first way because we cannot avoid using theories to do experiments, we need to define some background knowledge to perform them, so we consider better to use well-established theories to test new hypothesis instead of believing on the truth-value of a new hypothesis and test the observational theory. In any case, a falsifying hypothesis can be considered as problematic as a falsified one.
Moreover, after having decided which is the observational theory, we must decide the

truth-value of its results, that is if they are reliable or not. This is a major point because a wrong appraisal may define a problemshift as a degenerative one while in reality is is progressive, and viceversa. This case would be not as dramatic consequences as in the naive method, because we would not reject a theory immediately facing a counterevidence, but anyway it is an odd scenario.
In order to reduce this possibility, sophisticated falsificationists - contrary to naive falsificationists - allows an *appealing procedure.* This procedure allows a theoretician to question the verdict of a negative experiment. The experimenter, when performing an experiment, implicitly interpret results in the light of a certain theory or deduce the consequences from some theoretical assumption. If the observational theory is left implicit, the clash between an hypothesis and experimental results can be seen as a consequence of a *monotheoretical deductive model*, where the advanced hypothesis is under test and the observational theory is relegated to background accepted knowledge. The appeal procedure instead reveals the pluralistic character of the empirical test, making explicit the presence of an interpretative theory and putting the two theories on the same level. In this procedure the theoretician asks to the experimenter to specify this theory and makes it objectionable as its own hypothesis. This *pluralistic model* shows again that facts cannot reject a theory, because the clash between theoretical statements and experimental results represents actually an *inconsistency* between different theories. Concisely, Lakatos claims that "it is not that we propose a theory and Nature may shout NO; rather, we propose a maze of theories, and Nature may shout INCONSISTENT"(ibid.).
Now the question is: which one of the two inconsistent theories one should replace? In the framework of sophisticated falsificationism, the answer is obvious: one should try to replace first one theory and then the other one, and decide to retain the theory that gives the most excess in empirical content. Following this procedure, scientists should be able to avoid wrong decisions when facing a negative experiment. A conventional element remains when deciding which one of the theories constitutes a progressive problemshift, but we cannot pretend to remove all decisional elements in the scientific process. This only thing we can due is to reduce as much as possible this element in order to avoid big mistakes and and follow the 'safest' path. Also, a decision must be made when the theoretician uses the procedure appeal but no alternative theories are found to replace the new hypothesis or the observational theory, in order to account for the excess in empirical content. Finding a new observational theory may be difficult because of a long tradition of a certain kind of experiments and because it is a well established and successful theory. Anyway, in these cases one must take a temporary decision, for example deciding to temporarily account for the degenerative character of the proposed theory.
This process might seem to slow down the scientific process because of the complex relation between competitive theories and experiments, but it is very useful in avoiding mistakes. We must underline that it also left some decisions to common sense, that is also in the appeal procedure there is a conventional element: as well as we cannot include all theories in the background knowledge, in the same way we cannot claim to criticize

all of them.
Finally, in order to introduce Lakatos methodology of scientific research programmes, we appoint the so called 'tacking paradox'. We claimed that any auxiliary hypothesis able to increase the empirical content can represent a progressive problemshift. We did not constrain such hypothesis to be connected to theories in any way, so their possible disconnection from theories might make it difficult to eliminate them. The solution is to demand that such auxiliary hypothesis must belong more intimately to the original theory so that we can see some sort of continuity in this process, that is we do not see history of science merely as a series of theories with additional hypothesis able to make a progressive problemshift, but as research programmes. This difference may seem to be a straightforward linguistic solution but it has important methodological consequences as we are going to see in the next subsection.

1.2 Lakatos' methodology of scientific research programmes

The series of theories we have already discussed actually must be view from another point of view, that is they can be understood as scientific research programme. We can recognize different common elements in the series of theories belonging to a research programme, among which are two methodological concepts representing a guide in their development: the *negative heuristic* and *positive heuristic.*
The *negative heuristic* represents the methodological choice of defending the *hard core* of the programme. The consequence is that when scientists face a counterevidence they do not redirect the *modus tollens* [3] to the hard core, but instead "invent auxiliary hypothesis, which form a *protective belt* around this core, and we must redirect the modus tollens to these"(ibid.). Anomalies are then directed to these hypothesis, that we can modify or replace in order to account for experimental results. Obviously, if this modification or replacement increase the empirical content, that is if it predicts novel unexpected facts, it represents a progressive problemshift, otherwise it is a degenerating one. Notice that this is a methodological decision: scientists choose not to direct counterevidence to the hard core, which is considered 'a priori' irrefutable and unchangeable. We must underline that at each step in this process we require an increase in the empirical content, that is a theoretical progressive problemshift, but we cannot require immediately a progressive empirical problemshift as well. Corroboration may take long time, so we ask only for a 'consistently progressive theoretical problemshift' and 'intermittently progressive empirical shift'. We recognize that in order to avoid ad-hoc hypothesis built up to save

[3] if $A \Rightarrow B$ then $notB \Rightarrow notA$

the theory, we must require a theoretical constant progress, but we also recognize that we cannot require that new phenomena should be observable soon after their prediction; after all, experiments are constrained by historically factors like the technological possibilities, and we must account for it. We just require that when possible, that is when we are able to make an experiment, the theoretical content increasing hypothesis must be corroborated. Summing up, the negative heuristic allows the scientist to redirect refutations to the protective belt and to go on with the programme as long as it constitutes a progressive programme. When it ceases to produce an excess of empirical content, and starts to lag behind the facts, the scientist should decide to abandon its hardcore.
The concept of *positive heuristic* is likewise fundamental. It explains how the building process of scientific research programmes works. One may think scientists construct new theories trying to explain directly a certain anomaly and invent random auxiliary hypothesis until they finally explain the fact, but it is not the case. They follow a methodological plan, that is they have a long-term policy basically indifferent to known anomalies. Their aim is to start from a certain hardcore and step by step construct the research programme; the hope is that at a certain moment of this process they will be able to explain the anomalies. This research policy represents the *positive heuristic*.
Briefly, the negative heuristic does not allow the modus tollens to be directed against the hard core, while the positive heuristic represents a development project where the subsequent steps consist of increasingly refined and realistic models.
The scientist needs to follow a positive heuristic in order to have a guide when trying to construct a theory in an 'ocean of anomalies', so that its strategy is not to choose a random anomaly and trying to construct a theory to explain it, but rather to ignore actual anomalies and experimental data and follow the heuristic. His hope is that while proceeding in the development of the programme, he will be able to explain them, so that anomalies become corroborations of the programme. In this process the theoretician may construct a model contradicting contemporary observations, but he is well aware that it is a temporary and imperfect model, so he continues on his way. Importantly, it may also happen that when a research programme is undergoing a degenerative phase, a "little revolution or a creative shift in its positive heuristic may push it forward again"(ibid.).
Now we can see that while the naive falsificationist thought that only falsification represents a meaningful scientific discovery, we showed that it is verification of novel facts that makes a research programme going on, and this without caring of anomalies. This methodology return importance to verification and remove it from falsification, giving a much better and realistic account of scientific process.
A research programme, in the light of the above concepts, can be appraised on the grounds of its *heuristic power*, that means on its power of explaining anomalies and the amount of novel and corroborated facts it predicted in the process.
All the above discussion explains the *relative autonomy of theoretical science*, because scientists do not care about anomalies in the early stage of their research programme, but instead develop the positive heuristic regardless of any counterevidence. This con-

cept, unexplained by previous philosophies, is fundamental in the scientific process. If science really followed naive falsificationists' methodology, scientific success could never be reached; every time an anomaly occurs, scientists should reject their theory, but history of science teaches that *every* research programme is born falsified, so we the theory construction would be impossible. The relative autonomy of theoretical science is fundamental in allowing scientists to develop their theories without caring of anomalies at first; research programmes may take a long time to reach their final form, and before this moment they rightly may not be able to explain anomalies. This observation has a counterpart in the effectively smaller importance of experiments respect to the one usually accounted in historical accounts. We must remember that any scientific reconstruction is philosophically based, so that it cannot be impartial. Many times we find experiments referred to as 'crucial' experiments, but a better look at history of science shows that these are historical forgeries. As Lakatos points out:

> One of the most important points one learns from studying research programmes is that relatively few experiments are really important. The heuristic guidance the theoretical physicists receives from tests and 'refutations' is usually so trivial that large-scale testing - or even bothering too much with the data already available - may well be a waste of time [4]. In most cases we need no refutations to tell us that the theory is in urgent need of replacement [5]: the positive heuristic of the programme drives us forward anyway. Also, to give a stern 'refutable interpretation' to a fledgling version of a programme is dangerous methodological cruelty. The first version of a programme may even 'apply' only to non-existing 'ideal' cases; *it may take decades of theoretical work to arrive at the first novel fact and still more time to arrive at interestingly testable versions of the research programmes*[6], at the stage when refutations are no longer foreseeable in the light of the programme itself ([33], p.65).

This is a crucial point, because the relation between theory and experiment is not the trivial one often presented in textbooks, but reality is much more complex.
In the light of the above discussion, it is even more clear the importance of competition between research programmes. History of science can be seen as a proliferation of research programmes superseding each other. This replacement occurs when a research programme shows to achieve more empirical content than its competitors, that is when it is able to account for all the facts explained by other theories and also pass observational tests of novelty predicted facts. This process may take long time, both because

[4]an example could be the discovery of the Higgs boson in 2012. This particle was fundamental in the framework of the Standard Model, but the whole research programme was completed and considered a successful one much before this discovery. This has been possible because of its great empirical power, that made scientists sure that the Higgs boson would be found (there were some exceptions, such as Stephen Hawking). Even before the discovery, scientists were sure about its existence; the real surprise would have been to not discover it.

[5]for example Standard Model still continues to be confirmed by experiments, but we search for a new theory because of its limitations.

[6]my italics

the research programmes may become able to predict new facts after a long period, a period necessary for the development in the light of the positive heuristic, and because of limitation of contemporary experimental technologies. In any case, it is clear that being scientific history an history of competing programmes, 'theoretical pluralism' is a fundamental factor for its progress. As a consequence, the construction of new theories should not wait for the degenerating phase of a dominant research programme, but it should always be carried on. Proliferation of theories is demanded to strengthen the competition. This discussion also shows that even if we can talk with hindsight of a replacement of some research programme by another one with more heuristic power, more research programmes may coexist in the process. We are not forced to replace a programme by a better one if it still continues to show a progressive character. This is a clear consequence of this methodology, because a research programme may show its potential after a long time, and it may demonstrate itself to be better than a competing research programme previously considered more powerful. Time, in this sense, is a crucial factor. The scientific process cannot be appraised at any instant, because we cannot appraise theories but research programmes, that usually follow a long-time policy in order to reach their goals.

Historiographical accounts often shows a fake succession of events, such as crucial experiments instantly rejecting a research programme. This, of course, is a false narration. An experiment gains the appellation of 'crucial' only with hindsight, after a period when the anomaly it discovered allowed a research programme to supersed another programme. Often, the process is even more complex than this, so that very few experiments deserves to be called 'crucial'. The role of time in the appraisal of the scientific process shows that the idea of instant rationality is utopian. In this framework a research programme cannot be appraised straightforwardly at any instant. Obviously, we must evaluate it in order to carry on the scientific process, but this evaluation can be only temporary, and we must be aware that it may change with time. This is one of the conventional elements we cannot get rid of. Instant rationality is not able to account for all the events in history of science.

This process, embedded in time as all processes, allows to give two the crucial definitions. As Lakatos points out: "my account implies a new demarcation criterion between *mature science*, consisting of research programmes, and *immature science* consisting of a mere patched up pattern of trial and error. For instance, we may have a conjecture, have it refuted and then rescued by an auxiliary hypothesis which is not *ad hoc* in the senses which we had earlier discussed. It may predict novel facts some of which may even be corroborated. Yet one may achieve such 'progress' with a patched up, arbitrary series of disconnected theories. Good scientists will not find such makeshift progress satisfactory; they may even reject it as not genuinely scientific. They will call such auxiliary hypothesis merely 'formal', 'arbitrary', 'empirical', 'semi-empirical', or even *'ad hoc'*. Mature science consists of research programmes in which not only novel facts but, in an important sense, also novel auxiliary theories, are anticipated; mature science - unlike

pedestrian trial-and-error - has 'heuristic power'. Let us remember that in the positive heuristic of a powerful programme there is, right at the start, a general outline of how to build the protective belts; this heuristic power generates *the autonomy of theoretical science*"([33], p.87-88). The difference is in the way the programme develops, in the way progressive problemshifts are achieved. A research programme might make a progressive problemshift by using an arbitrary series of disconnected auxiliary hypothesis, and in this case we call it 'immature'. A mature research programme, otherwise, follows the positive heuristic, so that scientists know the auxiliary hypothesis they will add in the evolution of their models.
These concepts are of major importance to differentiate between real research programmes, guided by a long-term policy, and series of theories that, even if able to achieve progressive problemshifts, do not constitute research programmes in the very sense. I will remind these important concepts in chapter 3, where I will analyze the string theory research programme from a lakatosian perspective.

Two other concepts remain to conclude this brief review of Lakatos methodology of scientific research programmes. One is the historiographical methodology of scientific research programmes, linking philosophy of science to history of science, and the other one is the Zahar's extension of Lakatos' methodology.

Elie Zahar was a student of Lakatos, and they wrote together a paper in 1972-1973, then published in 1976 titled "Why did Copernicus's research programme supersede Ptolemy's?" ([33], p.168). Here I am not going to discuss the whole paper, but I only analyze the extension of the methodology of scientific research programmes proposed by Zahar. He only added a simple but important concept in order to explain why Copernicus won the battle before his theory was corroborated by experiments. The first verification of his theory was achieved after about seventy years the theory was proposed, but his supporters where already convinced of its superiority respect to Ptolemy's system. Why were they persuaded? Zahar identifies a new class of facts as novel facts, facts that could not be considered 'novel' in the Lakatos viewpoint. He claims that we can consider as novel facts also those already known facts that were not expected as a consequence of the initial hypothesis. For example, General Relativity was not constructed in order to explain the anomalous Mercury's perihelion, but to conclude the project started with Special Relativity, that is to demonstrate the equivalence of all observers. Einstein's first aim was not to solve this anomaly but in any case, thanks to Schwarzschild, it could be unexpectedly explained. This new concept explains why Copernicus' theory was considered better than Ptolemy's. Copernicus' assumption has many consequences for the description of planetary motions. This consequences where already known facts, but they were explained by Ptolemy's system using ad-hoc parameter adjustments. On the other hand, the same facts could be explained by Copernicus using one simple assumption, that was not proposed for the very aim.

As I have already discussed, philosophy and history of science are closely related. History of science aims to reconstruct the scientific progress, and this reconstruction is always philosophically based because of the impossibility of an objective narration. Historians choose the events they consider of major importance, and narrate them through the lenses of their own historiographical methodology. The methodology of scientific research programmes represents an accurate instrument for this reconstruction, being able to describe the scientific process as a rational series of events. Thus, it is natural to use this methodology in order to understand history of science. Having already discussed this methodology, we only need to add two other concepts: *internal history* and *external history*.

In a paper published in 1971, Lakatos argued that "(a) philosophy of science provides normative methodologies in terms of which the historian reconstructs "internal history" and thereby provides a rational explanation of the growth of objective knowledge; (b) two competing methodologies can be evaluated with the help of (normatively interpreted) history; (c) any rational reconstruction of history need to be supplemented by an empirical (socio-psychological) 'external history' " [32].

A deep analysis of the arguments sustaining these thesis would take us too far, so we limit ourselves to underline the main points.

First of all, in order to appraise a historiographical methodology, one should use a philosophical methodology. It means that, if Lakatos' methodology states that a research programme can supersed another one if the former shows an excess of empirical content respect to the latter, this appraisal should be made also for the historical reconstruction that the methodology provides to the historian, that is, the new narration must provide a rational explanation of a larger number of historical facts respect to competing methodologies, and should also predict novel historical facts. Now, having already analyzed the main methodologies in modern philosophy of science, one should be able to recognize that this is the case for an historiographical research programme based on Lakatos' methodology. We will not discuss it further, so we refer back to the quoted paper.

In order to give a brief explanation of points (a) and (c), we need to define 'internal' and 'external' history. 'Internal history' represents the rational reconstruction of historical events, which is, as we have already told, philosophically based. The historian will take from history the facts *he* thinks to be important in the view of his methodology, and explains them. The more facts a methodology is able to reconstruct as 'rational', the better will be the reconstruction. Obviously, scientists are human beings, scientific process cannot be completely rational. Humans are subjected to psychological bias, politics, false beliefs, culture and many other factors. They influence scientific research, and sometimes the rational path of science can be modified or deviated, showing exceptions. 'External history' aims to account for these factors, and for this reason is often called 'sociological' or 'empirical' history.

These two elements, together, should be able to give an explanation of history of sci-

ence. The demarcation between internal and external history is philosophically based: depending on the methodology one adopts, one can see an historical fact as belonging to internal history and another one belonging to external history. External history tries to explain historical events that appears irrational in the light of the adopted methodology so, as we have already claimed, the richer is internal history respect to the external one, the better is the rational reconstruction.
One criticism may come as a natural question about all we have claimed until now. We saw that time is an important factor is the evaluation of scientific research programme, so the natural question is: how much time should we expect in order to appraise a research programme as a progressive or degenerating one?
This question is implicit in the critics Kuhn moved to Lakatos for not giving a "criteria which can be used *at the time* to distinguish a degenerative from a progressive research programme"[31], and claimed that his methodology would be useless without such criteria, only *'verbal ornaments'*, as Feyerabend pointed out. A similar critics was moved by Musgrave, asking "at what point dogmatic adherence to a programme ought to be explained 'externally' rather than 'internally' "[33].
The answer given by Lakatos explains very well the logical fallacy this critics present: "One may rationally stick to a degenerating programme until it is overtaken by a rival *and even after.* What one must *not* to do is to deny its poor public record. Both Feyerabend and Kuhn conflate *methodological* appraisal of a programme with firm *heuristic* advice about what to do. It is perfectly rational to play a risky game: what is irrational is to deceive oneself about the risk"(ibid.).
This point is crucial in order to understand the role of philosophy of science. Philosophers do not give orders to scientists about what to do, but they try to appraise research programmes in the light of their methodology, to discern scientific theories from pseudo-scientific theories, to identify rational and irrational problemshifts and so on. They cannot oblige scientists to abandon a research programme because it is degenerating, and likewise they do not settle a time limit until which scientists can stick to a research programme. They only evaluate it from a different point of view compared to that of scientists, so that the scientific community can be more conscious of the scientific process. We think here it is the importance of such a work, in the fact that being scientists human beings they can act irrationally, so philosophy of science can shed lights on their work and indirectly supervise it; 'indirectly', we may say, because methodological appraisal is made with hindsight, so it is the whole rational reconstruction of history of science that gives advice to scientists, explaining why science is so successful.
This fact has important consequences in the process of science. Today theoretical physics' community has been divided by string theory, and string theorists too are in conflict among them. A rational reconstruction of string theory research programme may clarify the situation and help to solve contemporary debates, and maybe suggest how to proceed. In particular, philosophy of science can provide a criterion for 'scientific honesty', in the light of which there are no 'allowed' and 'forbidden' actions in the 'game of

science', but at most 'honest' and 'dishonest' moves. This difference is of importance, because some philosophers may try to impose some normative and a priori rules to the game of science. Popper, for example, is one of them, thinking that it was necessary to impose an *immutable statute law*, that is *a priori general rules* for scientific appraisal. Looking at his demarcation criterion, pretending to give a general rule to distinguish science from pseudo-science, this appears quite evident. The idea of the imposition of this 'book of rules' is quite ingenuous, because history teaches us that scientists have never had it, but anyway they have been able to achieve scientific progress. There is no 'statute law', and in fact philosophies based on this idea never succeeded to give a precise rational reconstruction of history of science.
We have arrived at a key moment, so I would like to quote Lakatos in order to be clear:

> Up to the present day it has been the scientific standards,as applied 'instinctively' by the scientific *élite* in *particular* cases, which have constituted the main - although not the exclusive - yardstick of the philosopher's *universal* law. But if so, methodological progress, at least as far as the most advanced sciences are concerned, still lags behind common scientific wisdom. Is it not then *hubris* to try to impose some *a priori* philosophy of science on the most advanced sciences? [...] I think its is. And indeed, the methodology of scientific research programmes implies a pluralistic system of authority, partly because the wisdom of the scientific jury and its case law has not been, and cannot be, fully articulated by the philosopher's statute law, and partly because the philosopher's statute law may occasionally be right when the scientists' judgment fails. I disagree, therefore, both with those philosophers of science who have taken it for granted that general scientific standards are immutable and reason can recognize them *a priori*, and with those who have thought that the light of reason illuminates only particular cases. [...] But this two-traffic need not always be balanced. The statute law approach should become much more important when a tradition degenerates or a new bad tradition is founded. In such cases statute law may thwart the authority of the corrupted case law, and slow down or even reverse the process of degeneration. When a scientific school degenerates into pseudoscience, it may be worthwhile to force a methodological debate in the hope that working scientists will learn more from it than philosophers' ([33], p.137).

As we will see, string theory represents a divisive research programme. In this situation, a philosophically based discussion can represent an instrument to be used in order to clear and perhaps solve the discussion. Scientists' judgement is fallible, so their case law might get wrong, blinded by psychological bias or other factors, and in this situations the philosophers' statute law may be used to get them back on track. I think, as Lakatos does, that "a good methodology - distilled from the mature sciences- - may play an important role for immature and, indeed, dubious disciplines" (ibid.).

The role of philosophy of science in contemporary theoretical physics will be even more clear in the chapter 4, where I will analyze Richard Dawid's philosophy.

Chapter 2

History of string theory

Discussing Lakatos' methodology of scientific research programmes we recognized the fundamental relation between philosophy and history of science. Philosophy gives to historiography the tools to select and rationally explain historical events, and history of science gives to philosophy the material to be analyzed in order to understand the scientific process and create such tools. Furthermore, Lakatos' methodology in particular gives an important role to the details of a research program's development. Being a research program evaluated using concepts like the request for the prediction and corroboration of novel unexpected facts, the competition among research programs and the difference between progressive and degenerative problem-shifts, that is concepts representing historical facts, Lakatos' methodology implies that a research program cannot be evaluated without analyzing it as a process embedded in time. In this chapter I am going to present a brief review of the history and main concepts of string theory. To this end I will refer for the most[1] to the book 'A brief history of string theory, from dual models to M-theory' by Dean Rickles [49]. This is a 'teleological' reconstruction, meaning that mostly those developments that led to the current situation in string theory are discussed, leaving little room for those facts that are superfluous in today's perspective. Many important facts that do not strictly relate to the aims of the narrative have been left out, so I warn the reader that it is only a partial reconstruction of a specific history. Both because an analysis of Rickles' reconstruction would take us too far and because it is certainly sufficient for the purposes of this work, I will take for granted his account and try to give (for the moment) an objective presentation of the main facts.
Then, in the following chapters, I will analyze string theory using Lakatos' and Dawid's methodologies.

[1] I strictly followed his reconstruction, deviating only at the end when talking about the string Landscape and the Swampland program.

Rickles divides string theory's history into four periods, each of them starting with the resolution of a technical problem reviving hope in the research programme. Following Rickles, I give a schematic anticipation of the four periods we are going to explore:

- Phase 1 [1968-1973]:

  - Phase 1A (Exploring Dual Models) [1968-1969]: Gabriele Veneziano found a formula describing hadronic scattering after which many scientists start working on dual models. This formula is generalized and technical issues are solved; among them, one is the elimination of ghosts which, however, introduces a tachyon. Virasoro found a dual model analogue to Veneziano's. Both an abstract operator approach and a stringy geometrical approach are introduced; however, the latter is considered to be only a useful model, not describing anything real. Also, a Feynman diagram approach is introduced. Until now only bosonic models are considered.
  - Phase 1B (Embryonic String Theory) [1970-1973]: Physicists found that dual models are consistent only in a space-time with 26 dimensions. Fermions are introduced, and the number of dimensions reduces to 10. An action based on the minimization of the worldsheet area is constructed. The 'no ghost theorem' is proved and a consistent relativistic and quantum theory of a free and massless string is constructed and demonstrated to reproduce dual model's physics. In this phase trust in hadronic string theory decreases because of many unsolved problems and because of the rise of QCD. Nevertheless, the low energy limit of hadronic string theory is performed, showing to reproduce Yang-Mills gauge theory and General Relativity. Until now, string theory was only a theory of hadrons.

- Phase 2 (Theoretical Exaptation) [1974-1983]:
  String theory is rescaled to describe quantum gravity rather than hadronic physics; hadronic string theory research continues and is integrated with QCD. Interacting strings are studied. Spacetime supersymmetry is introduced and shown to finally remove the tachyon from the physical spectrum. A chiral anomaly is found. Physicists try to compactify the extra-dimensions and reformulate superstring theory. In this period supergravity is introduced and vigorously pursued.

- Phase 3 (Superstring Phenomenology) [1984-1994]:
  The chiral anomaly is resolved by using a gauge group which is also phenomenologically promising. For this reason, many scientists start to consider superstring theory really a possible unifying theory. The same hopes are raised by the discovery of the heterotic string theory and compactification on Calabi-Yau manifolds. This last one, anyway, finally implies the loss of uniqueness. The (re)discovery of D-branes and the duality conjectures radically change the nature of string theory.

- Phase 4 (Beyond Strings) [1995-presents]:
string theory is discovered not to be only a theory of strings, but instead scientists found that the five string theories they had contain many objects with different dimensionality called Dp-branes. Furthermore, dualities are shown to link different string theories and they are all seen as different limits of a deeper unknown theory, called M-theory. The Bekenstein-Hawking formula is reproduced and a possible resolution of the information paradox is proposed. Maldacena presents a conjecture named Ads/CFT correspondence, opening important new lines of research. Instability of Calabi-Yau manifolds is found and resolved using D-branes, but as a consequence string theory now allows for a huge number of possible ground states, leading to the concept of 'Landscape' and an increasing important role of the anthropic principle. These last two concepts represent the main topics around which the debate between string theorists and their opponents is constructed.

Before going ahead with the first phase I give a brief account of the status of particle physics research before the discovery of Veneziano's formula. Particle accelerators were discovering a large number of new particles that theoreticians were not able to explain in the quantum field theory framework. QFT describes particles as quantized oscillations of their corresponding fields, so there is a field associated to each fundamental particle. The newly discovered particles, if supposed to be all fundamental particles, needed a corresponding large number of new fields. This proliferation of fields constituted an odd fact that theoreticians were not able to explain. The very concept of fundamental particles became in trouble in the physics of hadrons. Furthermore, QFT was facing the problem of infinities - then solved through renormalization techniques - and, more than anything, quantum field theory seemed to not be able to describe strong interaction because of the breakdown of the perturbation expansion for a large coupling constant. All these problems put theoreticians in crisis and, as a consequence, many of them tried to modify the classical approach, bypassing quantum field theory. This resulted in the *S-matrix theory*, based on *nuclear democracy* (a term coined by Gell-Mann) and the *'bootstrap'* approach. The S-matrix is a matrix whose elements give the probabilities for different pattern of free incoming particles and free outgoing particles, that is particles are considered to arrive from infinity, interact and then go to infinity again, and a certain probability is calculated for each process. In this viewpoint, what happens in the interaction region is completely ignored; this region is regarded as a kind of 'black box' inside which we refuse to look. Using this approach one tries to neglect all 'unobservable' processes, focusing only on observables such as the scattering amplitude. This approach avoids quantum field theory and underlines an instrumental approach, being the description of particles interactions abandoned. Formally speaking, the idea was to put physical constrains on the mathematical form of the S-matrix, using the minimal amount of empirical data; mathematical consistency was supposed to be enough to correctly define the S-matrix

and the physics it describes.
The main supporter of this philosophy was the Berkeley physicist Geoffrey Chew, who argued to abandon the concept of fundamental particles in hadronic physics, proposing to treat hadrons 'democratically', so no hadron was considered fundamental. As a consequence, Chew followed a 'bootstrap' approach: if no hadron is fundamental, they are all bound states of other hadrons, and they are held together by the exchange of hadrons too, so it is a self-powered process.
The consistency conditions imposed on the S-matrix were Lorentz invariance, analyticity [2], crossing, unitarity. This idea comes from a work by Landau [34] were he found a connection between Feynman diagrams - used to represent the contributions to the S-matrix as a perturbative series for weakly coupled fields - and the S-matrix regarding some 'singularity conditions': poles in the S-matrix were found to correspond to three graphs, and branch points to loop diagrams. This *mathematical* discovery led scientists to hope that imposing mathematical conditions on the S-matrix may be enough to achieve physical predictions about particle processes.
In this context, in 1959 Regge proposed a theoretical scheme accounting for hadronic processes [48]. It is from this 'Regge theory' that the pre-history of string theory can be dated back. Regge proposed to consider angular momentum as a complex variable in order to analyze the potential scattering problem. Following this idea, amplitudes become simple poles in angular momentum corresponding to the propagation of intermediate particles and whose location was determined by the energy of the system. So, one gets Regge poles corresponding to particles whose spin is linearly related to their mass, and tuning the energy one can get a so called 'Regge trajectory', a graph describing the properties of scattering amplitudes. Once the energy is tuned to integer or half-integer values of the angular momentum, new particles corresponding to that energy are predicted. Each trajectory represents a family of particles with same quantum numbers except for spin, so the classification of these trajectories allows to classify different families of particles.
This relationship between masses and spins were confirmed by experiments. Looking at Regge trajectories we see that they are unbounded from above, so they contain particles with increasing values of the spin, allowing to analyze the exchange of high-spin particles. this was a nice feature because QFT was not able to treat with spins greater than 1. In Regge theory the exchange of these kind of particles is represented by an exchange of composite objects named 'Reggeons'. Also, in order to explain the slowly rising cross section observed for collisions at high energies, a Regge trajectory was identified with a new particle, the 'Pomeron', carrying no quantum numbers and responsible for that behaviour.
Regge theory was related to Chew's bootstrap approach because of the correspondence between mathematical poles and physical particles. The poles in the S-matrix corre-

[2] it was already known that the S-matrix was an analytic function.

sponds to resonances, their masses are determined by the position of the pole in the complex energy plane, the residues of the poles give the couplings and the imaginary part of a complex pole gives the lifetime of the corresponding resonance. This being so, Chew argued that imposing mathematical constrains on the S-matrix one should be able to determine its expression, and being S-matrix related to physical properties by the above correspondence one should be able to extract physics from it and make predictions. No use of equations of motions should be made, and only a finite number of coupling constants was considered to be necessary; the idea was to use only general mathematical principles, neglecting dynamics. In this framework, the bootstrap approach was performed by generating a pole in a certain variable by summing an infinite number of singularities in some other variable, so physically a particle can be seen as a bound state of other particles, achieving nuclear democracy.

In this context a fundamental step was made by Dolen, Horn and Schmid in 1967 who introduced a duality principle now called 'DHS duality' or 'FESR duality'. As Rickles points out, "they noticed that Regge pole exchange (at high energy) and resonance (at low energy) descriptions offer multiple representations (or rather *approximations*) of one and the same physical observable process"([49], p.38). This means that a large number of resonances (poles) produced in the s-channel describes the same physics of an exchange of Regge poles in the t-channel. The two processes describe the same physics, so they can be considered equivalent. This duality principle represented the formalization of the bootstrap principle, stating an observational equivalence between a description without forces, but with resonance production, and a description with forces, mediated by particle exchange, corresponding exactly to Chew's principle of generating a pole by summing over an infinite number of singularities. Duality relates the Mandelstam variables s and t, representing respectively the energy and momentum transfer and corresponding to the s-channel with resonance production and the t-channel with Regge pole exchange, so the duality principle then ca be formalized as a symmetry of the amplitude in these variables: $A(s,t) = A(t,s)$.
The DHS duality can be summarized saying that "direct s-channel resonance particles are generated by an exchange of particles in the t-channel"[3]([49], p.39). As a consequence of this duality, only one of the two Feynman diagrams corresponding to s- and t-channels was to be considered in the calculation, they representing the same process. Another consequence of this duality is what Rickles, following Ralf Kramer [30], calls an "*epistemic gain*: if we know about the resonances at low energies, we know about the Regge poles at high energies." This duality was easy to be explained in the S-matrix framework, because the s- and t-channels had identical initial and final free states, and they represent the only factors considered in this approach. In any case, the formal demonstration of this duality had to wait until the Veneziano formula and, in a more important sense, until

[3]DHS duality is often called FESR duality, because it was implicit in the 'Finite Energy Sum Rules'.

string theory formulation, where it is explained by conformal invariance of the string worldsheet.
Another important step to be mentioned is the narrow-resonance approximation made by Mandelstam when trying to model the rising Regge trajectories [37]. Using this approximation, where hadrons are treated as stable particles, Mandelstam introduced two new constants, the Regge slope α and the intercept $\alpha(0)$, which are fundamental in dual models and string theory.
Before considering the Veneziano amplitude I would like to underline that dual models did not make any prediction, but they were nevertheless considered to be genuine physical theories, both because there were no alternatives and because they seemed to account for many inexplicable features of hadronic dynamics.

2.1 Phase 1 [1968-1973]

I have just presented the situation before 1968, when Regge trajectories were discovered and FESR duality elevated to the status of a principle connecting different kinds of processes. This duality, in any case, was not proved, and many people though that duality was not possible.
The next step was to include FESR duality in the framework of S-matrix theory to obtain a model able to describe hadrons.
In 1968, Gabriele Veneziano [59] found a formula capable to make this connection. He discovered that the Euler β-function was able to account for all the desired properties: it described the Regge trajectories and also had (almost all) the required properties for a scattering amplitude, among which the FESR duality. His formula represents the first example of a complete dual resonance model, even if it was achieved in a certain approximation scheme, that is the narrow-resonance approximation, which violates unitarity. In any case, including both Regge behaviour and S-matrix consistency conditions, it can be considered a solution to the bootstrap approach. Using Mandelstam variables s and t, Veneziano's formula can be written as follows:

$$A(s,t) = \frac{\Gamma[-\alpha(s)]\Gamma[-\alpha(t)]}{\Gamma[-\alpha(s)-\alpha(t)]} = B(-\alpha(s), -\alpha(t))$$

where $B(-\alpha(s), -\alpha(t))$ is the Euler Beta function. The singularities of this function represent a set of infinite poles pointing to locations of particles on Regge trajectories, reproducing the Regge behaviour. The FESR duality is also explicit, being $A(s,t) = A(t,s)$. The narrow-resonance approximation (or zero-width approximation) giving an infinite set of poles was able to account for both resonances and Regge poles,

and thanks to duality one could consider as equivalent a description in terms of infinitely many hadron states constituting the intermediate states and a description in terms of an exchange of such states, which is responsible for the strong force. Veneziano's dual resonance model was able to perform the boostrap approach and also contained FESR duality, but it was to be replaced by a more realistic and general model. The next steps were to add unitarity, eliminating the approximation scheme, generalize from 4-particle to N-particle amplitudes, consider both the tree-level and loop amplitudes, add spin and isospin and understand the physics behind it, or what it is a model of.
Veneziano's article caused a great mobilization in the theoretical physics research field. All physicists concerned with dual models were trying to generalize his model, and soon after the publication many steps were made in this direction. The N-point generalization was soon made by different physicists and groups, among which we recall Chan Hong-Mo, Bardakci and Ruegg, Goebel and Sakita, Koba and Nielsen. The latter two also reformulated Veneziano's formula using the so called 'Koba-Nielsen' variables. In their formulation, duality is expressed by Mobius invariance, and a later development by Fairlie and Nielsen was crucial for the understanding of conformal invariance. Fairlie, with Keith Jones, also discovered that imposing the (unphysical) condition $\alpha(0) = 1$ one finds a tachyonic ground state.
In 1969, Jack Paton and Chan Hong-Mo [44] added isospin to the Veneziano model. They made it assigning elements of SU(3) to the external lines of a scattering diagram, and this method will be the standard one to attach quantum numbers to the end-points of open strings; at this stage, in any case, there were no connections with strings.
Virasoro constructed a new dual model different from Veneziano's one, reducing to Veneziano's formula for intercept 2 [60]. They shared the same properties but Virasoro's formula possesses SL(2,C) invariance; The generalization of this new model was made soon as well.
The most important step in the generalization process of the Veneziano amplitude is the recovery of unitarity. This step is related to the addition of loop contribution to tree-level amplitudes, because the Veneziano amplitude was considered to be the lowest order term of a complete unitary theory, so unitarity was expected to be recovered when the whole perturbation series was summed up. To this aim, Fubini, Gordon and Veneziano developed an operator formalism used to analyze the factorization of tree diagrams; this formalism allowed them to construct loop diagrams by sewing together tree diagrams. Unitarity then was achieved demonstrating factorizability,and it was proved independently by Bardakci and Mandelstam [5] and by Fubini and Veneziano [15]. The most important formulation for future developments, in any case, was due to Nambu, who demonstrated that the Veneziano model could be factorized using an infinite set of harmonic oscillators [40]. This formulation allowed a new perspective, because it clearly pointed to a more physical picture of the Veneziano model, so it was a crucial step towards a realistic theory. Furthermore, the harmonic oscillators formalism, using creation and annihilation operators, is a paradigmatic and well understood formalism in theo-

retical physics, so it also allowed for new mathematical developments; the dual model's spectrum could be recovered using the above cited operators in order to construct an infinity of states forming a Fock space. The familiarity of this formalism, in addition to make thing much more transparent, was crucial for the following interpretation in terms of strings, because harmonic oscillators clearly pointed to an underlying oscillating system.
Up to this point, only the vertex for the emission of the lowest lying (bosonic) states were constructed, so the first thing to do in order to obtain a real model was to construct a vertex for the emission of a generic state. This was made by Fubini and Veneziano, and only later it was realised that these states admitted an interpretation in terms of strings.
In any case, this formalism also revealed a serious problem, that is the presence of infinitely many ghost states. It was Miguel Virasoro who partially solved this problem in 1969, devising an infinite-dimensional gauge algebra from the oscillator and using an infinite class of gauge conditions in order to eliminate the infinity of ghosts. The problem was only partially solved, because this procedure could be performed only in the unphysical case of unit intercept, $\alpha(0) = 1$, which results in a tachyonic ground state (with a massless first excited state). Virasoro was aware that this was an unrealistic case, but hoped to be able to generalize it later. So, the presence of ghosts was exchanged with the presence of a tachyon.
Soon after Virasoro's paper, Del Giudice e Di Vecchia showed that physical states were orthogonal to spurious states, and together with Fubini constructed the space of physical (later called 'DDF states') for the unit intercept case.
Summarizing, the new operator formalism clearly suggested a deeper investigation of the physical system responsible for generating the spectrum, but it was incomplete, not including fermions, and also presented many problems, the main one being the presence of a tachyonic ground state.

I have just discussed the birth of dual models, the phase 1A covering the years 1968-1969. In order to complete this pre-history of string theory as we know it, I am going to present the phase 1B, framed by Rickles in the period 1970-1973, when the concept of strings was introduced for the first time.
Until 1970, dual models were purely mathematical models, but once the operator formalism was devised, many physicists understood the possibility of a real physical description of hadronic processes. For this reason, the formulation of hadronic string theory was not unique, but different physicists arrived at it almost at the same time.

Leonard Susskind is one of the founders of string theory, and even today he is one of its strongest supporters. Before the discovery of Veneziano's formula, he was trying to analyze hadronic processes in the 'infinite momentum frame' (today called 'light-cone frame'), a Galilean-invariant frame where standard quantum mechanics - and its well-

understood tools - was applicable. Once he came to know that formula, he applied his method to the dual model, finding a spectrum with equally spaced energy levels clearly pointing to an underlying quantum harmonic oscillator system. The final step was to find a precise oscillating system able to reproduce Veneziano's formula. Even if he firstly thought that its harmonic oscillator formulation merely represented a mathematical analogy, he soon realised that his model was pointing to a precise picture: something like a rubber band, or a violin string.
In this picture, the distinct particles lying on a Regge trajectory were to be considered as different modes of oscillations of the same object, an oscillating string with a quark and an antiquark at the endpoints. Susskind immediately understood that the Veneziano amplitude could be obtained visualizing the process as a scattering between strings merging together forming a single string and separating again.
Susskind notes that this idea initially had little success, and argues that it was because he tried to investigated what physicists working on S-matrix theory and dual models refused to investigate, that is what was hidden in the interaction region, the 'black box'. However, Rickles supports that more likely it was because the idea of a fundamental string was in direct conflit with the bootstrap approach, where no fundamental objects where supposed to exist. For this reason he underlines that S-matrix theory, resulting in this hadronic string model, was not abandoned only because it lost the competition against QCD, but mainly because it resulted in a physical model contradicting its fundamental principles.

Yoichiro Nambu, at the same time, derived an expression for the internal energy of a meson pointing to a system such as a quantized string or a cavity resonator. He achieved this result by using the operator formalism of creation and annihilation operators, which he deeply analyzed when trying to reproduce the Regge behaviour in a non-local field theory framework, but without interesting outcomes[4]. Nambu initially did not understand the precise system his model described - it may be a string, or an hollow body, or something else - and only after having investigated the Koba-Nielsen representation of the beta function he realised that the underlying system was an oscillating string. Again, the idea of a rotating and oscillating string with quarks at the endpoints was a natural picture able to explain the Regge behaviour: rotating, a string experiences a centrifugal force stretching itself apart, increasing its tension; being the tension related to the energy per unit length, and so to the mass of the string, and the rotation representing the spin, this model was able to explain the spin-mass relationship found for mesons. Moreover, being the tension proportional to the length, this model was also supposed to account for the confinement of quarks.
At first, Nambu was undecided about the way to follow: the hadronic string model had

[4]The hypothesis that particles are not point-like objects, but extended entities, was made many times in history of science, but without any interesting result.

mathematical problems, but it seemed to be able to explain the Regge-behaviour and confinement; at the same time, a new infinite-component quantum field theory emerged, able to explain the stability of hadrons, but itself not being without any problem.

Holger Nielsen followed a line of reasoning very different from that of Susskind and Nambu to arrive to the concept of strings. Nielsen advanced the hypothesis that the Harari-Rosner duality diagrams (diagrams representing equivalent scattering channels) were a limiting case of infinitely many Feynman diagrams constituting something like a 'fishnet', which taking the limit of infinitely many particles takes the form of a string world sheet. Nielsen used an electrostatic analogy with a two-dimensional conducting disc in order to compute the Veneziano amplitude for N particles. The idea that led Nielsen to the concept of strings was quite intuitive: the strong interaction has a large coupling constant, so higher order Feynman diagrams are increasingly important in the expansion. For this reason, one can consider the $n \longrightarrow \infty$ limit (where n is the order of the diagram) to give the leading behaviour, a limit graphically equivalent to an infinitely dense network of Feynman diagrams viewable as a two dimensional surface, that is a string world sheet. Again, this picture pretended to explain the Veneziano amplitude through the merging and splitting of strings and also to account for confinement.
Talking about Nielsen, it should be mentioned a later work with Paul Oleson where they used a different analogy, that is a superconductor [42]. This work shows how Nielsen and Oleson tried to merge together the S-matrix theory with traditional quantum field theory. Quantum field theory was supposed to be a useful tool both in order to deepening the underlying structure of dual models and to solve and reinterpret technical issues. Furthermore, this work - where strings are viewed as magnetic fluxes between two quarks representing magnetic monopoles - was highly influential for the development of QCD.

At this point we can recognize two different approaches in the hadronic string models, a mathematical approach based on the operator formalism and a geometrical approach, the latter coming from the former but providing a physical picture. The geometrical interpretation of the Veneziano amplitude was not immediate, because strings were used at first only as a useful analogy, in a similar way to the use of quarks at the birth of QCD. Rickles dates back the concept (and coining) of string worldsheet to the paper by Susskind 'Dual-symmetric theory of Hadrons' published in 1970 [57], where for the first time duality was explained in terms of conformal invariance of the worldsheet. Soon after, Susskind, Kraemmer and Nielsen presented for the first time a quadratic worldsheet action. Susskind parametrised the points of the worlsheet using the space coordinate Θ and the time coordinate τ on the worldsheet and wrote the equation of motion in terms of a dynamical variable $X_\mu(\Theta, \tau)$.

The equation of motion reads:

$$\frac{\partial^2}{\partial\tau^2}X_\mu(\Theta,\tau)-\frac{\partial^2}{\partial\Theta^2}X_\mu(\Theta,\tau)=0$$

This is the generalization of the equation valid for a point particle, so here there is no geometrical description in terms of a one-dimensional object. It was Nambu who, in the same year, managed to generalize the action principle for point-particles to an action-principle for one-dimensional objects. This was achieved by replacing the minimization of a particles trajectory with the minimization of the area spanned by the string worldsheet. This idea came from the visualization of a string as a chain of infinitely many particles propagating in parallel, so that the infinitely many trajectories to be minimized look like a surface whose area is to be minimized as well.
Nambu wrote the action:

$$I=\frac{1}{4\pi}\iint\left(\frac{\partial X_\mu}{d\tau}\frac{\partial X^\mu}{\partial\tau}-\frac{\partial X_\mu}{d\xi}\frac{\partial X^\mu}{\partial\xi}\right)d\xi d\tau$$

giving

$$\left(\partial^2/\partial\tau^2-\partial^2/\partial\xi^2\right)X^\mu=0\quad\left(\partial X^\mu/\partial\xi=0,\text{ when }\xi=0,\pi\right)$$

where ψ and τ are respectively the worldsheet space and time coordinates. The action can also be rewritten as:

$$S_{N_{ambu}}=-T_0\int_{\pi_i}d\tau\int_{\sigma_1(\hbar)}d\sigma\sqrt{\left(\frac{\partial X_\mu}{d\tau}\frac{\partial X^\mu}{\partial\sigma}\right)^2-\left(\frac{\partial X_\mu}{d\tau}\frac{\partial X^\mu}{\partial\tau}\right)\left(\frac{\partial X_\mu}{d\sigma}\frac{\partial X^\mu}{\partial\sigma}\right)}=\int d\tau d\sigma S^2$$

where T is the tension of the string and it is related to the Regge slope via the formula: $\frac{1}{T}=2\pi\alpha'$. In this last formulation, σ and τ are the worldsheet coordinates, while $X_\mu(\sigma,\tau)$ (with $\mu=1,..,d$ where d is the dimension of spacetime) performs the embedding of the wordlsheet in the target (Minkowiski) spacetime. This action is invariant under the arbitrary reparametrizations $\delta X^\mu(\sigma,\tau)=\psi^\alpha\partial_\alpha X^\mu(\sigma,\tau)$, these transformations representing an infinite dimensional symmetry group for the action. This reparametrization invariance implies that it is always possible to find a suitable change of coordinates to gauge away the oscillations parallel to the motion, so that we can recognize transverse oscillations as the only physical modes. As a consequence, this suggests the possibility of finding an action free of ghosts. The worldsheet action is referred to as the 'Nambu-Goto action' because they contemporary found an almost equivalent expression for the action, even if only Goto published his result.
As I have already mentioned, one major problem of these first attempts to construct hadronic string models was the presence of ghosts. The Nambu-Goto action, with its reparametrization invariance, offered a way to solve this problem, suggesting to focus

on transverse modes only. This problem then was solved through no-ghost theorems, where different symmetries were imposed on the worldsheet depending on the chosen formulation. All of them, in any case, implied an infinite set of subsidiary conditions corresponding to the conditions already found by Virasoro. In order to work, no-ghost theorems forced the value of the intercept to be one, and this implies the existence of a tachyonic ground state.
Once the transverse modes were recognized as the physical modes, they were quantized, and Lorentz invariance showed to require 26 spacetime dimensions. Also, the no-ghost theorems independently provided by Brower, Goddard and Thorn, required $d = 26$.
The complete geometrical description of a propagating quantized string was given by Goddard, Goldstone, Rebby and Thorn (GGRT) in a paper in 1973 [17], where it was shown that, once quantized, the action gives the Regge trajectories of the dual models. The GGRT paper clearly described all the consequences of the Nambu-Goto action and the no-ghost theorem demonstrated the mathematical consistency of the theory, so hadronic strings were finally recognized as a real physical interpretation of dual models and no longer a mere analogy. In any case, hadronic string theory presented many problems, the most important being the existence of massless particles not observed, the large number of spatial dimensions and the presence of a tachyon.
Hadronic string theory, representing the microscopic interpretation of dual models, started to be studied separately from the latter, becoming an independent research programme. The GGRT paper published in 1973 accounted for a complete description of a free quantum relativistic string propagating in a 26-dimensional spacetime. The remaining important task was to add interactions, and it was carried out by Mandelstam who was able to formalize the already mentioned idea of the strings joining and splitting at each vertex using the operator formalism. Mandelastam showed that his interacting model was able to recover the dual resonance model and filled the gap in the dynamics of hadronic strings.
At this stage the first quantization of a relativistic string was complete, resembling a two-dimensional field theory on the worldsheet. The second quantization, allowing for the description of an arbitrary number of strings, that is describing a quantum field theory of strings, was performed by Kaku and Kikkawa in 1974.
The first one to find the condition of $d = 26$ dimensions was Claude Lovelace who, in 1970, found that this condition was demanded to avoid violation of unitarity. In particular, it was required in order to transform certain branch cuts into simple poles, thus allowing a particle interpretation in terms of a so-called Pomeron, later understood as a closed string and finally identified as the graviton (while Reggeons were later identified as open strings) . Being the number of dimensions 26 instead of 4, this result was not taken seriously at first, but in any case the fact that the number of dimensions was fixed by consistency conditions on the S-matrix was a great achievement in the bootstrap philosophy, where parameters where supposed not to be arbitrary, but an output given by consistency conditions. This bootstrap of spacetime was not taken seriously not only by

Lovelace but by the entire scientific community, until it became to reappear as a consequence of other consistency conditions such as the no-ghost theorem and the Lorentz invariance of the quantized theory. In particular, 26 was found to be the maximum number of dimensions for which no ghosts were present in the spectrum, and Brower also argued that the theory was non-renormalizable above that number. Furthermore, hadronic string theory was found to be affected by a conformal anomaly, that is conformal invariance was lost once the theory was quantized; as a consequence, the theory violated unitarity and was supposed to be non-remormalizable too. In order to restore the consistency of the theory conformal invariance was needed, and it was found that the conformal anomaly cancels only in 26 dimensions. For these reasons, the $d = 26$ result seemed to be not only an unphysical result to be accomodated, but something more significant. It is important to underline that it was the first time in the history that the number of spacetime dimensions was obtained as a result of the theory, and not put by hands.
Rickles dates the end of the first age of string theory around 1973/4 , when it was considered merely as a theory of hadrons. At the end of this period the generalized Veneziano model and the Shapiro-Virasoro model were known to be reproduced by a quantum theory of open and closed strings respectively; both the free and interacting theories were developed and the consistency conditions well understood. In any case, the hadronic string theory was not understood as a physical theory describing reality, also because of the large number of extra dimensions. It was mainly used as a useful picture, something to which the operator formalism could be related to achieve a better visualization of what one was doing.

One of the necessary steps towards more realistic duality models was the inclusion of fermions in the spectrum.
In 1970, Pierre Ramond found a way to construct a dual theory of fermions using a sort of 'correspondence principle' between point particle theories and dual models. As Rickles explains, "Ramond invoked a correspondence principle whereby operators in the point particle case are to be thought of as averages over internal motions of the hadronic system" ([49], p.100). Following this method, Ramond firstly generalized the Klein-Gordon equation recovering the bosonic spectrum, and also obtained Virasoro's conditions and algebra. Then, he extended this method to generalize the Dirac equation, obtaining a new kind of algebra containing both commuting and anti-commuting harmonic oscillator operators, the first example of a so called 'superalgebra'.
In the same year, André Neveu and John Schwartz were trying to add fermions too, constructing a new dual model with anticommuting operators. In their model the tachyon corresponding to the leading trajectory, that is the state with $M^2 = -1$, was eliminated and replaced by a new tachyon corresponding to the next trajectory, with $M^2 = -\frac{1}{2}$, which they interpreted as a pion and hoped to make its mass positive by finding some mechanism. While being unable to work out successfully their model, they understood

that Ramond's model was deeply connected with their, and maybe they could be merged together. Neveu and Schwartz, with the contributions and following generalizations of Charles Thorn, Edward Corrigan, David Olive and Mandelstam, successfully managed to incorporate Ramond's model in their model, finally obtaining a spectrum with both bosons and fermions. It is important to note that Ramond, Neveu and Schwartz never referred to the string picture in their works. Only after Mandelstam's work on the interactions in the above cited models - analyzed in terms of joining and splitting of strings - those models were explicitly interpreted as describing 'spinning strings'.
The Neveu-Schwartz-Ramond (NSR) model can be seen as the first superstring theory, but only in the sense that it incorporates both bosons and fermions; there was no explicit relation to supersymmetry [5].
The last fundamental step characterizing the first age of string theory is the analysis of the zero slope limit of dual models carried out by Joel Scherk in 1971.
In dual models there were a fermionic sector, with leading trajectory $\frac{1}{2} + \alpha'$, a mesonic sector, with leading trajectory $1 + \alpha'$, and a Pomeron sector, with leading trajectory $2 + \frac{\alpha'}{2}$. The idea was to take the limit $\alpha' \longrightarrow 0$ in order to make masses infinite, keeping only massless particles. In particular, Scherk carried out a perturbative expansion in a parameter $\lambda = \frac{g}{\alpha'}$ sending both the coupling constant g and the Regge slope α' to zero keeping λ fixed. As Rickles points out, this is equivalent to make the inverse procedure of Nielsen's approach giving an infinitely dense network of Feynman diagrams, the sort of 'fishnet' I told about discussing duality diagrams. Thus this inverse procedure represents a low energy limit of dual models giving as a result standard quantum field theory, as Scherk demonstrated recovering the ϕ^3 Lagrangian field theory from the generalized Veneziano model, so that the Veneziano amplitude represents the tree-level approximation of the ϕ^3 theory.
Most importantly, it was shown later that in the zero slope limit dual models describe standard classical field theories, namely Yang-Mills and gravitational field theories. This is a crucial step in the early history of string theory, named by Rickles 'theoretical exaptation', a term borrowed by evolutionary biology referring to the change in function of a certain aspect of physiology, where the analogy is evident: string theory - even if not immediately, as we are going to see - started to transform from a theory of hadrons to a unifying theory of all forces.
Thanks to the exaptation made by the zero slope limit , many problematic features of the hadronic string theory became naturally interpreted in the optic of a unifying theory: the massless modes, while representing an odd feature for a theory of strong short-range interactions, where re-interpreted as the gauge bosons and leptons, represented by open strings, and the graviton (before supposed to be a new particle, the already mentioned Pomeron). In particular, they were Neveu and Schwartz to show that dual models describe Yang-Mills theories in the zero slope limit, while Tamiaki Yoneya demonstrated

[5] Rickles also notices that at this stage the dimensionality of spacetime was not constrained

that Einstein gravity was the low energy description of the Virasoro-Shapiro model. For reasons I am going to discuss, the process of exaptation did not happen immediately after this work, but took some time. In any case, the analysis of the zero slope limit made an important connection between dual models and standard quantum field theory. In particular, being the strings length equal to $\sqrt{\alpha'}$, the recovering of classical field theories in the limit $\alpha' \longrightarrow 0$ suggests a natural interpretation of point-particle field theories as the low energy limit of a string field theory.
As a consequence of this discovery, a strong link between dual models and field theories was established, allowing for a better understanding of the former by using standard and well understood methods of the latter. It is interesting to note that dual models derived from the S-matrix and bootstrap philosophy, whose aim was to find an alternative and more general approach to quantum field theories, but they were finally found to be strictly related to them.

2.2 Phase 2 [1974-1983]

This second period, is often viewed as a period of crisis for string theory, but Rickles was able to show that this is not true at all. Interest in hadronic string theory started to decrease at the end of 1973, and many think this was due to QCD defeating dual models. While we can surely recognize QCD as a main factor, this was not the only one. Hadronic string theory had many problematic features, such as the existence of massless particles and the large number of extra dimensions. These problems were reinterpreted in the process of exaptation, but in the context of hadronic physics they contributed to decrease the trust in dual models once QCD appeared in the research landscape. Hadronic string theory was considered interesting because of its topological structure, but once it was shown that QCD was able to generate the same planar diagrams in the $\frac{1}{N_c}$ expansions, interest in this subject started to decrease. In any case, as I am going to discuss, hadronic string theory had an important part in the understanding of confinement, giving a clear qualitative picture in terms of a string holding together quarks at its endpoints, so to some extent hadronic string theory was absorbed by QCD.
In this period, work on string theory continued thanks to some physicists enamoured of its mathematical structure, thinking that it should have some role in the description of nature. Anyway, the most of the scientific community chose to pursue QCD, a standard quantum field theory less problematic than dual models. QCD has no tachyons in its spectrum, and describes strong interactions in four dimensions. Furthermore, dual models could not explain scaling, a fundamental characteristic of strong interactions. Also, being strings extended objects hadronic string theory had very nice UV properties, precisely because interactions were not considered to happen at a single point, but rather

'smeared' out. This feature, while of paramount importance for the trust in string theory as a quantum theory of gravity, was an odd one for a theory of strong interactions, where hard scattering events have to be explained. QCD was able to do so thanks to asymptotic freedom, and it was able to explain scaling differently from hadronic string theory. Furthermore, between 1967 and 1973 deep-inelastic scattering experiments carried out at SLAC pointed to a picture of protons as composite objects. They were subsequently recognized as quarks, at first considered by many only mathematical entities.
While hadronic string theory was not able to deal with experiments because of its soft scattering amplitudes and failed to give any prediction, QCD had predictive power, and in 1974 charmonium was found. The success of a new quantum field theory was a blow for the bootstrap approach, whose birth was due to the apparent inability of quantum field theories to deal with strong interactions. Dual models were not anymore necessary, and were to be replaced by a much more powerful canonical theory.
In any case, QCD presented some problematic features, the main one being the difficulty to explain why quarks where not kicked off protons in scattering experiments; in a few words, QCD was in trouble with finding an explanation for confinement. This problem was due to scaling: strong interactions become stronger and stronger as distance increases so the coupling constant increases as well and the perturbation expansion breaks down. QCD fails to describe large distance phenomenology, failing to give an explanation of quark confinement. Hadronic string theory, on the other hand, gave a better description at lower energies, and for this reason Nambu, in 1974, proposed to incorporate it directly in the QCD research programme, in particular in order to account for the Regge behaviour. The idea was to use the Nielsen-Oleson model, where strings were represented as vortex lines in a field theory. This vortex lines had no endpoints, so Nambu added monopoles at the endpoints in order to capture the flux and make the flux tube of finite length, recovering the picture of dual strings. This new system, constituted by two monopoles tied together by a flux tube in a superconductor, was analogue to a system with a quark/antiquark pair tied together by a string. The string picture was able to explain both Regge behaviour and confinement because of its tension: the bigger the spin is, the bigger the centrifugal force separating the quarks is, so the tension increases as well; an increase of tension means an increase of energy, and so of the mass, explaining the Regge plot, and at the same time when quarks try to separate the increasing tension binds them together, explaining confinement. As Rickles explains, "the string model of hadrons provides a neat qualitative account of 'soft' processes (the Regge phenomenology, along with duality), while the quark model provides an account of 'hard' processes (deep-inelastic scattering): they are complementary rather than competing" ([49], p.123).
This bridge between dual models and field theories was further investigated by 't Hooft, who demonstrated that the topological structure of the perturbation series of U(N) Yang-Mills theories in the large N limit is the same of the one given by dual models. He was able to show that interactions in a two-dimensional gauge theory correspond to inter-

actions of quantized dual strings, also recovering approximately the Regge trajectories. These works showed a connections between dual models and gauge theories, and further underlined the non negligible role of hadronic string theory in this period. In view of this situation, we can say that hadronic string theory was not simply replaced by QCD, but rather absorbed in the QCD research programme. In any case, if not for exaptation process, string theory would have been abandoned as an independent theory. Thanks to some strong supporters, string theory survived, starting to be considered something more than a mere theory of hadrons, and it is important to notice that this process started before QCD 'won the competition'.

After this brief discussion about QCD and hadronic string theory, we can come back to the zero slope limit and its consequences. Once string theory was finally recognized as a potential theory of all forces, the Regge slope was rescaled by $\sim 1/GeV^2$ to $\sim 10^{-38}/GeV^2$, in terms of length corresponding to a switch from $\sim 10^{-13}cm$ to $\sim 10^{-33}cm$, equivalent to a shift of the string tension of about 40 orders of magnitude. It is worth to remember once again the consequences of this exaptation: the massless modes, once viewed as a problem to solve, were reinterpreted as gauge bosons and leptons (open strings) and the graviton (closed strings); the UV soft behaviour - after some time - was seen as a fundamental feature able to tame UV divergences appearing in the point-like quantum field theories; finally, the extra dimensions, while being an odd fact in the optic of a theory of hadrons, could be understood thanks to string theory containing gravity: in General Relativity spacetime is dynamic, so the extra-dimensions was supposed to spontaneously compactify into extremely small spaces, thus being unobservable. In particular, compactification was fundamental in order to describe the four-dimensional phenomenology we observe, and also turned out to be a useful tool potentially able to solve different problems. While the string theory exaptation can be dated back to 1973/4, its potential was not immediately understood. Schwarz remembers[6] that one of the factors slowing this process was that at that time particle physicists were not concerned with gravity, studying processes were gravitational effects could be neglected, so they saw string theory only as a theory of hadrons which failed to explain strong interactions. This also explains why the spin-2 particle in the Regge theory's spectrum was not immediately recognized as the graviton, but as a new particle. Also, Rickles underlines that the research landscape was not yet mature for string theory; it was also thanks to parallel research programmes like supersymmetry and supergravity that this maturity was achieved and people started to take superstring theory seriously. One crucial step in the early history of string theory as a theory of all forces was the so called 'GSO projection'. In 1976 Gliozzi, Scherk and Olive (from which the acronym) found a method to get rid of the tachyonic ground state with $M^2 = -1$. They used a chiral projection to eliminate a large set of states, among which there was the tachyon

[6]Cf. [49], p.137.

too. In particular, the ground state of the NS (bosonic) sector contained a graviton, a massless antisymmetric tensor and a massless scalar, while the NSR sector (with also left-handed Majorana spinors) contained a massless scalar and a massless spin $\frac{3}{2}$ state, then recognized as the gravitino. In the NSR sector this projection left an equal number of fermions and bosons at each mass level, pointing to the existence of supersymmetry. In particular, the procedure followed shows the fundamental role of supersymmetry in the elimination of the tachyon and the influence of the contemporary work on supergravity. Gliozzi, Scherk and Olive discovered that string theory described supergravity in 10 dimensions in the zero slope limit. Once compactified, they found that pure supergravity in 10 dimensions gives supergravity coupled to matter in 4 dimensions, with the spin $\frac{3}{2}$ particle clearly pointing to supersymmetry. Being supersymmetry involved, the bosonic and fermionic states would have to pair up, allowing to project out of the physical spectrum all the states with no partner, among which the tachyon.
Finally, a consistent (tachyonic free) superstring theory was achieved, with Majorana and Weyl conditions giving a number of dimensions $d = 10$. However it is worth noting that at this stage spacetime supersymmetry was not explicit, but this was made manifest only in the '80s by Micheal Green and John Schwarz.

Supergravity was considered a promising theory in the 1970s, and superstring theory owes a lot to it. As we have already mentioned, supergravity prepared the ground for the acceptance of string theory, because in earlier period the scientific community did not show a good disposition towards unifying theories, and even less towards theories with extra dimensions. In any case, string theory also borrowed from it something else, like the tools for compactification schemes (using the so called 'Kaluza-Klein mechanism'), the GSO projection method and the classification of the different string theories adopted even today. Rickles, analyzing the number of publications about supergravity, dates the main period of work on this subject from 1975 "until 1988 at which point superstring theory was more secure and supergravity research became far less popular, due to the growing realisation that it would remain forever non-renormalizable because of the local degrees of freedom" ([49], p.148). The hope on supergravity was based on the possibility that supersymmetry might be able to get rid of the divergences found in gravitational quantum field theories, but it did not meet expectations. The work on supergravity, in addition to provide string theory with many tools, also showed with its failure that point-like quantum field theories were not able to describe the quantum nature of gravity, being unable to eliminate the infinities affecting them. This fact, together with other discoveries we are going to discuss, increased the trust in string theory as a viable theory. In 1981 Alexander Polyakov explored further the mathematical properties of string theory, leading to a greater understanding of them. Its work concerned the application of functional integration and its use in the development of a path integral formulation of string theory perturbation series. In this expansion, symmetries can be used to recognize

equivalent worldsheet manifolds in order to avoid over-counting [7]. The path integral formulation allowed for well-established tools from quantum field theory, caused connections with fields of mathematics such as statistical physics, and most importantly prompted the development of conformal field theory.
One of the major aspects of Polyakov's work concerns the critical dimension. He was able to show that a Liouville mode can be added to modify the number of dimensions. In particular, as we have already discussed the extra dimensions where required in order to cancel the conformal anomaly, but Polyakov demonstrated that this conformal anomaly can be eliminated also for d¡26 introducing a dilaton as a Liouville mode.
In any case, the most important work in string theory in the '80s was done by Micheal Green and John Schwarz. First of all, their aim was to make supersymmetry manifest, and in order to do this they used a new set of oscillator operators in the light-cone gauge formalism and constructed a physical Fock space without the necessity of any projection method. In addition to help in the computation of loop amplitudes, this new formalism made spacetime supersymmetry manifest, even if at the expense of manifest Lorentz invariance. This proof of spacetime supersymmetry in string theory also led to the classification schemes of superstring theories, based on the possible supersymmetry algebras in ten dimensions with spin ≤ 2. This classification scheme was borrowed by supergravity, and divided superstring theories in TypeI and TypeII. TypeI superstring theories have N=1 supersymmetry, are non-chiral theories with 16 supercharges, while TypeII theories have N=2 supersymmetries and can be further divided into TypeIIA, a chiral theory with 16 supercharges of one chirality and 16 of the other one, and TypeIIB, a non-chiral theory with 32 supercharges. TypeI theory were known to be 1-loop renormalizable, while TypeII theories were proved to be 1-loop finite by Green and Schwarz in 1981. TypeI superstring theories seemed to have a chiral anomaly, while TypeII theories, not being chiral, where not considered interesting, and for this reasons in the early '80s they were mainly studied because of their link to supergravity, reducing to the latter in the appropriate limits. Chiral anomaly, together with gravitational anomalies, constituted a very constraining factor in theory building, and were largely investigated in those years. They affected many theories and were also shown to affect supersymmetryc gauge theories in d=6 (mod4) and d=10 by Paul Townsend in 1984. For these reasons there was little hope of TypeI superstring theories to be anomaly free.
In TypeI superstring theories the chiral anomaly appeared in the exagon loop diagram. In 1984, Green and Schwarz published a crucial paper [19] demonstrating that the chiral anomaly cancels for the SO(32) and $E_8 \otimes E_8$ gauge groups. As Rickles points out, "while the open string theories could account for SO(32), $E_8 \otimes E_8$ did not appear in a string theory and so could be seen as pointing to the existence of an entirely new type"([49], p.160-161). This work by Schwarz and Green marks the so called 'first superstring rev-

[7] in particular, diffeomorphism and conformal invariance make worldsheet with an equal genus g equivalent. This recalls the already seen over-counting due to DHS duality in dual models.

olution', often considered to happen after a dormant period, but even if it was without any doubt a great discovery and a fundamental step in the history of string theory this narrations is not quite exact, because as we have already seen in the years before 1984 much work was made. Even if talking about a real revolution can be an exaggeration, the anomaly cancellation result had an incredible influence on the scientific community, and a large number of papers and works succeeded it. A consistent theory of all forces, free of anomalies and with a nice UV behaviour was found, so superstring theory was finally recognized as a viable unifying theory. One of the remaining important tasks to do was the compactification of extra dimensions, which had to be done in a manner to recover the phenomenology of our 4 dimensional world. It is in the optics of this further task that the anomaly cancellation result appears even more crucial. In fact the cancellation occurs for a phenomenologically interesting gauge group, $E_8 \otimes E_8$, with E_8 having a maximal $E_6 \otimes SU(3)$ subgroup, showing the possibility to recover the Standard Model gauge group[8]. This means that the anomaly cancellation occurs for a gauge group which can be related to low energy physics, as pointed out by Edward Witten in a paper published soon after Schwarz and Green's result. In the light of all these promises it is not surprising that superstring theory started to dominate the research landscape.

2.3 Phase 3 [1985-1995]

The anomaly cancellation result demonstrated string theory to be a consistent theory, and one of the gauge groups this cancellation occurred for, that is $E_8 \otimes E_8$, showed the possibility to recover the low energy physics of the Standard Model. The so called 'Princeton string quartet', formed by David Gross, Jeffrey Harvey, Emil Martinec and Ryan Rohm published a crucial paper [21] soon after this discovery, introducing a new string theory based on this gauge group, that is the 'heterotic string theory'. In this period the main goal was to use this new string theory together with compactification schemes in order to derive low energy physics. In absence of new experimental data, this was a major goal in order to give credibility to a candidate as a unifying theory. Already in 1985, critics were moved to string theory because of the impossibility to be tested against experiments, so theoretical progress was necessary to justify work in this research field. We will deepen these critics later, but it is important to remember that already from this period the scientific community started to be divided between strong supporters and opponents of string theory.

[8]One major problem before this result was that one cannot recover the SM gauge group only via compactification, because at least 7 extra dimensions would be necessary. Heterotic string theories have gauge groups already in 10 dimensions, before compactification, so the objective changed from adding symmetries to break them in the right way. For a deeper explanation , see [20], p.402-403.

Compactification is maybe the most important issue in the string theory context. The capability of recovering the physics as we know it depends on the ways we find to compactify extra dimensions, because the compactification scheme determines the low energy phenomenology. Compactification has been studied since 1975 by Schwarz and Scherk, and further investigated by Scherk and Cremmer the next year, who studied 'spontaneous compactifiation' solutions of General Relativity. As they points out in their paper, "these solutions represent a state where some of the matter fields have acquired position-dependent vacuum expectation values such that, in certain directions, space is so strongly curved that it closes upon itself. In other directions, where fields have constants vacuum expectation values, space-time does not close and is asymptotically flat. The non-compact dimensions can be thought of as the ordinary four-dimensional space-time, while the other, compact, directions are like an internal space" ([9], p.61). It is important to notice that, as the authors underlined in their paper, spontaneous compactification is a solution from General Relativity, which is embedded in string theory, so it is not an additional ad-hoc hypothesis. General Relativity teaches us that space-time is a dynamical object, so it is not surprising that compactification can arise in this framework.

Compactification was to be performed in a certain manner: dimensions were expected to compactify into a product space such as $K^{extra} \times \mathcal{M}^4$ where K^{extra} represents the compact manifold into which extra dimensions are compactified, and is expected to be of the order of the Planck scale in order not to be observable, and $\mathcal{M}^4$ is the ordinary 4 dimensions space we experience in our ordinary life. Being bosonic string theory defined in 26 dimensions and supersymmetric string theory consistent in 10 dimensions, we must account for this difference. The 16 dimensions of difference are then considered non-spatiotemporal, but instead expected to account for the particle's internal degrees of freedom. Only the remaining 10 dimensions are interpreted as spatio-temporal dimensions, so spacetime should have the structure $K^6 \times \mathcal{M}^4$. All the efforts pointed to discover the precise structure of the compact manifold K^6, and in trying to recover phenomenology from it [9].

A fundamental step forward in the search for such a manifold was made by Philip Candelas, Gary Horowitz, Andrew Strominger and Edward Witten in the paper 'Vacuum configurations for superstrings' published in 1985 [8]. In this paper they recognized so called 'Calabi-Yau' manifolds as the possible vacua satisfying the desired conditions for compactification. Quoting Yau himself, "if you want to satisfy the Einstein equations as well the supersymmetric equations - and if you want to keep the extra dimensions hidden, while preserving supersymmetry in the observable world - Calabi-Yau manifolds are the unique solution"([64], p.131). It is important to notice that Calabi-Yau manifolds have

[9] In particular, the so called 'Frenkel-Kac-Segal mechanism' was used to generate gauge groups by compactification for the heterotic string theory. In fact, being heterotic string theory a theory of closed strings, the standard procedure of attaching Chan-Paton factors to string endpoints could not be used.

SU(3) holonomy group demanded to have N=1 supersymmetry[10] in four-dimensions, and it is able to break E_8 down to a $E_6 \otimes SU(3)$, very promising for phenomenology[11].
Calabi-Yau manifolds were recognized as the correct manifolds for compactification, but to isolate a precise manifold was another story. The first manifold the authors used was not a correct one, because it gave four generations of particles instead of the known three generations. The compact manifold containing the extra dimensions is fundamental because it determines almost all the low energy physics: the particle content, the number of generations[12], the parameters' and constants' values and so on[13]. Furthermore, new spaces can be obtained by quotienting other spaces by discrete symmetry groups; in addition to modify the number of generations, this process also results in multiply-connected spaces and can be used as a tool per symmetry-breaking, another fundamental process for the low energy phenomenology.
After the isolation of Calabi-Yau manifolds, a problem was found threatening the whole construction: a conformal anomaly. Calabi-Yau manifolds are described by a Ricci-flat metric, which was found to breaks conformal invariance. Attempts where made to solve this problem, restoring conformal invariance considering small deviations from Ricci-flatness or relating non Ricci-flat to Ricci-flat metrics. These solutions, while not representing rigorous proofs, were accepted, with Doron Gepner clinching the case with solid mathematical arguments, as we will see later.
Initially the possibility of recovering low energy physics seemed to be close: a string theory with a gauge group containing the standard model gauge group was constructed, and a certain type of manifolds for compactification was found. Candelas, in 1988, used a computer algorithm to classify the possible Calabi-Yau manifolds, finding 7890 of them [7]. This was a large number, but not so large as to prevent the possibility of studying them and finding the correct ground state describing our world. In any case, as we are going to discuss, the following years saw a very different change of perspective.

The other fundamental issue faced in this period, linked to compactification, is the study of dualities.
One of the dualities we can find in string theory is 'target-space duality', or briefly

[10] One wants N=1 supersymmetry because "in four-dimensional supersymmetric theories with $N2$, the massless fermions always transform in a real representation of the gauge group, in stark contrast to what is observed in nature" ([20], p.415).

[11] To be precise, "Picking an SU(3) subgroup of $E_8 \times E_8$ can be interpreted as embedding the holonomy group H in $E_8 \times E_8$" ([20], p.477).

[12] The number of generations is given by the formula: $Ngen = |\chi(K)/2|$, where $\chi(K)$ is the Euler characteristics of the compact manifold K, a topological invariant.

[13] "Elementary-particle phenomena present us with many unsolved problems. If thes problems are addressed in the context of compactification of ten-dimensional string theory, then most of the puzzles can be translated into questions about the compact manifold K. For the most part we can write a dictionary translating questions about observed physical phenomena into questions about the compact manifold K. ([20], p.551)".

'T-duality'. This duality identifies mathematically different spacetimes - that is topologically and geometrically different - as physically equivalent spacetimes - that is they give the same physics and could be not recognized as different by experiments.
T-duality was first studied in 1984 by Kikkawa and Yamasaki where the existence of a minimal length was already found as a consequence [29]. In general, this duality shows that if we consider a closed string wrapped around a compactified dimension of radius R, the physical laws remain unchanged if you consider the $R \longrightarrow 0$ limit for winding number m and momentum n and if you consider the $R \longrightarrow \infty$ with winding number n and momentum m. To be more precise, laws are invariant under the exchange: $(m, n, R) \longrightarrow (n, m, \frac{\alpha'}{R})$ for $R \longrightarrow 0$. This duality works not only for compactification onto a circle, as it was firstly studied in the above mentioned work, but also for more complex schemes; in 1985 Norisuke Sakai and Ikuo Senda studied compactification onto a torus and, in addition to showing that T-duality still works in this case, they demonstrated modular invariance of the vacuum.
One important consequence of T-duality, already presented by Kikkawa and Yamasaki in their paper, is the existence of a minimal length. When the space compactifies strings wound around the compact dimensions tend to get smaller in order to reduce their tension energy, but simultaneously, because of the uncertainty principle, momentum gets larger, so there will be a certain point at which the two effects will be balanced. This point determines a minimal length, and should be achieved for $\frac{\sqrt{\alpha'}}{R_i} = 1$, where $\sqrt{\alpha'}$ is of the order of the Planck length. T-duality shows what happens for distances smaller than this minimal length: the equivalence of $R \longrightarrow 0$ and $R \longrightarrow \infty$ descriptions means that if one tries to probe distances smaller than the minimal length one is bounced back to distances bigger than that. The important point is that this behaviour has a clear explanation in the string theory framework, and such an explanation was given by David Gross through a mental experiment aimed to see how string theory works when used to probe sub-string distances [22]. This mental experiment concerns how strings work in scattering experiments at very high energies. Particle physics does not impose any limit for the resolution of distances through scattering experiments: the more we increase energy, the more increases our resolution power; ideally, if we could reach energies bigger than the Planck energy, we should be able to discover what lies beyond the Planck scale. This is due to the fact that the resolution power is related to the size of the probe we use, meaning that we can resolve distances up to the the size of the probe, and being particles point-like objects we expect to be able to resolve infinitely small distances, at least ideally speaking. David Gross showed that this is not the case for string theory. Strings are not zero-dimensional objects, but extended objects, and their size increases with energy. As a result, if we increase the collision energy in a scattering experiment, the energy makes the string size increase, so our probe gets even larger. This means that trying to resolve distances smaller than the string size, we make the last increase, then decreasing the resolution power. This represents a physical explanation of T-duality:

trying to do $R \longrightarrow 0$ we infinitely increase the string size, so we are left with a $R \longrightarrow \infty$ description. This is a purely stringy effect due to the extended nature of strings and their capability of winding around compact dimensions. This effect has physical implications related to the nature of spacetime. In fact, in 1988 Amati, Ciafaloni and Veneziano also analyzed the ultra-high energy/small distance regime to investigate the nature of spacetime at such scales. They found that, due to the soft behaviour of strings scattering amplitudes, string theory cannot resolve distances shorter than the string size. In this work they also argued that at such energy scales it could exist some sort of 'quantum geometry'.
T-duality relating different spacetime structures and the existence of a minimal length suggested new speculations about the nature of spacetime, underlying its fundamental role in string theory. In particular, it was clear that in order to understand the dynamical aspects of spacetime a background independent formulation was needed. String theory contains General Relativity, and spacetime dynamics is also of crucial importance for the compactification of extra dimensions, so it became clear that perturbation theory around a fixed spacetime was not enough to really understand string theory. This problem was faced by Witten and Strominger, who had proposed a (non-perturbative) superstring field theory as a possible formulation of string theory. This formulation was made in analogy to standard quantum field theory, with strings being created and annihilated by string field operators, but just like in standard QFT only a perturbation expansion around a fixed background could be achieved.
Before discussing the other dualities and their consequences, we must mention another 'stringy effects'. In addition to the standard Calabi-Yau compactification we find the so called 'orbifold compactification'. We have already and implicitly mentioned this compactificaton scheme when talking about new spaces obtained by quotienting Calabi-Yau manifolds with discrete group symmetries. The thus obtained 'orbital manifolds' (from which the contracted term 'orbifold') present singularities, differently from the smooth Calabi-Yau manifolds. Even if these singular points in the spacetime structure could be smoothed out, the interesting result found by Lance Dixon, Jeff Harvey, Cumrun Vafa and Edward Witten was that strings propagating on such orbifolds does not feel any singularity, any obstruction. This is, again, a purely stringy effect, because particles propagation would be effected by these singularities. Furthermore, the conservation of modular invariance for orbifolds results in additional 'twisted sectors' of states which can be used for alternative mechanisms of symmetry breaking and other methods to recover realistic theories.
Another work worth mentioning is the one by K. S. Narain in 1985, suggesting that $E_8 \otimes E_8$ and $SO(32)$ heterotic string theories may represent "two different vacuum states in the same theory"([41], p.378).
Narain, together with Sarmadi and Vafa, also studied 'asymmetric orbifolds'. They analyzed the possibility that right-movers and left-movers can be described independently on different orbifolds, meaning that the compactified dimensions might not be real space-

time dimensions.
All the attempts already described underline how much subtle and crucial is the concept of spacetime in string theory and how strictly related to the stringy nature of elementary objects it is.

Summarising, two anomaly free string theories were found. The heterotic string theory with gauge group $E_8 \otimes E_8$ was also promising for the construction of realistic models, because E(8) can break down to the Standard Model gauge group. Furthermore, Calabi-Yau manifolds were identified as the correct compactified spaces, and their SU(3) holonomy group can break down E(8) to $E_6 \otimes SU(3)$. Finally, if we consider the already mentioned multiply-connected spaces obtained quotienting by a discret group action, E(6) can be further broken down giving the Standard Model gauge group through a symmetry breaking mechanism where closed non-contractible curves (Wilson loops) around the holes act like Higgs bosons. Moreover, E(6) admits complex representations, without which chiral fermions could not be described.
The heterotic string is a theory of closed strings and can be seen as a hybrid of a bosonic string theory of left-movers in 26 dimensions and a superstring theory of right movers in 10 dimensions. The different dimensionality can be explained through the Frenkel-Kac mechanism applied to the 16 dimensions of difference: left-movers live in a product space of a 16-dimensional torus, accounting for particle's internal degrees of freedom, and a 10-dimensional spacetime, the same where right-movers live in. In this way a coherent spacetime picture is achieved.
All this nice features made string theorists to believe that a realistic 4-dimensional model could be constructed, and one of the main goals was to recover the three fermionic generations of the Standard Model.
A Calabi-Yau manifold giving three generations was finally discovered for the heterotic string theory, and the compactification scheme were chosen to preserve N=1 supersymmetry at the Planck scale.
As string physicists already knew at that time, having an anomaly free string theory and a compactification scheme was not enough: in order to match low energy experimental results one had to 'fix by hand' some parameters to select one specific vacuum solution. This is obviously opposite to the bootstrap philosophy claiming that no arbitrary parameters should enter theory building, but it was forced by the impossibility of experiments at the string scale: we cannot probe the spacetime structure at Planckian energies, so the only thing we can do is to use low energy results in order to select the correct ground state.
In any case, an high degree of degeneracy in the ground states was recognized as an obstacle to the realization of realistic models. The consequences of this fact were explicitly presented by Strominger in 1986 in the paper 'Superstrings with Torsion' [54].
This plurality was a treat to string theory's uniqueness claims, and more importantly to its predictive power. There were many approaches trying to tame this 'ground state

explosion', as Schellekens named it ([50], p.5), approaches followed today as well, even if modified by further discoveries. Many people hoped that some mechanism would be found able to select one single vacuum out of this plurality, maybe hidden in the non-perturbative regime.
Some steps towards uniqueness was made in those years, but of a different kind. In particular, following Rickles, we can recognize three kinds of pluralities in string theory: 'Plurality of Type 1' concerns the different kinds of string theories we know, such as TypeI, TypeIIA, TypeIIB, Heterotic SO(32) and Heterotic $E_8 \otimes E_8$, while 'Plurality of Type 2' concerns the degeneracy of ground states in a particular string theory of the above kind; 'Plurality of Type 3' will be introduced in the next chapter.
In a paper published in 1987, Doron Gepner tried to restore some sort of uniqueness trying to relate different ground states by gauge symmetries [16]. Rickles calls Gepner's framework a 'CFT-CY Correspondence', where CFT stands for conformal field theory and CY for Calabi-Yau manifolds. Gepner related the search for classical solutions to the string equations of motion to the analysis of 2D conformal field theories, showing that "a string theory compactified along the lines of a $\mathcal{M}^4 \times \mathcal{K}^6$ Calabi-Yau manifold approach corresponds to a two-dimensional free conformal field theory in four-dimensions and a two-dimensional field theory in six-dimensions: the conformal field theory determines the physics of strings"([49], p.195). The correspondence between conformal field theories and compactification schemes through Calabi-Yau manifolds costitutes an indirect proof that Calabi-Yau manifolds are the correct spaces for compactification, an argument that was not mined by the conformal anomaly we have already mentioned. In fact, being Gepner's CFT model an exact solution to the string equations, the correspondence with string theories compactified through Calabi-Yau manifolds gives an indirect proof that Calabi-Yau manifolds are exacts solutions of string theory equations as well.
Other groups were going to find similar results, and many of them found a strange link between the mathematical structure and the physical one. In other words, there seemed to exist different Calabi-Yau spaces giving the same CFT, that is different compact manifolds giving the same physical predictions. This ambiguity was resolved by finding a new symmetry, a mirror symmetry relating pair of manifolds which can be considered physically equivalent, thus avoiding double counting. As Rickles remembers, this 'Calabi-Yau manifold duality' or 'mirror duality' was made more precise by Greene and Ronen Plesser in 1990, when they also linked this one to other dualities; Strominger, Yau and Zaslow later explained clearly this link in a paper in 1996 entitled 'Mirror Symmetry is T-Duality' [56].
T-duality and mirror duality belong to a series of discoveries pointing to hidden links between string theories that culminated in 1995, As we have already discussed, these facts together with other stringy effects were a strong evidence of the subtle nature of spacetime in the context of string theory, and the importance of a background independent formulation was and has always been recognized by string theorists, differently from what some false narratives account.

I would like to summarize the main features characterizing string theory at the end of this period following Rickles, who quotes a list of the them as it was given by Murray Gell-Mann in a talk in 1987 ([49], p.200-201). Gell-Mann stated that string theory offered:

- an elegant, self-consistent quantum field theory,
- generalising Einstein's general-relativistic theory of gravitation treated quantum mechanically,
- in the only known way that does not produce infinities,
- parameter free,
- based on a single string field,
- but yielding an infinite number of elementary particles,
- some hundred of which would have low mass (although we don't know why they would be so very low!),
- including particles with properties like those of electrons, quarks, photons, gluons, etc.,
- with the underlying symmetry system essentially determined,
- and with the symmetry breaking connected with the behaviour of some extra, but perhaps formal dimensions.

This was the state of affairs at the end of the '80s, and it was going to change deeply in the mid-90s.

Quoting Polchincki'words from an interview with Rickles in 2009, "string theory went through this tremendous wave of activity in the 1984 to 1987-1988 period. from 1988 to 1995, there was a perception that it had slowed down. Now, in retrospect, huge amounts of stuff were done in those days: mirror symmetry, D-branes, Neveu-Schwarz branes, supergravity. Huge amounts of stuff being done, but nobody knew that it all fit together [14]" ([49], p.207). In this period the main goal was to investigate the non-perturbative regime of superstring theory, and as Polchinski remembers all the necessary tools were already on the table, waiting someone able to recognize their importance and interconnections.

[14] this quotation shows how a correct appraisal of a scientific research programme can be made only on a-posteriori, 'in retrospects'.

A crucial element for the investigation of non-perturbative string theory was the so called S-duality, first introduced by Anamaria Font, Dieter Lust, Luis Ibanez and Fernando Quevedo in 1990 in the paper 'Strong-weak coupling duality and non-perturbative effects in string theory'[14], where they conjectured a strong-weak coupling duality in the compactified heterotic string theory. This work can be seen as a generalization of a previous work by David Olive and Claus Montonen [39], who conjectured the existence of a electric-magnetic duality, that is of a Lagrangian dual to the Georgi-Glashow Lagrangian, where the roles of magnetic monopoles and electric charges can be interchanged through an S-duality mapping $e \longrightarrow \frac{1}{e}$ without any change in the physical predictions. This map clearly shows the strong-weak coupling duality: a perturbation expansion around the electric charge gives the same physical results of a dual field theory in a non-perturbative regime (if e is small, $\frac{1}{e}$ is large) where perturbative objects (the electric charges) are exchanged with non-perturbative objects (the magnetic monopole solitons). In order to analyze the theory at all values of the coupling constant, one needs states with a good behaviour under the renormalization of the coupling constant, and the electric and magnetic solutions at issue are such states, named BPS states ('Bogomolnyi-Prasad-Sommerfield' states).
In any case, the work by Olive and Montonen did not concern string theory, but it was a predecessor of the S-duality introduced in 1990. S-duality was later recognized as a fundamental tool for the study of the non-perturbative regime, because of the possibility of the interchange $g \longrightarrow \frac{1}{g}$.

The most of what we know today about dualities and non-perturbative objects in string theory was discovered in 1995, when Polchinski and Witten caused the so called 'second superstring revolution'.
D-branes[15] were already known before 1995, but it was only with the paper 'Dirichlet Branes and Ramond-Ramond Charges' by Polchinski [45] that their fundamental role was finally recognized.
In fact, Polchinski published in 1989 a paper with Dai and Leigh [10] where they applied 'for fun' T-duality to open string theories [16], finding that TypeIIA, TypeIIB and TypeI string theories were all dual (so discovering a little part of the web of dualities provided in 1995 by Witten) and that D-branes are a direct consequence of T-duality. Both these results were achieved by using a simple and elegant thought experiment. If we consider closed strings in a box, T-duality is not difficult to figure out, because closed strings can wound around compactified dimensions, and we have already seen its consequences talking about the minimal length. Open strings, instead, cannot wound, so which are the

[15] One usually finds the terminology 'D-branes' even if the complete terminology is 'Dp-branes' because they can have different dimensionality, so the 'p' refers to the dimensionality of the brane (while 'D' refers to Dirichlet boundary condition, for reasons we are going to see).

[16] 'for fun' because at the time the heterotic string theory, which is a theory of closed strings, was considered the string theory describing the world, taking attention away from the other string theories.

consequences of T-duality? Polchinski found that when one tries to shrink, for example, a k-dimensional box to a point, the only surviving states are those moving in the (d-k)-dimensions. This implies that compactification of k dimensions results in a new theory with a (d-k)-brane in it.
We see that before 1995, the concept of Dp-branes was already known, with different string theorists analyzing it, but their importance was not fully recognized.
Micheal Green also published in 1991 a paper untitled 'Space-time duality and Dirichlet string theory' [18] where he analysed T-duality in a theory of orientable open strings with toroidal compactification, finding an open string version of T-duality implying an equivalence of a description in terms of Neumann boundary conditions and a description in terms of Dirichlet boundary conditions in the dual torus.
I want to propose a similar but simpler example[17] showing how D-branes are related to T-duality, Neumann and Dirichlet boundary conditions. Suppose we have an open bosonic string with one compactified dimension $X^{25} \equiv X$. We can write it as:

$$X_{\mathrm{R}}(\tau-\sigma) = x_{\mathrm{R}} + \tfrac{1}{2}\ell^2 p^{25}(\tau-\sigma) + \tfrac{i\ell}{2}\sum_{n\neq 0}\tfrac{\alpha(n)}{n}\mathrm{e}^{-in(\tau-\sigma)}$$

satisfying Neumann boundary condition:

$$\partial_\sigma X(\tau,\sigma)\Big|_{\sigma=0,\pi} = 0$$

With compactification into a circle of radius R, we have $p^{25} \equiv p = k/R$, where k is the Kaluza-Kelin excitation number (the momentum is quantized in the compactified dimension). Then, we find that the T-dual $\widetilde{X}$ of X is:

$$\widetilde{X}(\tau,\sigma) = \tilde{x} + \ell^2 p^{25}\sigma + \ell\sum_{n\neq 0}\tfrac{\alpha(n)}{n}\mathrm{e}^{-in\tau}\sin n\sigma.$$

satisfying Dirichlet boundary condition:

$$\partial_\tau \tilde{X}(\tau,\sigma)\Big|_{\sigma=0,\pi} = 0$$

Dirichlet boundary conditions clearly show that the endpoints of the string cannot move in the 25th circular direction, so they can only move in the remaining 24 dimensions. This means that we can identify $\widetilde{X}$ as a hyperplane of dimensionality 24 where the string endpoints can move, thus defining a D24-brane. Summarizing, T-duality switches Neumann boundary condition to Dirichlet boundary condition, and this implies that a Dp-brane exists, where the string endpoints can move freely while being stuck in the compact direction.

[17] I take it from [38], a nice book for a simple and schematic introduction to string theory.

In 1995 Polchinski gave a clear definition of D-branes, but he understood their potential only after Witten's talk at *Strings 1995* conference. In this talk (whose content was later reported in a paper [62]), Witten showed that the 5 types of string theories (and 11-dimensional supergravity) are all connected by dualities, so they can be understood as a unique theory. Furthermore, he conjectured the existence of a 'higher-level' 11-dimensional theory reducing to the 5 string theories through specific compactification schemes and giving 11-dimensional supergravity in the low energy limit. He called this theory the 'M-Theory', but he did not give a clear definition of what this theory is.
This conjecture was anticipated by different previous results. For example, in 1995 (before the concept of M-theory was presented) Paul Townsend proved that TypeIIA superstring theory can be obtained by compactification of an 11-dimensional supermembrane theory[18]. For simplicity, I report a picture summarizing the web of dualities relating all these theories, avoiding a deeper discussion that would take us too far from the aims of this work:

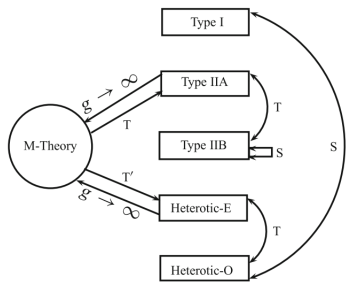

Figure 2.1: T' refers to T-duality with compactification on an interval. *Image source* [38], p.268

The 11-dimensional supergravity, supposed to be the low-energy limit of the M-theory, should be added to this web of connections, and we find it in the famous picture used by Witten to show these relations:

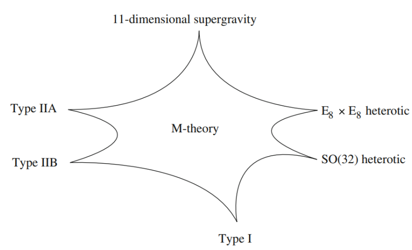

Figure 2.2: *Image source* [61], p.1128

[18]At this stage p-branes, that is 'supermembranes', were already used but they were objects with different dimensionality not yet recognized as Polchinski's Dp-branes, which have specific features

where the different low-level theories lye in different regions of the M-theory's parameter space, appearing in certain regimes. To be even more precise, we also know that TypeIIA and TypeIIB supergravity theories in 10-dimensions represent low-energy limits of respectively TypeIIA and TypeIIB superstring theories. The line of reasoning can be understood, at least partially, looking at these relations: there seemed to exist an 11-dimensional theory giving TypeII string theories (through compactification onto a circle) and Heterotic string theories (through compactification onto an interval), with TypeII string theories having TypeII (10-dimensional) supergravity theories as low-energy limit. For this reason, a 11-dimensional theory should exist, and it should also be a UV completion of 11-dimensional supergravity.
Given all these arguments, the emergence of an additional dimension in the M-Theory is not so much surprising. In fact, the 11-dimensional theory contains D2-branes that, once compactified onto a circle, appear as strings. Doing the reverse procedure, we can imagine that strings appear only because we are considering the perturbative 10-dimensional theory. Once we increase the coupling strength going beyond perturbation theory, the radius of the circle increases, so we 'discover' an additional dimension and find strings to be compactified 2D-branes. This reasoning also shows in which sense D-branes are non-perturbative objects.
As Polchinski wrote, "the perturbative expansion of the IIA amplitudes in powers of g_s is an expansion around the zero-radius limit of the KK [Kaluza-Klein] compactified d=11 theory, so the eleventh dimension is not visible at weak coupling. The d=11 theory is not a perturabtive string theory, but is one limit of a theory of quantum gravity, which includes string theories as other limits"([46], p.21).
M-theory's conjecture is surprising also because it relates theories with different spacetime dimensionalities, different compactification schemes and different kinds of D-branes (to make an example, TypeIIB superstring theory compactified on a circle was shown to be equivalent to 11-dimensional supergravity theory compactified on a torus by Schwarz). All of this underlines that M-theory should be a non-perturbative theory characterized by a rich structure invisible from a perturbative point of view.
We know that M-theory should be an 11-dimensional theory, and we can define it indirectly through its links to the other string theories, but we do not have still today a clear idea of what this theory really is. The position of M-theory in the web of theories was itself considered differently by different physicists: while Witten thought it was a theory giving the other string theories and supergravity theory in different limits, people like Greene, Morrison and Polchinski believed that it was just another theory located in a certain region of the parameter space, not a most comprehensive one.

I have already mentioned that Polchinski's paper in 1995 clarified the concept of D-branes, describing its properties and dynamics. Dp-branes where recognized as p-branes with open string endpoints attached, and he also calculated the interactions between parallel D-branes in terms of strings stretched between them. In this paper, he extended an already known duality, the *open-closed strings duality*, to D-branes dynamics. This duality, proposed in 1986 by Kawai, Lewellen, and Tye, relates open strings and closed strings amplitudes, and was extended by Polchinski in the presence of D-branes too. In particular, he showed that the interaction between parallel D-branes through stretched open strings between them admitted an equivalent description in terms of an interchange of closed strings, that is a gravitational interaction.

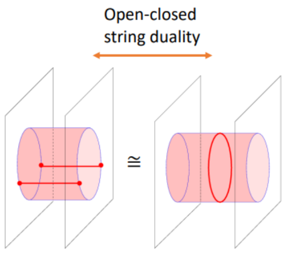

Figure 2.3: On the left, a one-loop for open strings stretched between two D-branes; they are created, propagates and annihilates, 'drawing' a circle. On the left, a tree-level diagram describing the same process in terms of an exchange of a closed string. *Image source* [12], p.12

This work was strictly linked to a key ingredient in Witten's talk. Witten argued that a certain kind of *classical* objects one finds as extremal black hole solutions in supergravity (the low energy limit of closed string theory), that is compactifications of p-brane solutions, corresponds to *quantum* BPS states. These BPS supersymmetric states were not precisely indentified, but they were supposed to exist because of the non-perturbative character of black holes. Black holes become infinitely massive at infinitely small coupling, so they could not be described by weakly coupled ordinary strings, but rather solitonic solutions should exist. These were identified with the already mentioned (and already used in quantum field theory) BPS states, carrying the so called 'Ramond-Ramond charges'. BPS states were supposed to describe extremal Reissner-Nordstrom black p-branes, but their precise nature was yet unknown..
Polchinski, listening to Witten, understood that these objects, carrying Ramond-Ramond charges, were D-branes. Being D-branes defined in the open string theory, and black p-branes in supergravity (low energy limit of the closed string theory), the open-closed string theory was fundamental to prove this relation. Polchinski calculated the mass, the number of preserved supersymmetries, and so on, finding agreement with Witten's calculations of BPS states' properties. In particular, he found that D-branes were the correct sources of Ramond-Ramond charges, and their dimensionality was equivalent to that of p-branes, so he called it Dp-branes. Finally, D-branes were recognized as BPS states.
The open-closed string duality, on the other hand, showed that "the open strings at-

tached to the D-branes can be described in a dual representation containing only closed strings. Under this duality, the D-branes, on which the open strings end, are replaced by a curved background geometry - often a higher-dimensional black hole" [12].
The importance of these results was not immediately understood by Polchinski, but others recognized it and started to work on D-branes and their applications. A serious discussion of all these results would take us too far, so I will limit to mention the main discoveries.
In 1996, Witten generalized Polchinski's work analyzing the interactions between N intersecting D-branes. To be more precise, he studied bound states of D-branes, and found that their low energy description is a theory with U(N) gauge symmetry group. Thus, Witten related the world-volume theory of D-branes and a gauge theory, a result that will be used by Maldacena for his conjecture. Polchinski's and Witten's works were really important for two main reasons. First, as I am going to discuss briefly, they allowed for the calculation of the Bekenstein entropy of black holes, by constructing them as a stack of coincident D-branes. Second, the relation between the world-volume theory of a D-branes bound state and a gauge theory with U(N) group allowed for a connection between string theory and gauge theories, culminating in the Maldacena conjecture. Also, new tools were available for a better understanding of high energy string theory dynamics. In fact, "Michael Douglas, Daniel Kabat, Philippe Pouliot and Stephen Shenker (1997) exhibited, in August of 1996 [...] the consequences of open-closed duality for the relation between the world-volume theory of the D-branes and supergravity. Namely, they argued that short-distance phenomena in open string theory are described by the infrared behaviour of the world-volume theory. As Witten had shown, the world volume theory was an ordinary quantum field theory. The novelty lay in the evidence that this low-energy effective action could describe short-distance phenomena in string theory and M theory: such as scattering between D-branes at Planckian energies. Yet, because D-branes are fine-grained non-perturbative objects, the infrared dynamics of the associated world-volume theory turned out to capture well the short-distance behaviour of string theory"([12], p.14).

We know almost nothing of this theory, but all this developments clearly showed that, quoting a popular sentence by Robbert Dijkgraff, "string theory is not a theory of strings". In fact, new objects of different dimensionalities appeared on the scene in the 1990s, namely the Dp-branes already discussed, showing that strings are just one kind of objects we can find in the rich structure of M-theory. Their important role and their applications suggested that, likely, Dp-branes degrees of freedom are the fundamental one for a correct formulation of string theory, even if such a formulation has not yet been found.
A different approach to M-Theory was developed by Susskind, Banks, Shenker and Fischler [4]. Rather than defining M-Theory starting from an 11-dimensional theory able to recover the different types of string theories via appropriate reductions, they tried

to define it as an 11-dimensional theory emerging from the strong coupling limit of a 10-dimensional string theory. This approach, called the 'Matrix model', "leads to a definition of M-Theory as the eleven-dimensional theory on a flat ('decompactified'), infinite background space" ([49], p.220). Even if this approach allows to investigate the non-perturbative regime of a string theory, the problem of having a fixed background remains.
The meaning of spacetime undergone a great change under all these developments, and it became increasingly clear that a background independent formulation was needed, and that the concept of spacetime - and especialy of spacetime dimensions - needs to be drastically modified.

2.4 Phase 4 [1995-today]

After the 'second revolution of superstrings', the power of Dp-brane machinery was used to analyze the black holes physics. In fact, black holes physics is the natural subject to be analyzed by a candidate as a quantum theory of gravity. Dp-branes and dualities, at this points, seemed to offer powerful tools in order to investigate such scenarios.

I have just mentioned that string theory contains solitonic solutions (BPS states) corresponding to extremal black hole solutions, and identifiable with Dp-branes[19]. One of the greatest results of string theory is without any doubt the calculation of the Bekenstein entropy of a certain kind of black holes. Entropy is related to the microstates of a system, so we need a method to count these microstates in order to get the entropy. In 1996, Strominger and Vafa[20] successfully calculated the Bekenstein entropy of 5-dimensional extremal black holes [55]. These black holes are not realistic models, being 5-dimensional and extremal, meaning that a zero-temperature approximation was used, but in any case it was the first time that the entropy of a black hole was correctly calculated. In order to find this result, Strominger and Vafa identified the microstates with BPS states, so their aim was to count the degeneracy of these BPS bound states. Knowing from the the previous work by Polchinski that BPS states are generated by D-branes, the problem was shifted to count the bound states of Dp-branes.
After this result, other string theorists tried to apply the same methods and generalize this result. For example, Curtis Callan and Juan Maldacena generalized the above result to near-extremal 5-dimensional Reissner-Nordstrom black holes [6].

[19] The gravitational force they generate can be visualized as due to open strings merging together to form a closed string which can leave the D-brane and representing the gravitational radiation.

[20] For a very nice review of Strominger and Vafa's calculation see [12], from which also the previous part concerning Polchinski's and Witten's results is taken.

These results were crucial not only because it was the first time that the Bekenstein entropy was calculated, but also because a new possibility for the resolution of the Hawking paradox opened up. Stephen Hawking discovered that black holes are thermal objects, emanating a thermal (casual) radiation. This means that if one could throw into a black hole pure quantum states, it would radiate them out as mixed (thermal) states, this resulting in a violation of unitarity and a loss of information. This is known as he 'black hole information paradox', and Hawking argued that black holes evaporation violates quantum mechanics.

As Rickels remembers, "Strominger and Vafa pointed out the potential relevance of their work to the black hole information paradox. They suggest that D-brane machinery might be used to directly compute the low-energy scattering of quanta by an (extremal) black hole, to check for unitarity or its violation. They note that S-type dualities could be utilised to make this a possibility, turning a strongly coupled problem to a weakly coupled one. Studying the Hawking radiation in terms of open string excitations, one finds that unitarity is indeed preserved"([49], p.223).

The resolution of the paradox was not conclusive, being all these results obtained for a very special kind of black holes, but another important contribution was given by another crucial 'discovery'.

In 1997, Juan Maldacena published a paper [36] that started something like a 'third superstring revolution'. In this paper, he presented a conjecture now known as 'Maldacena conjecture' or AdS-CFT correspondence (where AdS stands for 'Anti-deSitter and CFT for Conformal Field Theory). Maldacena conjectured that a string theory (which necessarily contains gravity) on AdS_5 is dual to a conformal field theory (a quantum gauge theory without gravity) defined on the conformal boundary ∂AdS_5 [21]. This new duality, based on the already mentioned open-closed strings duality[22] (and the relation between D-branes and black holes), states the equivalence of a quantum theory of gravity in the 'bulk' and a gauge theory without gravity on the boundary for a certain space with negative cosmological constant [23]. This is not a realistic model, because our Universe is supposed to have a negative cosmological constant, but it suggests many deep concepts and applications. First of all, it was used to argue once again that the black hole's evaporation violates quantum mechanics: if a theory with gravity is dual to a gauge theory without gravity, it means that black holes can be described by the ordinary laws of quantum mechanics as well. This conjecture was indeed used in the context of

[21]The AdS-CFT correspondence is strictly connected to the holographic principle, because it states that a certain theory defined on the boundary of a certain space is enough to describe the physics inside that space (in the bulk).

[22]I remind that D-branes at weak coupling represent a gauge theory, that is an open string theory, but the open-closed strings duality states that an equivalent description might be done in terms of closed strings, describing the gravitational interaction.

[23]It is quite evident the relation with the holographic principle, a principle more and more important in contemporary theoretical physics.

black holes, showing that a black hole is described in the dual theory by a plasma of hot quarks. The evaporation of a black hole is then described by a classical evaporation of a thermal system, and since in the latter case there is no loss of information, the same would be in the former case.
The paper where this new duality was proposed is one of the most cited papers in the history of physics, and had an incredible impact in the research field. One of the most important consequences of this conjecture was that it opened up the possibility of a non-perturbative definition of string theory. Gauge theories are well known, and if really exists a strong-weak coupling duality between string theory and gauge theories, the latter may help to shred light on what string theory really is.

This was the situation at the end of the '90s, before a further led to a very different scenario.
After the (first) 'ground state explosion', the main goal was to find some mechanism or principle to get the vacuum corresponding to our Universe among all the others. The situation was not very clear but in any case not dramatic: the estimated number of ground states was of the order of thousands, so the hope to classify them and to find the correct one was still alive.
A complication arose when the stability of the compactified dimensions was considered. It was found that without any mechanism constraining them, these dimensions would decompactify because of the instability of Kahler and complex structure moduli, determining the size and shape of Calabi-Yau manifolds. This problem was solved via flux compactification: a quantized flux and D-branes [24] were used to 'freeze' such a moduli, making the compactified dimensions stable [27]. The problem is that this new kind of compactification scheme results in an additional 'ground state explosions', bringing to the 'Plurality of Type 3' we mentioned above. This time, however, the explosion was much more dramatic: an estimated number of $\sim 10^{500}$ (or even more) possible ground states is supposed to exist, constituting the so called string 'Landscape' [25].
The Landscape divided the string theorists' community into two main factions: those, such as David Gross, still trying to find some principle enabling us to pin down one single vacuum corresponding to our Universe, and those believing that the ground states of the Landscape correspond to really existing Universes.
This second approach is the most followed, and its main supporter is Leonard Susskind, who also spread the concept of 'Megaverse' (the set of universes described by the Landscape) in the common imaginary. The concept of Landscape is strictly linked to the anthropic principle, without which it would be useless.

[24] again, we notice the fundamental role of D-branes machinery.

[25] It is worth mentioning that the string Landscape also contains the ground states of the F-Theory, a 12-dimensional theory conceptually similar to the M-Theory that was developed by Cumrun Vafa. As it is for the M-Theory, the extradimensions can be compactified to get realistic models, and this procedure leads to a huge number of possible configurations as well.

The line of reasoning is based on the observation that we live in a very special Universe whose parameters seem to be incredibly fine-tuned in order to allow for our existence; an even very small change of a certain parameter and life in our Universe would be impossible. This looks like a 'miracle' or something made by a Superior Intelligence if we suppose our Universe is the only possible one. String theory, in any case, shows the existence of $\sim 10^{500}$ different consistent ground states, and if we assume they are really existing Universes, the lucky conditions we find in our Universe have a simple explanation: being such a huge number of Universes, all with different values of the physical parameters, the existence of a Universe with parameter values life-friendly is not anymore surprising, but a statistical issue. Using this line of reasoning it seems to be possible to solve different problems of contemporary physics, such as the incredible fine-tuning of the cosmological constant. In any case, as I will discuss in the next chapter, things are not that simple.
A strictly related issue to the string Landscape is the theory of eternal inflation, belonging to inflationary cosmology.
The concept of inflation was developed independently by Aleksej Starobinskij and Alan Guth in the '80s, and it provides a cosmological model assuming that our Universe went through a period of exponential growth. Inflation is able to explain different features of our Universe, the main one being the isotropy of the Cosmic Radiowave Background, whose homogeneity between regions of the Universe supposed to be causally disconnected was not yet explained. Inflation is accepted by the most of scientists, but not by the entire scientific community.
A particular type[26] of inflationary model is the one based on the concept of 'eternal inflation', containing the so called 'Bubble Theory' proposed by Andrej Linde. The eternal inflation model proposes a background Universe undergoing an eternal and exponential expansion; inside it, quantum fluctuations of the vacuum can make 'bubbles of alternative vacua' to appear and, in certain cases, these alternative vacua can expand and generate other Universes, in a process resembling the Big Bang. This process can be visualized referring to the concept of Landscape. This term refers to the relative minima and relative maxima we can find in the energy potential, which can be visualized as 'valleys' and 'peaks'. Here, the relative minima represent 'false vacua', that is metastable vacua, describing a certain region of the multiverse. Being metastable, quantum fluctuations might make these vacua to decay into more stable vacua through tunneling effect. This process corresponds to the generation of 'bubbles of alternative vacua' in the theory of eternal inflation.
In any case, details are not so much important for the following considerations, but the only thing of interest is that eternal inflation predicts the existence of a Multiverse, in agreement with the string Landscape.
I will come back to this topics and its consequences in the next sections. For this mo-

[26] Even if many physicists think that inflation necessarily implies eternal inflation, things are not that simple, as Smolin argues in [52], chapter 11, so I prefer to discuss them separately.

ment, I only want to stress that these concepts led to a strong debate and division inside the scientific community, as it is evident from the publication of by now popular books, such as the already mentioned books by Peter Woit, Lee Smolin, Leonard Susskind and Brian Greene. Many scientists argue that anthropic reasoning is not science, while its supporters claim that scientific methodology might need to be revisited. We see that this debate is more about scientific methodology than physics, and so it represent a suitable place to apply the philosophical tools we have already discussed. For this reason, the main goal in the next sections will be to achieve a better understanding of this debate and frame it in the correct philosophical scenario.
Another important recent development is represented by the so called 'Swampland programme', first proposed by Cumrun Vafa and related to more general issues about quantum gravity (QG). Researchers in this programme try to identify those features that a consistent theory of quantum gravity does *not* have, this also helping to figure out how a consistent one should look like.
The line of reasoning is that not all Effective Field Theories (EFTs) are consistent with a QG UV-completion, and being easier to understand if a theory does not admit such a completion the programme's heuristic points to discard these theories rather than isolate them. The EFTs consistent with QG are considered to belong to the Landscape, a QG Landscape originally different from the string Landscape, while the others are considered to belong to the so called Swampland. In these terms, this project tries to conjecture some features that are common to the theories in the Swampland; doing so, one might appraise if a certain theory is a possible candidate for a quantum theory of gravity or not.

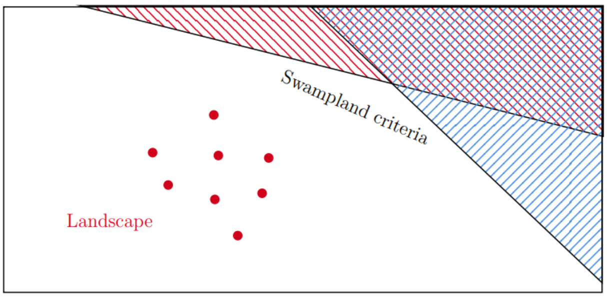

Figure 2.4: Swampland conjectures enable to exclude EFTs not compatible with a QG UV-completion, isolating the Landscape. *Image source* [2], p.147

This goal, in addition to isolate indirectly the Landscape, might also suggest which features a theory of quantum gravity should have. Obviously, if the number of EFTs in the Landscape were infinite this programme would be quite useless, because of the difficulty to recognize some patterns, so the research programme is based on the assumption that there is a finite number of EFTs consistent with a QG UV-completion. Also, being

string theory the most developed candidate as a quantum theory of gravity, it represents the main tool that is used in order to develop and propose those conjectures. If string theory was only one of many possible QG theories, all the information we may get would not be general properties of QG theories, so another strong assumption the programme is based on is that string theory is the only quantum theory of gravity. This strong assumption, identifying the Landscape with the string Landscape, is called the 'String lamppost principle', and it ensures that by using string theory in order to find general features of QG we are not getting completely wrong. In any case, in the Swampland programme not only string theory but also other general considerations and tools such as black hole physics are used in order to understand if a certain theory belongs to the Swampland or not. For this reason, this line of proceeding may help to get information about theories of quantum gravity in general.
For our purposes, it is important to understand that this programme does not belong to the Landscape programme we have discussed above. There, it was assumed that many different universes exist, each one described by a different theory of QG (that is, by a different string theory). The Swampland programme, instead, belong to the line of research still believing that a unique theory exists. Quoting from [2]:

> The core of the Swampland programme is the uniqueness of string theory. In string theory, as we increase the cutoff, the landscape of theories that seemed to be disconnected, become connected. It is believed that increasing cut-off high enough would lead to one single theory with a single connected moduli space. Given that different Calabi–Yau manifolds lead to different EFTs, this implies that all different Calabi–Yau manifolds must be transformable to each other using specific geometric transitions. In fact, this statement is a well motivated math conjecture often known as Reid's fantasy.

Summarizing, the Swampland programme uses string theory and other tools such as black hole physics and general arguments in order to conjecture some general features that a consistent quantum theory of gravity should not have, also trying to get a hint about the properties that a well defined theory of quantum gravity may have. These principles might represent a good guide in theory building, or might be used to to get a hint about the structure of M-Theory or something else. Maybe, they will be discovered to unify into a single general principle allowing us to find the much desired theory of quantum gravity.
To conclude this topic, I would like to point out that some of these principles are based on considerations about naturalness, that is some general features of QG theories are proposed in order to explain some issues that appear as problematic from a natural point of view. In any case, I will come back to this topic in the next session.

To conclude this review of the main recent approaches in string theory, it is worth mentioning the so called 'brane cosmology'. This model is based on an idea that we can

already find in [25], where Polchinski speculates that our Universe might be a 4-brane embedded in a higher dimensional spacetime. If the extradimensions are extanded, we can visualize our Universe as a 4-brane moving through the 'bulk'. One of its nice features is that it may explain why the gravitational force is so much weaker respect to the other forces at our energy scale. In fact, closed strings (representing gravitons) would not be stuck to our world-brane, but they would be free of moving in the extradimensions as well. This would imply that the gravitational force is leaking its power in the extradimensions, differently from the other forces described by open strings stuck to the world-brane. At short distances, instead, the leak would be very small, so the gravitational force would be very much stronger.
In any case, this model does not represent a much followed line of research and it is not related to methodological debates, so I will not discuss it further.

At this stage this brief review of string theory's history might be considered almost complete, because no other important developments have occurred after the concept of Landscape entered the scene. In the next chapters I am going to propose a philosophical analysis of this history, further investigating those issues that might be recognized as the most important from a methodological point of view.

Chapter 3

String theory from a Lakatosian perspective

Having summarized the history of string theory research programme, we now have all the tools to appraise it in the light of the methodology of scientific research programmes. I am going to make the following steps: recognize the hard core and the protective belt of auxiliary hypothesis, identify the positive and negative heuristics, look at progressive and degenerative problemshifts; then, I give a rational reconstruction of the internal history of string theory research program, and finally analyze external history in order to account for apparently irrational developments.
I would underline that, when not otherwise specified, with 'string theory' I mean the whole web of theories already completed in the late '90s, that is the the general framework of M-Theory. With the terminology 'string theory research programme' I refer to the research programme which started with the exaptation already mentioned and split into two research lines that I am going to discuss individually: the research line still searching for a unique M-Theory, and the one based on the real existence of the Landscape.

3.1 Hard core

The pre-history of string theory is not considered as belonging to the string theory research programme at all, but it is nevertheless important for a rational reconstruction of the research programme itself.
First of all, the *hard core* of the research programme has to be identified. In general, it is not an untouchable set of assumptions, but it may change with time in the light of new discoveries, because some necessary conditions might be unknown at first and discovered following the positive heuristic. Obviously, the way this change happens is fundamental for a correct appraisal of the research programme.
Early after the exaptation period, little was known about string theory. The hard core

can be identified with one single assumption: particles are actually manifestations of one-dimensional objects' vibrational modes.
Then, when supersymmetry was found to remove the tachyon, giving a consistent theory, it was added to the hard core. This claim might be considered wrong, because supersymmetry might not be strictly necessary; indeed there have been attempts to formulate string theories without supersymmetry[1]. We do not know precisely what string theory is, so we do not know if supersymmetry is necessary or not for its formulation. In any case, string theory as we understand it today contains supersymmetry, so I think it should be added to the hard core. I would like also to remind that many calculations are performed assuming supersymmetry[2], this showing that it plays a fundamental role in the applications and is universally assumed as a fundamental feature of string theory. In particular, low-energy supersymmetry is usually considered. To give an example, in the classical textbook 'Superstring theory, Vol.2' by Green, Schwarz and Witten, the authors explicitly give motivations to assume low-energy supersymmetry ([20], p.412). Among them, there is the possible solution of the already mentioned 'hierarchy problem'. Here, they states: "We dot know how to solve this problem, but one necessary ingredient is presumably that the ordinary $SU(2) \times U(1)$ Higgs doublet must remain massless at the compactification scale and indeed to within extraordinary precision. [...] Of course, if we suppose that the $SU(2) \times U(1)$ Higgs doublet would be exactly massless in the limit of unbroken supersymmetry, then the tiny but nonzero scale of $SU(2) \times U(1)$ breaking in the real world must be related to a small scale of supersymmetry breaking". I will come back to this topic at the end of this chapter.
Also, an implicit assumption in the hard core concerns the validity of quantum mechanics. This is not a trivial assumption because, as I have already mentioned, black holes evaporation seemed to violate quantum mechanics and, for example, it was considered by Hawking as invalid for such a case. Also, quantum mechanics is at the base of dualities: Polchinski's thought experiment demonstrating T-duality was based on a paradigmatic reasoning in the framework of quantum mechanics, that is the 'put in a box' experiment with uncertainty principle. For these and other reasons, quantum mechanics should be considered as belonging to the hard core of string theory research programme. In any case, this fact may change with future developments, in fact a "new version of Heisenberg's principle [involving] some non-commutativity where it does not usually arise [...] may be the key to the thinning of the degrees of freedom that is needed to describe string theory correctly" as it was speculated by Joseph Atick and Edward Witten in 1988 ([3], p.314). String theorists assume the validity of quantum mechanics but also recognize that likely it should be replaced with another fundamental theory in order to give a clear description of string theory. I have already mentioned a certain kind of 'quantum

[1] As an example, see [13]. Also, in [58] p. 357-358 Susskind wonders if supersymmetry is more or less probable in the Landscape.

[2] For example, the Bekenestein entropy was calculated by Strominger and Vafa in the maximal supersymmetric case.

geometry' that may be part of this future developments. As I pointed out in the introduction, quantum mechanics is an incredibly successful framework but many scientists feel unsatisfied of its unintelligible character. In any case, if quantum mechanics will be replaced, it is expected that the new theory would reduce to quantum mechanics in a certain limit, and in general it should account for all its achievements, as it is demanded by Lakatosian methodology. Quantum mechanics belonging to the hard core means that every evidence contradicting it will be addressed to the protective belt, and not that it cannot be replaced by a more general framework.
The same cannot be said for General Relativity. Its validity does not represent a fundamental assumption for string theory, but it is rather derived (and also corrected at high energies) from string theory. If General Relativity turns out to be wrong, string theory would not suffer of any counterevidence.
The three elements I have included in the hard core so far are the same we can find in the analysis given by Lars-Goran Johansson and Keizo Matsubara in the popular paper 'String Theory and General Methodology; a Reciprocal Evaluation' [26]. Now, I would like to add another element to the hard core, that is the assumption that string theory represents a unifying theory with no free parameters. In fact, as we have seen talking about the pre-history of string theory, it was the S-matrix theory which lead to dual models and, then, to hadronic string theory, so the last was embedded in the same bootstrap philosophy. This philosophy, I remind, demanded the absence of arbitrary parameters: the constant values should be an output, and not an input. String theory, after the exaptation process, was still following the same philosophy; in fact, in [28], David Kaiser talks about superstring theory as a "sign of the S-matrix programme's afterlife". We may say that string theory represented the highest expression of the bootstrap philosophy: consistency conditions gave as outcome the dimensions of spacetime too. The absence of free parameters and arbitrary structures was also a great virtue giving strong support to the research programme. In any case, this fact will be better understood when talking about the lost of uniqueness.
Summarizing, the hard core was constituted by four elements:

- the fundamental objects are one-dimensional strings[3] rather than point-particles;
- supersymmetry exists;
- quantum mechanics is valid;
- string theory has no arbitrary parameters.

[3]Obviously, when Dp-branes were discovered, they were added to the list.

3.2 Protective belt

Now, the *protective belt of auxiliary hypothesis* has to be identified. It contains all those hypothesis against which counterevidence is redirected in order to protect the hard core. The most evident one is without any doubt the hypothesis that extradimensions are so small that they cannot be detected using the available technology.
Another hypothesis, namely a set of conjectures, is the validity of the dualities connecting the different types of string theories, so that the 'Plurality of Type 1' is not really a plurality, but the different theories can be viewed as different formulations of one single theory. The reasoning behind this assumptions lies in the fact that certain calculations are performed for specific string theories. For example, the black holes entropy is calculated in the context of the type IIA supergravity, the low energy limit of Type IIA string theory, which is non-chiral, so it would not represent a realistic theory of our Universe. If the calculation is performed in the context of a non-realistic theory, one may question the result loses credibility, even if it is the correct one. But if dualities are assumed as valid, this line of reasoning would be mistaken, because Type IIA string theory would be connected to the phenomenologically promising heterotic string theory. In general, one calculation made in a specific string theory would be considered as valid for the whole set of theories. Dualities are of fundamental importance for another reason. Coming back to the example of the black holes entropy, this calculation is performed using closed-open strings duality. In general, assuming the validity of these conjectured dualities is equivalent to secure the validity of mathematical procedures, that is to give a solid foundation of the obtained results.
One may ask why dualities are not included in the hard core. The reason is that they are not strictly necessary. Before the mid '90s heterotic string theory was considered the string theory describing our Universe, because the other types were not phenomenologically promising (for example, TypeIIA theory is not chiral). The different string theories were 'disconnected', and even if their presence asked for some kind of explanation, there was no reason to abandon the whole framework just because some of them did not represent realistic models; it was enough to have one theory among the others able to recover the right phenomenology. When dualities were discovered and well understood, they connected the different string theories and provided powerful tools for applications and for a better understanding of the entire framework. In any case, there is no reason to include this concept in the hard core. The existence of different string theories does not represent in itself any inconsistency needed to be fixed adding duality conjectures.Their role appears as crucial only in the more general context of M-Theory. The conjecture of its existence relies on this web of dualities, and in case they were found wrong, the M-Theory conjecture would be rejected. If we identify string theory with M-Theory, we should add dualities to the hard core, because they would be of crucial importance for the entire research programme. While this step would be rational, I consciously decided

not to do this. As we have already discussed, M-Theory's role is ambiguous even among the string theory community. Polchinski, Greene and Morrison thought of it as another theory in a certain region of the parameter space, and not as a higher-level theory. In this different viewpoint, a disprove of dualities would not invalidate the whole project, but it would leave the theories divided, with only some (such as heterotic string theory and M-Theory) of them again considered to be the correct one, and it would invalidate only those calculations and concepts based on dualities. For these reasons I did not include these conjectures in the hard core, but if a string theorists believes that the high-level M-Theory is the only way to achieve the research programme's goals, he would be implicitly adding them to the hard core.
These two elements are accounted also by Johansson and Matsubara, but I think a further conjecture should be added, that is the finiteness of the perturbation expansion. This issue is reported by Lee Smolin in his already mentioned popular book [52]. Smolin states that many string theorists though that the perturbative expansion was demonstrated to be finite, but it was not true (and it is not still true) at all. In particular, it was rigorously proved only up to the third order of the perturbation expansion. Even if the demonstration for the third order was achieved only in 2001, string theory was considered to be finite already in 1991. Dean Rickles [49] says that "there was word circulating that Mandelstam had discovered a proof in the mid-to late 1980s. He finally supplied an explicit proof in 1991 (supplying formulas for the n-loop amplitude, for Bose strings and superstrings) that 'could be put on a computer (but may require an unreasonable amount of computer time)' " . Smolin decided to make e deeper investigation, asking to many string theorists and mathematicians. He found that many string theorists thought finiteness was definitely proved by Mandelstam, or that, even if a conclusive proof had not yet been found, it was evident anyway, not needing a further investigation. On the contrary, he says that some mathematicians thought that Mandelstam proof was not complete, and that maybe string theory was not finite at all. Finally, he reports [52] that Carlo Rovelli asked to Mandelstam himself, and he "explained to Rovelli that what he proved was that a certain kind of infinite term does not appear nowhere in the theory. But in reality, added, he did not demonstrated the finiteness of the theory itself, because other kinds of infinities may appear. So far terms of this kind have never appeared, , but it is also true that no one has ever proved that this cannot happen". This fact underlines an important aspect related to the awareness of scientists of the situation of the scientific research programme they are working for, suggesting further considerations about scientific methodology. For this moment, the aspect of our interest is in the absence of a rigorous mathematical proof, so I think it is correct to include among the protective belt the conjecture of the perturbation expansion's finiteness, whether string theorists are aware of it or not.

Johansson and Matsubara include among the protective belt the existence of the Megaverse, writing: "explain the value of the constants of nature assuming a landscape of universes". I do not agree with this assumption. Adding this element to the protective belt means identifying string theory research programme with the Landscape programme only, but I have already discussed that it is only one among different approaches, even if it is the most followed one. We cannot neglect the existence of those lines of research still searching for uniqueness; it would be not only unfair, but also inaccurate. This fact is strictly related to the reasoning which led me to add the existence of unique theory with no arbitrary parameters to the hard core. In order to understand this point, we need to come back to the first and second 'ground state explosions'. When compactification was found to imply the lost of uniqueness because of the arbitrariness of shape and size of compact manifolds, many string theorists were still hoping that uniqueness would be restored. This fact has an evident epistemological interpretation in the light of the methodology of scientific research programmes: many string theorists, facing a theoretical counterevidence, added to the protective belt the auxiliary hypothesis that a certain yet unknown principle or mechanism able to restore uniqueness would exist. This line of reasoning also strengthen the argumentation justifying the additional element I added to the hard core. In fact, string theorists would have had no reason to conjecture the existence of such an unknown principle/law/mechanism if the uniqueness and absence of free parameters would not be present in their research programme's hard core. Auxiliary hypothesis are made for the very purpose of avoiding counter-evidence to be directed to the hard core, and for this reason I think such a conjecture deserves to be added to string theory's hard core.
After the first 'ground state explosion', even if there were already attempts to use the anthropic principle to make predictions, an organic research programme based on the real existence of the Megaverse was absent. It appeared when compactification via fluxes caused a second and much more catastrophic 'ground state explosion'. At this point, while 'uniqueness researchers' still believed in the possibility to recover uniqueness, many string theorists decided to assume the real existence of a Megaverse corresponding to the ground states in the Landscape. They accepted the existence of a huge number of universes described by different string theories with different parameter values, abandoning the assumption of a single theory giving these values as an output. I would like to underline that I am using the terms 'lost of uniqueness' and 'arbitrariness of parameters' as equivalent. In fact, accepting the existence of many universes characterized by different parameters values means accepting that many string theories consistent with different parameters values are possible, that is not a single consistent string theory exists. For this reason, the assumption of no free parameters can be viewed as an assumption about uniqueness.
In any case, I think that the account given by Johansson and Matsubara cannot be considered correct for the Landscape as well. Landscape researchers accepted that uniqueness was definitively lost, assuming the existence of a multiverse in order to explain the

features of our Universe. Now, there is a subtle epistemological issue. One may think that the only difference between these two main branches of string theory research programmes differ only for an auxiliary hypothesis in the protective belt, but this is not the case. In fact, the Landscape hit the hard core, because it was accepted that string theory would not be absent of arbitrary parameters; on the contrary, the existence of many different ground states with different constants values was taken to be a virtue, as I am going to discuss later. The existence of the Landscape, being opposite to the assumption of no free parameters and also constituting the basic assumption in the Landscape programme, should be considered the basic element in the hard core of a different research programme. For this reason I am not going to consider the Landscape as an element in the protective belt of the 'old' string theory research programme, but as an element in the hard core of a different research programme[4].

Finally, I would like to remind that the Landscape programme is linked and supported by the theory of eternal inflation. They represent different research programmes, and the corroboration or disprove of eternal inflation would not be directly addressed to the Landscape programme (even if it would have a strong impact), so we can consider almost all their features (the hard core, protective belt etc.) as separated. There is only one element that I think it is worth to render explicit. Both eternal inflation and the Landscape predict the existence of many other Universes, so it is natural to wonder why we do not see them. The reason is that they are located beyond the 'cosmic horizon', an 'event horizon' due to the expansion of the Universe. I think that this assumption, common to both research programmes and strictly resembling the one relative to the size of the compact dimensions, should be added to the protective belt.

Summarizing, I present two different hard cores and protective belts corresponding to two different research programmes which I would like to call 'standard string theory research programme' and 'Landscape research programme'.

For the first one we have:

- Hard core:
  - the fundamental objects are one-dimensional strings rather than point-particles;
  - supersymmetry exists;
  - quantum mechanics is valid;
  - string theory has no arbitrary parameters.

[4]A similar situation might occur if supersymmetry will be disproved. In such a case, supersymmetry might be removed from the hard core, and string theories without supersymmetry might be developed (or, maybe, string theory would be abandoned).

- Protective belt:
  - compactified dimensions are too small to be observed with contemporary technology;
  - validity of dualities;
  - finiteness of the perturbation expansion;
  - existence of a principle/law/mechanism able to restore uniqueness.

For the second:

- Hard core:
  - the fundamental objects are one-dimensional strings rather than point-particles;
  - supersymmetry exists;
  - quantum mechanics is valid;
  - the Landscape's ground states correspond to really existent universes constituting the 'Megaverse' (or 'Multiverse');
- Protective belt:
  - compactified dimensions are too small to be observed with contemporary technology;
  - Validity of dualities;
  - finiteness of the perturbation expansion;
  - the other universes are located beyond the cosmic horizon.

3.3 Negative and positive heuristics

Now, Lakatos methodology provides us with other important concepts: the *negative heuristic* and *positive heuristic*.
The negative heuristic has a simple role: it does not allow any counterevidence to be directed against the hard core, but instead it redirects counterevidences to the protective belt. I would like to remind that the negative heuristic represents an implicit methodology we can find everywhere in history of science: theories develop in an 'ocean of

anomalies', and if scientists allow anomalies to be directed against the hard core, no theory could be developed. All theories are born falsified, but scientists hope to solve the anomalies at some time in the future, following the positive heuristic.
The positive heuristic represents a guide in the theory building process. All theories follow a standard path: their development is characterized by increasingly refined models, and at each step some anomalies are expected to be solved. In the context of string theory, we can identify the positive heuristic and these subsequent models looking at its history. When the theoretical exaptation was achieved, the positive heuristic was clear. String theorists had a free theory defined in d=26, with a tachyon in the spectrum, not including fermions, representing a potential quantum theory of all forces, and without free parameters.
The first steps were clear, and underlines how scientists work to face anomalies. First of all, an interaction picture was developed, being initially a theory of free relativistic strings. Then, fermions were added, leading to supersymmetry. Supersymmetry, as a consequence, eliminated the tachyon in the spectrum. This step shows clearly the methodology followed by scientists: usually anomalies are not faced directly, but the refinement process is expected to solve them automatically. Anomalies are seen as consequences of the theory being still incomplete. Supersymmetry also reduced the number of dimensions from 26 to 10, this suggesting that maybe dimensional reduction would arise automatically during the theory building process. In any case, compactification of extra dimensions was already investigated in the early period of hadronic string theory, so getting a 4 dimensional theory was already part of the positive heuristic. A chiral anomaly was found, and this technical problem was resolved leading to heterotic string theories. Again, on the to-do list there was the recovering of the Standard Model, and heterotic string theories were very promising for this aim. At this point, huge efforts were made trying to recover low energy phenomenology by using the $E_8 \otimes E_8$ gauge group and compactification techniques. At this step a big problem started to arise, that is the lost of uniqueness due to compactification. Uniqueness was a major motivation for research in string theory, as it was the absence of free parameters. In any case, dualities were discovered, pointing to a reduction of the 'Plurality of Type1'. Scientists were still hoping that other discoveries may help to recover uniqueness, so this anomaly was expected to be solved following the positive heuristic. Dualities led to the (re)discovery of Dp-branes and to the M-Theory conjecture. Now, being string theory a candidate as a quantum theory of gravity, we can obviously find in the to-do list the study of physical scenarios where quantum gravity is necessary, such as black holes physics. Found non-perturbative objects, the theory was ready to approach such problems, and achieved some success[5].

[5] The Bekenstein formula was recovered for special kinds of black holes, that is for maximal supersymmetric five dimensional near-extremal black holes. Here we can identify another heuristic at a sub-level: calculations are usually performed in the simplest cases and then extended to more complex one, following a process of refinement towards more realistic models. In fact, at first extremal five dimensional black hole were studied, and then near-extremal five- and four-dimensional black holes. As I am going

Also, the M-Theory conjecture suggested a way to recover uniqueness. At this stage the positive heuristic was clear: try to give a non-perturbative formulation of string theory and find the M-Theory (for those scientists viewing M-Theory as a higher-level theory, these goals were strictly connected). Also, Dp-branes deeply modified the whole understanding of string theory: string theory was not anymore a theory of strings alone. This fact stressed the importance of achieving a non-perturbative formulation of string theory: some profound principles were expected to exist that should allow for a clear formulation and for the selection of the correct ground state, so the research programme pointed to this direction. Again, similarly to the discovery of the chiral anomaly, a problem was found about the stability of the compact dimensions. This time, however, the solution (compactification via fluxes) led to another 'problem': the Landscape. At this point the research programme divided: some string theorists saw the Landscape as a virtue, while others were still searching for uniqueness. Also, the heuristics were different: Landscape supporters tried (and still try) to find the ground state corresponding to the vacuum of our Universe, while uniqueness supporters were still searching for some principle selecting the correct ground state. These scopes might seem similar but they are very different: the former accepted the status of the theory as conclusive, thinking that finding the above mentioned ground state was the last step, while the latter were searching for some principle to get a complete formulation of the same theory, and selecting one ground state; in brief, we can say that the former were searching for a ground state, while the latter were searching for a theory.

The last 'discovery', the Maldacena conjecture, suggested novel paths for a non-perturbative formulation of string theory, not yet achieved.

This brief account makes clear how the theory was developed during the years and how anomalies were faced, clearing the role of the positive heuristic. Also, this underlines how a research programme may change (for the better or for the worse) during its development because of unexpected problems or discoveries: the chiral anomaly, whose presence was not known from the start, led to a very promising gauge group; dualities and D-branes provided powerful tools and pointed to a unification of the five different string theories; the lost of uniqueness because of compactification schemes had not dramatic consequences, but the same lost because of stabilization via fluxes led to an incredible change in the research programme.

String theory research programme, before the apparition of the Landscape, was characterized by a clear heuristic; after this event, as I argued above, the research programme divided, and different heuristics were developed.

Wanting to write down the to-do list in the positive heuristic, I would like to add also the 'Swampland programme'. Even if this programme is based on precise assumptions,

to argue, while it is true that this result does not represent strong evidence in favour of string theory, one cannot expect that results are immediately obtained for the realistic case. Phenomenology follows a process of the same kind of theory building, and in both cases time has an important part.

it clearly belongs to the 'standard string theory research programme', which one might consider to include all those approaches rejecting the existence of a Megaverse. Furthermore, it makes very clear the difference between 'standard' and 'landscape' research programmes. In fact, in this programme the uniqueness of string theory is made explicit and elevated to the status of a principle, the 'String lamppost principle'. I think this fact also supports the claim that uniqueness of string theory should be considered to belong to the hard core of the 'standard string theory research programme'.

The Swampland programme has a clear positive heuristic, so it should be seen as a precise programme embedded in a more general research programme, where many different low-level heuristics[6] are allowed in order to achieve the goal.

I have just discussed the positive heuristic followed by string theorist until the emergence of the Landscape. While we can recognize different approaches, such as the Swampland programme, the general positive heuristic followed by the 'standard string theory research programme' has remained unchanged up to now: string theorists working in this research programme are still following some 'mathematical miracle' able to restore uniqueness, that is allowing for a complete formulation of string theory and for the recovering of the ground state corresponding to our world. This is a very generic heuristic, due to the fact that the incomplete and ambiguous status of string theory does not suggest any clear path to be followed. As a consequence, there is not a precise to-do list, but rather many different approaches, each one trying to achieve some information. Being 'standard string theorists' concerned about the incomplete formulation of string theory, their heuristic is mainly concerned with theoretical issues. They recognize that, at the actual stage, string theory is not able to recover our low energy physics, and are searching for something enabling them to achieve this goal.

A conceptually similar but practically different situation is faced by string theorists working on the Landscape. In fact, they are mainly concerned with the possibility of finding our ground state in the Landscape, or at least to obtain some information through statistical considerations and anthropic reasoning. Obviously, the acceptance of the existence of a Megaverse clearly implies a loss in predictive power, and the anthropic principle is used in order to fill this gap. 'Landscape researchers' recognize the impossibility of calculate constants of nature from first principles, thus using anthropic reasoning in order to explain their values. In any case, this is not as easy as it might seem, and I am going to explain why.

The Landscape is strictly linked to eternal inflation. As I have already discussed, the latter predicts the existence of an infinite number of Universes, because of the continuous generation of 'bubbles of false vacua' expanding and generating other universes.

[6]The positive heuristic of a research programme includes a list of goals to be achieved in order to develop the programme. In any case, the ways these goals could be achieved might be different not yet known. This means that in order to achieve a specific result one might need to make many steps, following what I called a 'low-level' heuristic.

The Landscape, on the other hand, predicts the existence of a huge but finite number of vacua. As a consequence, an infinite number of universes should exists corresponding to each valley we can find in the Landscape. This infinity of possibilities makes statistical considerations very difficult to be done. For example, suppose we would like to obtain the mass of the Higgs boson analyzing the distribution of its values in the Landscape. If the number of universes was finite, one might find a certain distribution and claim that in our Universe the value of the Higgs boson's mass corresponds to the most probable value we can find in the Landscape [7]. The problem is that the Landscape is supported by eternal inflation, implying that an infinity of universes exist, so we would have to be able to compare different infinities, not anymore different numbers. The difficulty to compare infinities would ruin any kind of predictive power (if we can consider such a method to have real predictive power). Landscape string theorists are well aware of this and other problems. I took this example from the popular book 'The cosmic Landscape', by Leonard Susskind [58]. In this book he presents and explains the Landscape programme, discussing its virtues, its problems, and answering to the critics. I postpone a very general review of this and Smolin's popular book to the next section; for this moment, it is important to notice that from this book (which I take as a paradigmatic example) it clearly emerges the absence of a clear path to follow.

The arguments in favour of the Landscape are methodological and philosophical, mainly concerned with issues related to naturalness, and the only key ingredient seems to be the anthropic principle, to which almost all the explanatory power is addressed, as I am going to explain.

It is important is to understand the difference between the two situations just discussed. In fact, while 'standard string theorists' are still searching for their theory, 'Landscape researchers' mainly aim to achieve some information of any kind from the Landscape, using the anthropic principle as the main instrument. In both cases, experiments do not belong to the to-do list. In the first case, it is quite obvious, because researchers are still searching for a complete theory to be tested. In the second case, we cannot talk about any serious attempt to test the Landscape hypothesis. For example, again in [58], we can find some proposals[8], but they do not strictly concern string theory and, even less, the Landscape. There may exist new proposals, but as a matter of fact Landscape researchers are not following a clear and definite heuristic.

Given this situation, the positive heuristics related to these research programmes are pretty simple and very generic, and might be summarized as follows:

[7] Here I am not concerned with the validity of this method, whose discussion would take us too far.

[8] For example, Susskind argues that finding our Universe to have a negative curvature would mean that we live in a 'bubble', as predicted by eternal inflation. In any case, this concerns mainly eternal inflation and not the Landscape, even if they may be seen as strictly connected. Another proposal refers to the existence of cosmic strings, but their existence is predicted also by other theories.

- *Standard string theorists:*
  - find a complete formulation of string theory or some principle/law/mechanism to select our ground state from the Landscape;
  - find some general features of theories inconsistent with quantum gravity in order to isolate (and indirectly achieve some information about) the consistent ones (Swampland programme);
- *Landscape researchers:*
  - try to achieve some information about our Universe from the Landscape;
  - use the anthropic principle to give an explanation to the specific values and structures we can find in our Universe.

3.4 Theoretical progressive and degenerative problemshifts

Now, the most important issue has to be discussed. Looking at the history of string theory we can recognize, not without ambiguities, the progressive and degenerative problemshifts[9], allowing to appraise each historical development as a rational or irrational step and also to give a general evaluation of the research programmes. As I have already discussed, a good philosophy of science should be able to give a rational explanation of the most of the historical development of a certain research programme; it will be the task of external history to account for those issues that rest unexplained by internal history. The evaluation of these problemshifts cannot be completely objective, but I think convincing arguments can be proposed.

String theory firstly appeared as a theory of hadrons. The first problemshift we can recognize is the explanation of the Veneziano amplitude. Dual models were found to be able to recover Veneziano's formula, but they were constructed for this very aim, so this cannot be considered a real success. We might consider it a progressive problemshift if dual models also predicted novel facts, or if they unexpectedly explained already known facts. Dual models led to string dual models, but initially strings were not considered to be real objects, this picture being considered as a mere (but useful) analogy. When strings started to be considered as physical objects, they showed to be able to explain

[9] I am talking about *theoretical* problemshifts, string theory not being directly verifiable.

other features of strong interactions, the most important being confinement. We have already seen how important hadronic string theory was for the understanding of those features QCD was not yet able to explain. This one falls within those cases accounted by Zahar's extension of Lakatos methodology. In fact, confinement was not a newly predicted fact, but it was already known. In any case, strings were introduced deductively from Veneziano amplitude, and confinement had not been considered at all. For this reason this step can be considered as a progressive problemshift. I would like to stress that I am still talking about hadronic string theory, so this is not an argument in favour of the string theory research programme as we know it today. By the case, this shows that the introduction of a physical string picture was a rational step.

When QCD 'replaced' [10] hadronic string theory, the means to perform the theoretical exaptation were already available, and indeed it was already starting. QCD was developed during the '70s and '80s, with asymptotic freedom discovered in 1973. The zero slope limit, which caused the theoretical exaptation, was investigated by Scherk in 1971. The exaptation process took some years to be completed, but it cannot be said it was carried on merely in order to save the theory. The timing was both fortunate and unfortunate: quoting Rickels, "it is interesting to note that Alton Coulter brought out a paper explaining the relationship between massless spin-2 fields and gravitational theory the same year as Scherk's first paper on the zero slope limit, 1971"([49], p.137, footnote 11), so the immediate identification of the Pomeron with the graviton was possible, but not trivial. Also, quoting Davide Olive, "this idea of unification of gauge and gravitational interactions was much discussed by the community in CERN Theory Division in the year 1971-1972 even though this was before the discovery of asymptotic freedom and the formulation of the Standard Model"([43], p.352). Hadronic string theory was 'defeated' by QCD, the physicists were completing the formulation of the Standard Model, string theory was consistent only in 26 dimensions and also had a tachyon in its spectrum; on the other hand, it was found that string theory might be able to unify gauge and gravitational forces. Given this situation, the slow process of exaptation can be justified without arguing that it was a mere manoeuvre made in order to save the theory.

It is true that there were few physicists working on string theory, and they were so enamoured of its mathematical beauty that they thought the theory should have a physical meaning. Their efforts explain how string theory could survive [11], so it is an 'external' explanation, but it does not explain how string theory could become a serious candidate as a unifying theory. In fact, string theory was not seriously considered until the anomaly cancellation result. This is an important fact explained by internal history: people were not immediately impressed, but it was the resolution of a technical problem that made string theory more appealing.

[10] or 'absorbed', as we have already discussed.

[11] For example, Rickles mentions "Murray Gell-Mann's influential role in keeping at least one string theorist, John Schwarz, working on strings"([49], p.128).

In any case, the exaptation resolved a problem affecting hadronic string theory and also transformed a not very welcome of its features into a great virtue. The problem I am talking about was the presence of massless particles in the spectrum, an odd thing for a short-interacting force. These particles were reinterpreted as the photon and the graviton, and it is important to remind that the exaptation was not performed in order to solve this very problem. On the other hand, the nice feature I mentioned above is the good UV behaviour of string theory scattering amplitudes. This behaviour did not represent a nice feature for hadronic string theory, being the objects at issue strongly interacting, but it is a very nice one for a quantum theory of gravity. The hypothesis of particles being extended objects was made many times in the previous decades, but had never given good results. Also, to find a finite or renormalizable theory of quantum gravity was one of the main and most difficult tasks of theoretical physics. This time, a (probably) finite theory of hadronic physics was found to be able to describe the gravitational force. Finally, in those years many attempts were made in order to find a unifying theory, and string theory was found to be such a theory. All these facts were completely unexpected, so they represent impressive progressive problemshifts (in the Zahar's sense).

In 1970, following the positive heuristic, the RNS model included fermions to the spectrum. As a consequence, a mathematical discovery was made: the first example of a 'superalgebra'.

In 1976 the GSO projection eliminated the tachyon from the spectrum. It was performed for this very aim, but also had an unexpected consequence: the dual spinor model was clearly pointing to supersymmetry. The GSO projection marks the birth of the Type I superstrings, but spacetime supersymmetry for string theory was considered only in 1979 by Green and Schwarz, giving birth to the complete classification of superstring theories we know today. Supersymmetry also implied a dimensional reduction from 26 to 10 dimensions. This 'chain' of theoretical steps can be considered as a progressive problemshifts: the elimination of the tachyon led to supersymmetry[12], which gathered many supporters at the end of the last century (and is still considered probable by many physicists) because of its potential (for example, it may help to solve the hierarchy problem, to unify the coupling constants, and also provides some candidates for dark matter), and supersymmetry led to a dimensional reduction making some physicists believe that further constrains might give a 4-dimensional theory.

The most important problemshift, in my viewpoint, is represented by the anomaly cancellation result. We have already discussed this issue, but I would like to remind what happened in 1984. A chiral anomaly was found to affect chiral superstring theories. The anomaly was found to cancel for the gauge groups $SO(32)$ and $E_8 \otimes E_8$, and the second one, as already discussed, was a very promising gauge group, able to return the gauge

[12]Without analyzing the historiographical reconstruction, one may risk to appraise a certain feature as an ad hoc assumption, falling in error. For example, one might think that supersymmetry was implemented for the very aim of eliminating the tachyon, but we know things did not go this way. This example shows how such an analysis might influence the appraisal of a research programme.

group of the Standard Model. It is not difficult to understand why superstrings became incredibly popular after this result: a consistent and probably finite unifying theory was found to be anomaly free for a phenomenologically promising gauge group. This was a theoretical progressive problemshift, one of the most important in the history of string theory, giving birth to the 'first superstring revolution'.
Furthermore, soon after this result, Calabi-Yau manifolds were found to be the correct manifolds for compactification. This represented, I think, another progress: these manifolds have SU(3) holonomy gauge group, a useful gauge group to break the symmetries to the SM gauge group. This was not an input, because Calabi-Yau manifolds were found asking for 6-dimensional compact spaces with Ricci flat metric preserving supersymmetry, then it was not obvious that the right compact spaces would have had a phenomenologically promising algebraic group. This is, once again, one of those facts increasing scientists' trust in the theory.
I would like to underline that, even if hadronic string theory was born from experimental data, string theory developed without any empirical guide. As I have already discussed, the methodology of scientific research programmes is able to explain the relative autonomy of theoretical science. String theorists followed the positive heuristic, and the research programme, in absence of empirical data, was pushed forward by theoretical progressive problemshifts.
Once superstrings were found to be anomaly free, and once heterotic string theory constructed, the remaining task was to compactify the extra dimensions in order to recover our 4-dimensional physics.
Compactification for string theory was studied since 1975, and in 1976 Scherk and Cremmer were studying 'spontaneous compactification' in General Relativity. String theory does not explain why only d-3 spatial dimensions compactify, so the compactification of the extra dimensions is something like an ad hoc adjustment made to match empirical evidence (that is the fact that we experience only 3 spatial dimensions). String theorists know that the (d-3)-spatial dimensions being compactified should be an output of string theory, and they hope to find some criterion or law explaining why only 3 spatial dimensions are extended. In any case, It is important to remind that the very mechanism of spontaneous compactification was not constructed ad-hoc, but it was found to be a solution of General Relativity, which is a low energy limit of string theory. Superstring theories contains gravity, describing the dynamics of spacetime, and spontaneous compactification was a dynamical phenomenon of spacetime. For this reason I think compactification was not completely ad hoc, having a certain mathematical and logical support. After all, to directly define a theory in 3 spatial dimensions is not less ad hoc than compactifing a theory to 3 spatial dimensions, they are both ad hoc formulations made to match the observable world. I would like also to remind that theories with extra-dimensions had already been proposed many times during the last century,

the most famous attempt being the Kaluza-Klein theory[13]. This theory shows the phenomenological potential of extra-dimensions, whose compactification is able to explain and provide low energy physics. This was the objective of superstring theorists after 1984: compactify the heterotic string theory breaking symmetries in such a way as to recover the Standard Model gauge group.

The appraisal of the rationality behind the compactification scheme is not trivial. A theoretical progressive problemshift occurs when an auxiliary hypothesis, added to explain a certain fact, implies the prediction of novel facts or the unexpected explanation of an already known fact (as proposed by Zahar). For this reason we have to understand if compactification, in addition to account for already known facts, also gave such a consequence. I think it can be said with certainty that this procedure did not predict novel physical phenomena. The analysis of Calabi-Yau manifolds and orbifolds showed interesting new facts, for example it was shown that strings propagates onto orbifolds without feeling any obstruction, something that does not apply to point-like particles, a purely stringy effect.

In any case, they do not represent novel physical effects, but rather mathematical properties of the theory itself, without any physical consequence.

The discussion about the explanation of unexpected already known facts is much more ambiguous. It was clear from the start of the exaptation process that dimensional reduction was needed in order to recover low energy physics. Being the goal to recover the Standard Model, the achievement would not appear as a progressive step if no other unexpected feature could be explained. But we are not looking at the whole research programme. A research programme is considered better if it has an excess of empirical content with respect to the other one. If string theory was able to recover *uniquely* the Standard Model, even if the procedure achieving this goal would not have led to novel facts, we would have ended up with a unifying theory describing both the SM and the gravitational force; the excess of theoretical content would have been evident. Also, string theory at that time had another attractive feature with respect to the SM. I have already mentioned that the SM have 19 free parameters, an odd feature expected to be absent in a unifying theory. String theory, in fact, seemed to have no free parameters at all. It was not only the excess of theoretical content, but also the way this might be achieved represented a very attractive feature. Moreover, string theory seemed to be able to explain other features that rest unexplained by the Standard Model. The properties of the compactified manifold determine all the low energy physics: the constant's values, the number of fermionic generations, and all the other parameters. They are all inputs in the SM, while string theory promised to explain them as properties arising from the topology and geometry of the compact spaces, thus giving a physical interpretation so far absent. One major example, already mentioned, is the number of generations. String theory is able to recover this number and explain it as due to topological features of

[13]General Relativity, in a different sense, is another example.

the compact manifold (the Euler characteristics). This one and other explanations do not represent neither new facts neither unexpected facts, because compactification was carried on for this very aim. In any case, the possibility to give a physical interpretation to these already known facts was a great virtue. Also, the same is true for the absence of free parameters. In fact, as Lakatos points out, "we should not abandon it if, supposing its rival were not there, it would constitute a progressive problemshift. And we should certainly regard a newly interpreted fact as a new fact, ignoring the insolent priority claims of amateur fact collectors. As long as a budding research programme can be rationally reconstructed as a progressive problemshift, it should be sheltered for a while from a powerful established rival" ([33], p.70-71).

A research programme able to account for another research programme's results and also giving a previously unavailable physical interpretation of them has more heuristic power even if the interpretations at issue do not predict any novel fact. Scientists are not unconcerned with similar results[14], as it is evident reading the following sentence: "it is still very satisfying to see the group SO(10) and the correct fermion representation emerging in a natural way. These successes are unlikely to be entirely accidental, and the relation between the number of generations and the topology of K is very possibly the seed of an eventual explanation of the origin of flavour"([20], p.409).

These reasonings explain why compactification schould not be considered as a mere ad hoc explanation of already known facts, even if it does not represent a real progressive problemshift in the Lakatos (or Zahar) sense. String theorists, in any case, were found to be too optimistic.

I argued that the unambiguous recovering of the Standard Model, together with an explanation of all its arbitrary values and features, would have constituted a rational motivation for an increasing trust in the research programme, a real victory I would say. But, as we know, things did not turn out that way. Compactification was found to result in the lost of uniqueness: the arbitrareness of constants in the SM appears in string theory as an arbitrareness in the parameters describing the shape and size of the Calabi-Yau manifold. This is the already mentioned 'ground state explosion', but it was not too dramatic. The possible number of ground states was estimated to be something like ~ 8000, a little number compared to the subsequent $\sim 10^{500}$ ground states in the Landscape. Many string theorists believed that the right ground state might be 'fished' testing the different ground states against experimental evidence. This procedure would have been carried on by inserting by hand the already known experimental values. In this way the theory lost the much-dreamed-of uniqueness, and it was considered definitely not able to give as an output the physical constants. String theory is a descendant of S-matrix theory and the bootstrap philosophy, aiming to the absence of arbitrary parameters. Finally, this approach failed also for the well promising string theory. It was

[14] As an example, if a novel satisfying interpretation of quantum mechanics would be given, it would constitute a progressive problemshift in the scientists' minds.

a bad blow for string theory, because one of its more attractive features was lost. Before discussing the roles of dualities and the M-Theory conjecture for the topic at issue, I want to stress that the lost of uniqueness would not have been too much dramatic if the Standard Model was finally recovered, even if using a certain amount of insertion by hand. In fact, string theory nevertheless would have given a physical interpretation to many unexplained features (the number of generations, even if inserted by hand, would have had a topological interpretation).

A great consequence of compactification are dualities, namely mirror-duality and T-duality. These dualities imply that apparently different descriptions really give the same physical effects, so they are physically equivalent. They provide powerful tools for the investigation of string theory's mathematical structure, and they have been of crucial importance for the following developments of the research programme. Their epistemological status is ambiguous. They do not represent a prediction made from compactification, representing only (still unproved) mathematical conjectures, and they neither give an interpretation to already known facts. In general, dualities represent very powerful mathematical tools that helped string theorists to discover and clarify unknown mathematical connections between different string theories, suggesting even more unknown connections (M-Theory), and also to discover new objects hided in the non-perturbative regime (Dp-branes). Dualities underlines a crucial concept in contemporary theoretical physics, that is the great impact that a mathematical discovery can have for the development of a research programme, further explaining how the autonomy of theoretical science works. They led to a dramatic change of perspective: Dp-branes appeared to represent the real fundamental degrees of freedom, with strings being only one of many different cases. Also, the web of connections they implied led to the hypothesis of the M-Theory, so that string theories were considered by many as different limits of a single theory. This 'second superstrings revolution' is a very special case in the history of science: an incredible change occurred within the research programme itself. Some mathematical discoveries were able to bring to light an incredible amount of new connections, structures and tools, an incredibly rich structure which had remained hidden until that moment.

In this optics it is not difficult to understand the renewed confidence that string theorists had in the programme: they found that the five theories they have in hands, may reduce to one single more general theory, and they also found very powerful tools to test string theory (at the theoretical level).

The methodology of scientific research programmes does not assess them. Dualities do not consist of auxiliary hypothesis made in order to match a certain empirical evidence, subsequently providing novel predictions or unexpected explanatory connections, but represent mathematical relations giving very strong consequences. Lakatos' methodology does not apply to this case because the difference between progressive and degenerative problemshifts was given in order to recognize 'scientific' or 'acceptable' hypothesis from ad-hoc hypothesis made to explain some observed fact, but in this case there were no facts to be explained. Even if a serious discussion about the epistemological status of conjectures is outside our scope, I think it is worth to investigate a little further this issue, both because of their crucial role in the string theory research programme and because they are important for the appraisal of the research programme's rationality. My suggestion is to appraise the value of a conjecture looking at its consequences. In this case, we find many and strong consequences.

T-duality has many implications: it implies the existence of a minimum length scale, the existence of D-branes, and it is a key ingredient in the web of dualities relating the five types of string theories, as it is for S-duality.

As already mentioned, dualities point to a new kind of unification, soon after uniqueness was lost. The M-Theory conjecture made string theorists still hoping that some fundamental principle underlying M-Theory would have selected the correct ground state. The awareness that the biggest piece of the puzzle was absent, made people believe that some miracle might restore uniqueness.

Finally, the combined use of Dp-branes and dualities have crucial consequences for the applications. The most important example is the calculation of the Bekenstein entropy, making significant use of the open-closed strings duality. It was the first time that the microstates were counted and the entropy was found without leaving any arbitrary parameters, so this result gave new confidence in the programme. The calculation, as already told, was performed only for maximally supersymmetric near extremal five- and four-dimensional black holes, cases far to be realistic. Anyway, the application to the simplest cases is a natural way to proceed, suggested by the positive heuristic. We cannot claim that these results can be generalized to more realistic cases, nor that string theory would fail in such cases, but a great success for a specific and idealized case is nevertheless a success, rationally increasing confidence in the programme.

AdS/CFT correspondence is another consequence of dualities and Dp-branes, related in particular to the open-closed strings duality. Again, I suggest that the validity of a conjecture might be appraised by looking at its consequences. From this viewpoint, the AdS/CFT correspondence has important consequences, both for the applications to other fields of research (like QCD) and for string theory itself. It points to a non-perturbative definition of string theory and a better understanding through the connection to quantum field theory it allows for. Also, it potentially resolve the 'black hole information paradox', and it is linked to the holographic principle, which seems to be more and more important in contemporary theoretical physics. The fruitful applications of this correspondence are not affected by the critics of string theory's opponents, while its status inside the string theory's framework is. In fact, this correspondence works for a non-realistic spacetime, similarly to the Bekenstein entropy calculation. As it is for the latter case, many attempts have been made to generalize and refine this conjecture in order to apply it to more realistic cases, and much work is still being done today.

Dualities led to successful results, so the great trust about their validity is not so much surprising. In fact, one might argue that this success would be inexplicable if they were completely wrong. I will come back to this argument in the next chapter, when talking about the 'no miracle argument'. The epistemological status of these conjectures, together with their incredible consequences with respect to our understanding of spacetime, deserves deeper investigations, but it would take us far from our scopes. In the optics of the methodology of scientific research programmes, we can feel satisfied once we have managed to provide a rational motivation of the trust string theorists have in them.

Now, the most important issue has to be discussed: the emergence of the Landscape and its consequences. Here, we are concerned with the ways string theorists responded to the discovery of a huge number of possible vacua; the scope is to appraise the solutions they found in order to make their research programme go on. As I have already argued, this moment caused the split of string theory research programme into the 'standard string theory research programme' and the 'Landscape programme', differing by the ways they reacted to this event. The situation, to date, is not so much different, so an epistemological appraisal of this historical step is crucial for a clear understanding of contemporary string theory research. It is interesting to notice that the splitting into two lines of research was caused by a different methodological choice, by a very different philosophical interpretation of the same result. It is in such a cases that a good methodology might demonstrate to be very useful in order to clear the situation and frame it in the correct ground. Below, I am going to discuss the two research programmes separately, trying to provide such an evaluation.

As already mentioned, those string theorists rejecting the Landscape still believe that there exists an unknown theory describing uniquely our Universe. In their optics, the Landscape is not real, but a set of solutions, and a certain principle or law should exist allowing for a clear formulation of M-Theory and selecting the correct ground state corresponding to our Universe. I would like to remind that not only we do not know what really M-Theory is, but we neither know what in general a string theory is: we do no have any fundamental law, equation or principle, and we do not know which are the fundamental degrees of freedom for a correct formulation.
This line of research represents the natural continuation of the string theory research programme as it was from the start: a single unifying theory with no free parameters is still searched. The only thing that has changed is that the objective seems to be much more difficult. In the last two decades, no important steps forward have been made, so string theory is still incomplete as it was in the early 2000s. The epistemological considerations we have already made for the string theory research programme until the end of the '90s rest unchanged, but the sterility of this approach demands an explanation. The explanation I give in the light of the methodology of scientific research programme in based on the already discussed concepts of progressive and degenerative problemshifts. I told that once a research programme faces a counterevidence, scientists do not allow it to

be directed against the hard core. In this case the counterevidence is a theoretical result, consisting of $\sim 10^{500}$ ground states. I have discussed that in the hard core there is the assumption of the existence of a unique unifying theory with no arbitrary parameters, so this counterevidence was mining the ragion d'etre of the whole research programme. In such a case, scientists react redirecting anomalies to the protective belt, adding auxiliary hypothesis. In this case, string theorists conjectured the existence of a unknown principles able to recover uniqueness. Now, a progressive problemshift occurs when such an auxiliary hypothesis not only explains the anomaly, but also predicts novel facts, leading to an excess of empirical (or theoretical) content. In this case, however, there seems to be a problem. We do not know if the auxiliary hypothesis, one day, will be found to work: if some principles will be found able to recover uniqueness, the auxiliary hypothesis will save the hard core and the research programme as a consequence. I have already argued that if string theory would be found to recover the Standard Model but not to give novel predictions, it would constitute anyway a progressive problemshift, both because it would describe the Standard Model and the gravitational force in a single framework, so having an excess of empirical content with respect to the SM itself, and both because it would give a new physical interpretation to the already known features of the SM. In any case, no such principles have been found so far, and two decades are passed. We do not know at this stage if string theory will be able one day to recover the SM, so we cannot define it as a progressive or degenerative research programme, because we do not know if this auxiliary hypothesis will save the research programme from the second 'ground state explosion'. We only know that, if it will be the case, it will constitute a progressive problemshift, even if no new predictions will be made.

For these reasons, the 'standard string theory research programme' cannot be defined neither progressive nor degenerative. It may seem absurd not to define a research programme unable to give any result during five decades as a degenerative research programme; it may seem that lakatos' methodology has been used in order to save the status of the research programme itself. But the great virtue of lakatosian methodology lies in its relational character. A research programme can be evaluated at each step looking at progressive and degenerative problemshift, but its final evaluation has to be made only with respect to other research programs. This methodology is able to explain how is it possible that string theory has not been abandoned even if it failed to give new results: a research programme is 'falsified' only when it is superseded by another more powerful research programme. This might explain why string theory is still dominating the research field: not only there does not exist any better theory, but no other unifying theory is available at all. String theory is the only candidate so far, so even if it is failing to meet the expectations, even if it cannot be considered as a progressive research programme, even if someone might consider it to represent a degenerative research programme, it cannot be 'falsified' in the lakatossian sense. It is an important point, because it overcomes a big problem: time. One might say that, being string theory unable to give new results in so much time, it should be abandoned. But this cannot be an objective

argumentation. Others may say that the time limit to abandon a programme is 40 years, other that it is 23 years, and so on. Furthermore, it does not take into account the actual status of experimental physics. Obviously, string theory being unable to give predictions at low energy scales and the experimental technology being unable to test energy scales at the Planck scale both contribute to this stalemate situation. Time is an important component but its role cannot be appraised objectively; many contingent factors might influence the status of a certain research programme. The risk is that scientists use the ambiguity of this issue to argue against the research programme. Moreover, history of science shows that it happened many time that successful research programme had to wait a long time to show their success. For example, the Copernican theory was proposed in 1543, and we may say that it was verified only in 1610 by Galileo, that is after nearly 70 years, and this because of the experimental difficulties.
Reminding Lakatos' words, a new research programme "may start by explaining 'old facts' in a novel way but may take a very long time before it is seen to produce 'genuinely novel' facts". In addition to the experimental difficulties, also the mathematical ones are of importance. We have already seen that important revolution occurred when a technical problem was resolved. This underlines the crucial role of mathematics in contemporary theoretical physics: experiments do not represent the only difficulty, but mathematics is even more crucial. Furthermore, the already mentioned revolutions and mathematical discoveries stress the relative autonomy of theoretical science. Even in absence of empirical data, mathematical discoveries were able to push the research programme forward. This concept is important in order to understand that also the mathematical difficulties highly influence the development of a research programme. String theory has a very complex mathematical structure, and a lot of new mathematics have been developed for the very aim of understanding it. For this reason, I think that such difficulties can contribute to slow down a research programme's development. Also, string theorists might expect that some mathematical discovery may push forward again their programme. It would not be a surprise, because it happened many times also in the research programme itself. Again, quoting Lakatos[33], "the real difficulties for the theoretical scientist arise rather from the *mathematical difficulties* of the programme than from anomalies".
For all these reasons, I think that a great virtue of the methodology of scientific research programmes lies in its ability to completely bypass the role of time. It is not important how many years have been passed, so one cannot demand for a clear time limit after which a research programme should be considered as a degenerative one. The only important thing is the relation between the research programme and other research programmes. I repeat that it does not make sense to say that string theory should be rejected because it fails to give predictions. Things may change, and as we have seen time is not a reliable component. String theory will be completely 'falsified' when a better theory will overtake it, and this fact is independent of time.
Obviously, this does not mean that time should not be considered. Reminding an already given quotation, "One may rationally stick to a degenerating programme until it is

overtaken by a rival *and even after*. What one must *not* to do is to deny its poor public record." This point and its sociological consequences are of crucial importance and will be discussed when talking about external history.

The Landscape programme represents the most followed programme in contemporary string theory research. It is linked to many other topics, the major being the anthropic principle, naturalness, and eternal inflation.
In order to appraise this programme from a methodological point of view, we must look at the motivations leading to it. When the hypothesis of the real existence of the ground states in the Landscape was introduced, the anthropic principle was already a debated topic. In 1987, Stephen Weinberg used it in order to estimate the value of the cosmological constant. His calculation was based on anthropic arguments: if the cosmological constant was bigger than a certain value, life would not have been possible. A decade later, λ was measured and found to be in good agreement with the value estimated by Weinberg. Here we are not concerned with an appraisal of the validity of anthropic reasonings, being the scope to evaluate the methodological rationality of the Landscape programme. For this reason I postpone the discussion about this principle to the next session, where it will be of importance in order to analyze the contemporary debate around string theory and frame it in the right scenario.
In this case, the anthropic principle was used in order to address the puzzling issue represented by the cosmological constant. In the same period, inflation and eternal inflation were proposed, and found many supporters. Given this situation, we can try to give an explanation of the birth and rise of the Landscape programme.
When the Landscape was found, three main possibilities were available: reject string theory, reject the Landscape and go on searching a unique theory, or accept the real existence of the Landscape. These options might seem to be all equally acceptable, but things are not that simple.
I have already discussed the hard core and protective belt the string theory research programme was based on. String theory was born as an attempt to find a unifying quantum theory of all forces. The positive heuristic, at first, was clear: recovering the Standard Model and the constant's values leaving no free parameters. Even if anthropic principle was already in use and eternal inflation had already been proposed, before the Landscape was found string theory was not concerned with these topics, it representing the most promising theory to solve any problematic issue. No mention of other Universes, no mention of the impossibility to explain why constants have the values they have; rather, the heuristic had always been to find an univocal compactification scheme automatically returning all those values. Then, when stabilization of moduli was shown to give something like 10^{500} possible configurations, string theorists were in trouble. Discussing the methodology of scientific research programmes, I explained that all research programmes develop in an ocean of anomalies, neglecting them and hoping that following the positive heuristic the programme will turn all these anomalies into corroborations. For this rea-

son, to neglect the existence of the Landscape and go on with the research programme was a rational step, performed by addressing this problematic to the protective belt in order to save the hard core. We cannot yet evaluate this move as a progressive or a degenerative problemshift, but we can recognize a standard and rational way to proceed. The same cannot be said for the Landscape programme. I want to stress that here I am not evaluating the programme itself. Philosophy of physics is not physics, its task is not to appraise a research programme on scientific grounds, but the evaluation concerns methodology and rationality. As already discussed, Landscape researchers changed the hard core sacrificing uniqueness and assuming the real existence of an infinite number of Universes whose ground states are described by the Landscape. This is, in my opinion, a quite evident ad-hoc answer. To understand this point, I want to make even more clear that the possibility of the existence of other universes was not taken into consideration during the historical development of string theory. In Peter Woit's book one can find a quotation by Wolfgang Lerche clearly explaining this point:

> Well, what I find irritating is that these ideas are out since the mid-80's; in one paper on 4d string constructions a crude estimate of the minimal number of string vacua was made, to the order 10^{1500}; this work had been ignored (because it didn't fit into the philosophy at the time) by the same people who now re-'invent' the landscape, appear in journals in this context and even seem to write books about it...the whole discussion could (and in fact should) have taken place in 1986/87 ([63], p.245).

This hypothesis was added only once a theoretical counterevidence forced string theorists to make a methodological decision. The term 'counterevidence', in this case, has a clear meaning: the Landscape was pointing to the opposite direction with respect to the motivations behind the string theory research programme. For this reason, I think it is not incorrect to say that the counterevidence was directly integrated in the hard core in order to save the theory. In the light of the methodology of scientific research programmes, this might be appraised as an ad-hoc hypothesis. In any case, this is a crucial and subtle point, deserving a clear explanation.

A research programme can be viewed as a sequence of increasingly refined theories sharing a common set of fundamental assumptions, constituting the hard core. The theories are constructed starting from this hard core, which includes all those statements that scientists are unwilling to refuse. They are developed following the positive heuristic, and anomalies are faced directing them to a protective belt of auxiliary hypothesis, rather than to the hard core. It is important to understand that the hard core, as well as the protective belt, is not untouchable, but further statements might be added. In fact, following the positive heuristic, scientists might find some principle or anything else to be necessary (or nearly) for the research programme, so they will add it to the hard core. For the string theory research programme, an example is given by supersymmetry, which was recognized to be fundamental (even if, maybe, not necessary); another example is the existence of Dp-branes, extending the statement assuming the existence of strings by including also higher dimensional objects. Changing the hard core, in some cases, naturally represents an odd move. There might be cases where a certain statement has to be removed from the hard core. For example, if supersymmetry was demonstrated not to exist, it would be removed from the hard core. This would have many bad consequences for the research programme. String theorists might decide to continue developing their research programme, restarting from string theories without supersymmetry, but most results would no longer be valid, decreasing a lot the credibility of the programme. Another case - the important one for us - is the replacement of a statement in the hard core with another one, in contrast with the previous. I think such a move might be considered an ad hoc adjustment. The hard core includes all the assumptions the research programme is based on, the 'ragion d'etre' of the programme itself; their validity is so important that scientists do not allow anomalies to be directed against them, constructing a protective belt of auxiliary hypothesis. If a scientist allows an anomaly to be directed against the hard core, he is mining the pillars on which the research programme stands. In any case, this move is not inherently an odd one: if the replacement represented a progressive problemshift, it would be 'acceptable'. Coming back to string theory, I argued how the search for a theory with no free parameters was a major objective, and how the existence of the Megaverse was not assumed until the appearance of the $\sim 10^{500}$ ground states. The absence of arbitrary parameters was an important element in the hard core. String theory, coming from dual models, is in a certain sense the successor of the bootstrap approach, and maybe its highest expression: also the number of dimensions is constrained by mathematical coherence. This feature was considered a great virtue, because the absence of free parameters was demanded for a theory describing physics beyond the Standard Model, which has 19 parameters. This feature of string theory also appears in the already presented list of string theory's virtues given by Gell-Mann in a talk in 1987, further underlining its importance.

When the Landscape was found, a methodological move of the above type was made. Facing a 'theoretical counterevidence', some string theorists decided to sacrifice the assumption of a unique theory with no free parameters in favour of a Megaverse with infinitely many combinations of parameter values. These values were no longer considered to be univocally determined by the theory, but rather a 'discretuum' of theories describing different universes was supposed to exist.

The hard core, then, was radically changed, replacing an assumption with a nearly opposite one; this replacement is stressed by the division of the research programme in two different programmes, where one is the natural continuation of the traditional programme retaining the old hard core.

As discussed above, it is not surprising that many scientists considered this move an ad hoc adjustment. A theoretical problem threatening an hard core statement was resolved by eliminating the latter, and assuming the former. This can be rationally interpreted by many as an ad hoc move, being a clear example of scientists trying to save a theory.

In any case, even if these arguments explain why many scientists are strong opponents of the Landscape, I remind you that it does not mean that this step is intrinsically an odd one. The Landscape is a result of string theory, and it might be rejected or accepted. Even if its acceptance, as discussed, might appear as a desperate act made in order to save the theory, nothing prevents scientists from believing in a certain result thus radically transforming the whole research programme. This move would be 'acceptable' if it led to a progressive problemshift. Lakatos methodology allows us to recognize 'acceptable' and ad hoc stratagems by looking at progressive and degenerative problemshifts. For this reason, in order to give a complete appraisal of the Landscape hypothesis, one should look at its consequences: if it was able to lead to an excess of empirical content, one might evaluate it as a progressive problemshift. Evidently, it is not the case. The Landscape alone, at least in the light of our contemporary knowledge, has no predictive power. The unique kind of 'explanations' it can provide are due to the anthropic principle, which is empowered by the infinite number of possible universes.
There exist two different formulations of the anthropic principle, the strong one and the weak one.
The strong anthropic principle states that our existence is necessary, that the Universe is designed for us. In this optics, we hare considered to be very special observers, because the Universe is aimed at our existence, and has features necessarily leading to the development of our lifes. In this form, the anthropic principle appears as an act of faith, and it is generally rejected.
The weak anthropic principle simply states that we have to consider our existence when building theories or estimating parameters. In fact, certain values of physical parameters would not allow life to develop, so they are not all equally probable but are constrained by the possibility of our existence. In bayesian terms, the weak anthropic principle uses our life as an evidence to update probabilities. Furthermore, life is not considered necessary, but only a casual event. In this latter form, the anthropic principle is quite trivial, and is generally considered to be valid[15]. The debate, anyway, concerns its status of *scientific principle*. In fact, landscape researchers pretend to consider this principle a scientific one, even if it cannot give any precise prediction. In fact, Weinberg's prediction of the value of the cosmological constant does not represent a real prediction, but an estimation made using the anthropic principle to assign a higher probability to a certain range of values; there are no novel phenomena predicted, only a range of values for a specific constant. Also, they claim that this principle, together with the concept of a 'Megaverse', has 'explanatory power'. In fact, landscape researchers argue that the incredible fine-tuning of parameters we observe, which seems necessary for our existence, can be rationally explained only by assuming the existence of a Megaverse. The existence of infinitely many universes with different constant values, together with the evidence represented by life, 'explains' why and how is it possible that constants of nature have

[15]When talking about the anthropic principle I will always refer to the weak form.

the incredibly fine-tuned values we observe.
This line of reasoning might give an explanation to the fine-tuning we observe, but I have already discussed the difficulties the Landscape is facing also with respect to this aim, due to the difficulty to address a certain probability to a certain ground state. The anthropic principle, in the Landscape and eternal inflation frameworks, might be able to explain only a few metaphysical facts: we exist because there is an infinite number of Universes and so, statistically speaking, it is not surprising to find at least one of them able to support life; this 'explains' the incredible fine-tuning of constants we measure without resorting to the concept of an Intelligent Design. This state of affairs clearly shows why some physicists talk about Landscape and anthropic principle as a 'religion', these representing, up to now, only metaphysical arguments.
All these arguments should make clear that the hypothesis of the real existence of other Universes does not predict any novel fact, neither it explains new unexpected facts. In fact, there is no necessity of the Landscape in order to estimate constants values through the anthropic principle. We have already seen that Weinberg estimated the value of λ years before the Landscape. So, even if one wish to say that the anthropic principle is able to make predictions, this does not support the Landscape. The real existence of other Universes would only serve to avoid the possibility of an Intelligent Design. Though this is a very interesting point, it is metaphysics, not science.
This discussion should make clear two main facts: first, the radical change in the hard core explains why many scientists consider the Landscape hypothesis as an hoc stratagem, made to save the theory; second, even if this assumption could have represented an 'acceptable' step, it has failed to give novel physical predictions, so it might be appraised as a degenerative problemshift, and the Landscape research programme as a degenerative one. The situation, in fact, is very different from the one characterizing the 'standard research programme'. Standard string theorists are still waiting for a complete formulation of string theory, which might resolve the 'problem' of the Landscape; in this optics, assuming the existence of a certain principle or something else able to restore uniqueness is a natural way to neglect an anomaly and proceed with the research programme, hoping to resolve it in the future, following the positive heuristic. For this reason we cannot appraise this decision as a progressive of degenerative problemshift.
For the Landscape research programme, the situation is quite different. The replacement of an element in the hard core is not a common way to proceed. An anomaly is usually faced by adding auxiliary hypothesis and going on developing the programme, following the positive heuristics. The deviation from this natural methodology represented by the transformation of the hard core not only explains why many scientists consider it to be ad hoc, but also strongly demands for an excess of content, something that the Landscape have failed to provide until now.
In order to avoid all the possible critics one might move to this line of reasoning, I would like to remind, once again, what a methodological appraisal is. One might say that the existence of the Landscape is inevitable, that anthropic reasoning is the only way to

explain the incredible fine-tuning we observe in our Universe, or that eternal inflation supports the Landscape. These arguments might be acceptable or questionable, but they would be off-topic. The point is that my aim here is to appraise the rationality and acceptability of a research programme looking at its structure and historical development, not at technical and theoretical issues, which are up to physicists. Doing the contrary would go against the 'scientific expert principle'. Philosophy of science represents something like a super-partes instrument to investigate the rationality of the scientific process, helping scientists to develop a more critical mind with respect to their own methodology, to be more conscious and to establish a well founded methodological debate. For these reasons, I am not obviously stating that the Landscape programme would not achieve any success without any doubt. Maybe Landscape researchers have hit the mark, or maybe not, but here I am not concerned about that.

I argued that 'standard string theory research programme' might not be considered as a degenerative one. Counterevidence was faced adding the hypothesis that it would be solved once a complete formulation of string theory would be achieved, or some fundamental principle was found. This solution cannot be appraised as progressive or degenerative, because we already do not know neither if, once this formulation would be found, it would turn out to predict novel facts or to account for the Standard Model, nor if such a formulation will be found, nor if such a formulation exists at all. We can only say that this line of research, if things will remain as they are, will be probably overtaken by another research programme. I have also discussed the Landscape research programme, claiming that it constitutes a degenerative research programme. I remind that this evaluation is not conclusive because, this being valid for both of them, a "little revolution or a creative shift in its positive heuristic may push it forward again" [33]. In any case, the stalemate the two programmes are facing, due to the absence of a clear positive heuristic, allows to introduce another interesting concept formulated by Lakatos, which also explains why many physicists strongly criticize string theory despite its partial successes. Quoting Lakatos:

> My account implies a new demarcation criterion between *mature science*, consisting of research programmes, and *immature science* consisting of a mere patched up pattern of trial and error. For instance, we may have a conjecture, have it refuted and then rescued by an auxiliary hypothesis which is not *ad hoc* in the senses which we had earlier discussed. It may predict novel facts some of which may even be corroborated. Yet one may achieve such 'progress' with a patched up, arbitrary series of disconnected theories. Good scientists will not find such makeshift progress satisfactory; they may even reject it as not genuinely scientific. They will call such auxiliary hypothesis merely 'formal', 'arbitrary', 'empirical', 'semi-empirical', or even *'ad hoc'*. Mature science consists of research programmes in which not only novel facts but, in an important sense, also novel auxiliary theories, are anticipated; mature science - unlike pedestrian trial-and-error - has 'heuristic power'. Let us remember that in the positive heuristic of a powerful programme there is, right at the start, a general outline of how to build the

protective belts; this heuristic power generates *the autonomy of theoretical science* ([33], p.87-88).

I think this concept fits very well the situation at issue, also showing the progressive character of the very methodology of scientific research programmes. String theory, I think, reflects quite well the features of an immature science. Many physicists consider string theory more as a mathematical framework than a real theory. The web of dualities supporting the M-Theory conjecture and the subsequent emergence of the Landscape led to a situation where string theorists tried to proceed blindly, analyzing the mathematical structure of the theory in order to achieve some hint about the way to proceed or some information about what string theory really is. The lost of a definite positive heuristic blocked the theoretical development, leaving us with an incomplete theory where no fundamental principles or laws or degrees of freedom can be recognized. It is understandable not only why many scientists think of it as a mathematical framework, but also why its opponents think of it as a degenerative research programme. Even if, at least for the 'standard research programme', such a judgement might be premature, such claims are quite comprehensible. The immature character of string theory is reflected also in the strong trust in conjectures it is based on. One might argue that string theory has a beautiful mathematical structure, and that its results are certain, but the reality is quite different. Although I have given a rational explanation for the trust string theorists have in these conjectures, the fact remains that string theory is based on unproved mathematical statements.
The 'standard researchers' are well aware of the immature character of their research programme. Their hope is really that some discovery will lead to a complete and mature formulation of the theory. David Gross (whom I take as the main representative of the 'uniqueness' approach) thinks of string theory as a research programme in continuous development, demonstrating to be well aware of the status of string theory[16]. The immature status of string theory is mainly due to the absence of any fundamental principle or law, making it appear more like a mathematical framework rather than a clear physical theory. In the closing sentence ending the already mentioned book by Witten, Schwarz and Green, published in 1987, one can read: "The truth is that while much is known about string theory, the roots of this subject lie hidden. We do not know what principles unify the many surprises that make string theory possible. We do not know why propagating strings, or world-sheet path integrals, are a proper starting point for a generalization of nonabelian gauge theory and general relativity. The answer to such questions may lie in directions not yet contemplated"([20], p.551-552). The authors are

[16]As Rickles remembers ([49], p.235), in a 2005 talk David Gross revisited a list of eight fundamental questions about the status of string theory that he had already proposed at a meeting on 'Unified String Theory' in 1985. Two of these questions were: "What picks the correct vacuum?" and "What is string theory?". These questions make explicit both the 'uniqueness approach' and the awareness of the incomplete status of string theory; also, the fact of this list of questions being re-proposed (although in a revisited form) after 20 years makes evident both the continuation in time of this approach and the chronicity of string theory's immaturity.

well-known supporters of the 'uniqueness approach', and this quotation clearly shows not only in which sense string theory is immature, but also that finding its roots was an important objective already in the end of the '80s. From those years to the present day, even if the discovery of Dp-branes has improved our understanding of string theory, no basic principles or laws have been found, and 'standard string theory researchers' are still searching for a complete formulation of the theory.
Landscape researchers, instead, accept string theory as it is, with its Landscape of universes, trying to justify their choice referring to eternal inflation or anthropic arguments. I think 'standard researchers', given the situation, get a point when they say that this choice is premature. They recognize that string theory is immature, arguing that taking the Landscape for real, given this sorry state of affairs, is not a wise way to proceed. I agree with them. We do not know if the conjectures string theory are based on are valid, nor if it is really a finite theory of quantum gravity, nor if string theory as a well definite theory really exists. Arguing for the existence of a Multiverse starting from an immature theory and supporting this hypothesis only through metaphysical reasonings and eternal inflation cosmology, yet not proved and nor universally accepted, can be rationally considered to be premature.
This does not mean that a Landscape researcher should abandon its research programme. 'Scientific honesty', in this methodology, means that a research programme should be abandoned when superseded by a more powerful (with more heuristic power) research programme, and we know there are no such research programmes in the contemporary theoretical physics scenario. Reminding an already given quotation by Lakatos, "one may rationally stick to a degenerating programme until it is overtaken by a rival *and even after.* What one must *not* to do is to deny its poor public record. Both Feyerabend and Kuhn conflate *methodological* appraisal of a programme with firm *heuristic* advice about what to do. It is perfectly rational to play a risky game: what is irrational is to deceive oneself about the risk". I think this is an important point, because many string theorists, and especially Landscape researchers, usually give a pretty false narration about the status of their research programme. Many of them talk about string theory as the only possibility, saying that its results are indisputable. In the light of the above quotation, it is not surprising that their opponents accuse string theorists of 'scientific dishonesty'.
This situation has led to remarkable consequences that I will investigate further in the next chapter.

3.5 External history

I have proposed my personal interpretation of string theory research programmes in the light of the methodology of scientific research programmes. I think this reconstruction explains almost all the rationality of this development, and how we arrived at today's situation. A few facts remain to be explained that internal history cannot rationally reconstruct. So, it is external history, also called 'sociological history', that should take this task. The main fact to be explained is why contemporary theoretical physics is still dominated by string theory research programme. In fact, I have just argued that the most followed research programme, the Landscape programme, can be appraised as a degenerative one, and also the search for uniqueness has experienced a stalemate during the last decades. Looking at the number of publications, we can see that a very large number of papers continues to be published annually:

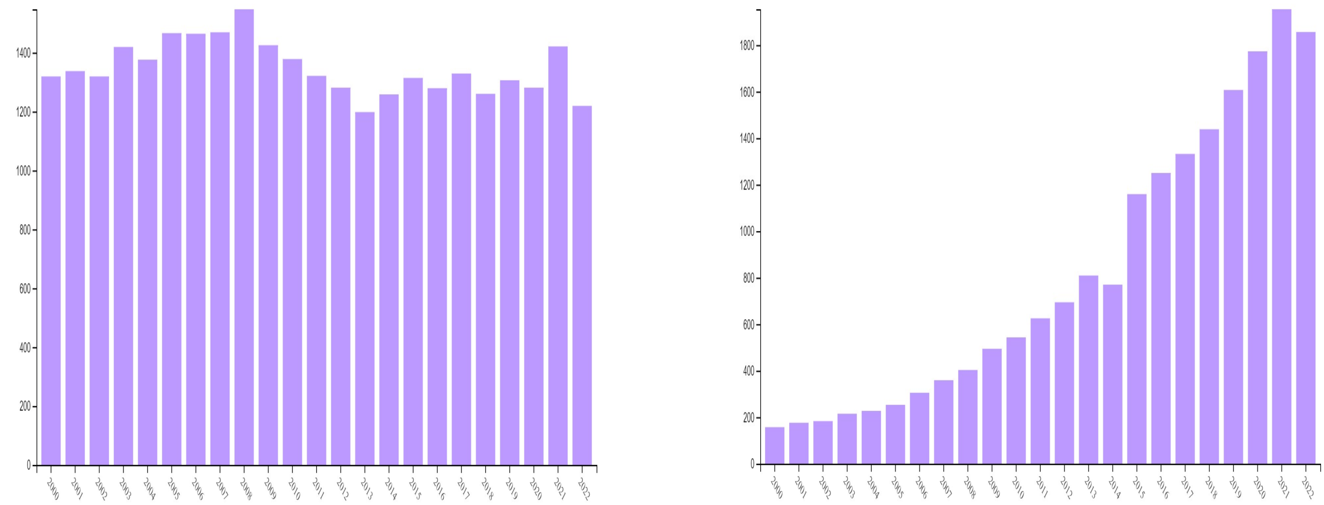

Figure 3.1: Number of publications per year related to 'string theory', 'superstring theory' or 'M-Theory' to the left and 'string Landscape', 'string theory Landscape' and 'Landscape programme' to the right. *Image source*, Web of Science.

We see that, for string theory as a general subject, the number of published papers is still incredibly large and has not decreased during the last two decades, while papers concerning the Landscape has experienced a great increase in number.
I would like to remind that string theory is importantly based on supersymmetry, and low-energy supersymmetry is usually considered. Supersymmetry was expected to be discovered by LHC in the last years, but no evidence has been detected so far. In order to resolve the hierarchy problem, supersymmetric particles were expected to be found below a certain energy, which has been passed without giving any result. The resolution

of this problem was one of the major motivations for supersymmetry, so trust in this theory has slightly decreased in recent years, even if many hope to find it at somewhat higher energies. The decrease of trust in supersymmetry should have led to a decrease of trust in string theory as well, because - as I have argued at the beginning of the chapter - the latter is strongly based on the former. Even if string theory does not fix any energy limit for supersymmetry breaking, and even if the 'hierarchy problem' does not represent really a problem, it being not discovered when it was expected to happen should have diminished the confidence in superstring theory. Looking at the above graphs, the situation seems to be quite different.

Given the epistemological status just described, and the above presented experimental situation, one should ask how is it possible that research in string theory is still so dominant

There are popular books, such as the already mentioned books by Lee Smolin [52] and Peter Woit [63], accounting for odd sociological dynamics characterizing the string theorists community. In particular, Lee Smolin devotes the last five chapters of his book to the 'sociology of theoretical physics' and to the aim to restore an healthy ethics in the scientific community. Smolin describes the string theory community as affected by the so called 'group thinking' and other cognitive biases, an aggressive and closed minded community thinking to be the only holder of truth, and hierarchically organized. Now, I have not enough information to argue neither in favour nor against these arguments, so I will not investigate them further. It is a very important issue deserving a deep analysis in order to reconstruct the external history, and sure it cannot be carried on only by reading two string theory's opponents popular books. In these books one also finds a description of the administrative organisation of academies and worrisome dynamics related to publication mechanisms (such as the paradigmatic 'Bogdanov affair', [63], p.217). I have not the necessary competencies to make a constructive critics about them, but it seems that they strongly contributed to the reaching of the actual situation we observe in theoretical physics research. If the scientific community really cares about the success of the scientific process, it should make self-criticism and fix these problems as soon as possible.

These books were published in 2006, and the situation might be changed during the subsequent years; in any case, the social and institutional dynamics they describe might be able to explain why we observe the above graphs.

I told I am not concerned with the issues just presented, but there is one particular topic very important in the light of the methodology of scientific research programme: the proliferation and competition of scientific research programmes. Lee Smolin rightly argues that when a certain theory is not universally considered by the scientific community as corroborated, scientific ethics asks to encourage the proliferation of other research programmes and different approaches, it being of paramount importance for the greeting of the community itself. When a theory cannot be considered as proved, the presence of other theories can lead to an healthy competition, fruitful both because results in a certain field of research can help to develop other fields, and because it helps to avoid dogmatic and premature beliefs. Smolin (referring to string theory) also claims that a research programme, in the light of these arguments, should not be allowed to monopolize the research funding before the theory is corroborated or universally considered as very reliable.

In the framework of the methodology of scientific research programme, this concept is even more important. As already discussed, 'falsification' occurs when a research programme shows an excess of empirical (or theoretical) content with respect to its predecessor. This concept underlines the relativist character of falsification, and the fundamental role played by the coexistence of different research programmes. Pluralism must always be encouraged, and not only when a research programme starts to enter a degenerative phase.

With regard to string theory, Peter Woit reminds that in 2001 a science reporter published an article for the New York Times titled 'Even Without Evidence, String Theory Gains Influence', claiming that string theorists were receiving all the benefits usually ascribed to successfully verified research programmes, such as tenured faculty positions, federal grants, prestigious awards and so on ([63], p.230). He also remembers that the MacArthur fellowship awards, in the years immediately following their creation in 1981, were assigned to nine physicists of which eight were string theorists. This and other facts contribute to the picture given by Lee Smolin of a research programme monopolizing the research resources, even if this research programme has never been proved by any experiment. I think this is an important point. The methodology of scientific research programmes underlines the importance of proliferation of research programmes, and such dynamics in the scientific community go exactly against this principle. Furthermore, the actual status of string theory demands even more strongly to encourage such a pluralism of theories and approaches. Some string theories speculate that string theory might benefit of new discoveries from other research fields, but the situation just outlined makes very difficult the development of other research programmes; as a consequence, it also goes indirectly against the development of string theory itself.

Lakatos methodology, incentivising the coexistence and competition of different research programmes, explains why the scientific community should encourage their development. This fact underlines how the study of philosophy of science might help the scientific community to stay on the right track. Methodologies, in fact, are constructed trying to explain how science could achieve its success during the ages. Doing this way, they indirectly suggest how to proceed in order to make science continuing to be successful. Besides, theoretical physics failing to give any important result in the last decades suggests that something is going wrong. It is not very clear how much this failure might be due to the impossibility of doing experiments at the Planck scale and how much it might be a consequence of contingent 'external' factors, but the slowing down of theoretical progress is quite evident.

Today, string theory is considered to be the only candidate as a unifying theory. In fact, other approaches to quantum gravity exist, but no other interesting theories of all forces are yet available. For this reason, string theory has no competitors. This might be interpreted as a good or an odd fact. As I will discuss in the next chapter, Richard Dawid thinks that this is not due to all the 'external' factors I have just analyzed, but that this situation suggests that string theory is the only possible unifying theory. This interpretation, based on the already mentioned 'limitations to underdetermination', represents an optimistic viewpoint. In any case, our ignorance should always encourage the development of other research programmes. Dawid's hypothesis is a very interesting one, but Lakatos methodology (and history) rightly suggests to always encourage the proliferation of theories. This suggestion protects the scientific community from wrong hypothesis, making the scientific process to achieve success in any case. Moreover, even if string theory has no competitors as a unifying theory, there are other candidates as a theory of quantum gravity only. String theory, if compatible with them, might benefit from their achievements. We have already seen that such dynamics occurs very often in the history of science. For example, the competition between hadronic string theory and QCD was very fruitful, leading to the former to be integrated in the latter. Also, supersymmetry and supergravity were crucial for the string theory's development[17]. As an example, Loop Quantum Gravity (LQG) one day might be found to be absorbed by string theory, solving its conceptual problems related to space-time and allowing for a background independent formulation.

As already discussed, the consequences of an inaccurate evaluation might contribute to slow down the scientific process, also leading to debates where the different factions fail to find a common ground. The lack of empirical data near the Planck scale not only makes theory building more difficult but also represents the absence of a final judge in the competition between theories. Lakatos methodology helps us to appraise historical steps as rational or irrational and, in general, to evaluate a research programme as a progressive or degenerative one. It gives a so called 'demarcation criterion', providing new crucial interpretations of concepts like 'falsification' and 'scientific honesty'. This methodology, however, does not provide any tool to evaluate a theory as a reliable one or not on purely theoretical grounds. Historically, experiments have always been the only instruments allowed to corroborate or to disprove an experiment; in any case, we have seen that theoretical physics also benefits of a certain theoretical autonomy, with theoreticians not caring so much of experimental results. Finally, the actual experimental situation strongly asks for a methodology providing tools for a non-empirical assessment, that is without empirical considerations; the relative autonomy of theoretical science points to the possibility that such a methodology might exist. Richard Dawid proposes such a methodology for a non-empirical theory assessment. Even if experiments

[17]The GSO projection mechanism, as well as the classification of the different types of string theories, were inspired by work on supergravity

remains the final judge of theories, their absence asks for another instrument to evaluate the reliability of theories. Such an evaluation is crucial because a wrong one would lead to dramatic consequences in the scientific process; an implicit evaluation is always made by researchers or institutions when they have to distribute awards, funds and chairs, and their methodological ignorance might lead to wrong choices, which impact on the scientific development. In fact, even if Lakatos methodology asks for the proliferation of theories, even more in absence of new empirical data, the scientific community cannot handle the development of infinitely many new theories and approaches; scientific research demands for teamwork, and the number of physicists is not really huge, so there is a finite number of research programmes that can be carried on at the same time. For this reason, the scientific community have necessarily to decide which research programmes to promote; maybe, methodological awareness might help to make the correct choices. To this aim, and in the actual situation, a methodology of non-empirical theory assessment might be very useful. Richard Dawid's philosophy might be useful to frame the actual debate in the correct ground, a necessary step to solve the conflict, but its validity as a methodology for non-empirical theory assessment is questionable and deserves to be discussed. A serious analysis is also demanded by this philosophy being used by some string theorists to support their research programme, as it is accounted by Sabine Hossenfelder in her popular book, 'Lost in math' [24]. String theorists know about Dawid's philosophy, arguing in favour of a 'post-empirical' methodology to appraise research programmes. String theory's opponents, on the other hand, look at it as an attempt to change the canonical scientific method in order to save a failed research programme.
This debate and Dawid's philosophy are discussed in the next chapter.

Chapter 4

Richard Dawid's non-empirical theory assessment

In the first part of this chapter I am going to present Dawid's philosophy, giving a summary of his book 'String Theory and General Methodology'[11]; then, in the last part, I will discuss and criticize Dawid's application of his philosophy to string theory.

4.1 Richard Dawid's methodology

As I have already mentioned, contemporary physics is facing a crisis due to the absence of empirical data leading the way. In this situation, the assessment of theories based on theoretical considerations has become increasingly important, and methodological debates arise as a consequence. Scientists opposing string theory stick to the traditional concept of scientific method, that is the necessity of empirical evidence to believe in a theory, while its adepts pretend to defend the availability of string theory on purely theoretical grounds. This scenario underlines the importance of clarifying the conceptual framework in which the debate in embedded. In the work 'String Theory and the scientific method' [11], Richard Dawid tries to provide such a framework, developing a methodology of *non-empirical theory assessment.* Dawid claims that the division string theory brought inside the scientific community is not only based on purely scientific foundations, but rather on a meta-level, that is about general methodology of theory assessment. For this reason a philosophical discussion is necessary to frame the debate in the correct ground. String theory's opponents claim that string theory cannot be considered as 'science', because it does not respect the traditional scientific method based on hypothesis and proofs, but string theorists strongly believe in their research programme, and string theory is the dominant one in theoretical physics research, so we should explain why we face this apparently contradictory situation. Furthermore, considering the lack of experimental data, the scientific community may benefit from a methodology of non-empirical theory assessment.

Here we are going to summarize the main concepts of Richard Dawid's methodology; then, in the fourth chapter, I will focus on its application to string theory and the relationship with Lakatos methodology of scientific research programme.

First of all, Dawid recognizes that, as we have already mentioned, the contemporary debate concerns the scientific method. We can frame the critics in the context of the canonical method: a research programme should reach its completion in a reasonable time, and it must be able to give some predictions that must be tested in a reasonable time as well. In this viewpoint, empirical data have a crucial role, because they are considered the unique judge capable of assessing theories. Also, empirical data are considered necessary for the theoretician to not lose the right track in the theory building process. Without any data, being guided only by theoretical considerations, it is easy for the theoretician to get lost - its detractors think.

We know string theory is not theoretically complete, and it is far from giving testable predictions. for this reason it is rejected by its opponents. Its supporters, on the other hand, claim to defend the theory's viability on purely theoretical grounds, so they implicitly (or explicitly) pretend a modification of theory assessment. After all, no scientific method is God-given, so that claim - even if we may disagree - seems comprehensible. The traditional scientific method has always been successful until now, but the particular status of theoretical physics may ask for a different perspective.

In any case, the two factions are not able to recognize the real ground on which their debate is founded, so they fail to solve the controversy through a rational and constructive discussion. The debate arose because of incompatibility of ideas, and for this reason Dawid feels the necessity to analyze it further. His work, then, explores the role and potential of non-empirical theory assessment in contemporary physics. As we will see, it is founded on the assumption that empirical and non-empirical assessments are equally important. As Laudan puts it in [35], the traditional empirical paradigm does not resemble the real process of science. Scientists base their evaluations not only on empirical grounds, but also and importantly on theoretical considerations. Also, Kuhn claimed that empirical data are always interpreted in a theoretical framework or, better, in a certain paradigm, so that there are no bare data.

The main concepts constituting Dawid's network of arguments are the *limitation to scientific underdetermination*(I will call it LSU), supported by the *no alternatives argument (NAA)*, the argument of *unexpected explanatory coherence (UEA)*, and the *meta-inductive argument from the success of other theories in the research programme (MIA)*; the marginalization of phenomena.

Dawid recognizes three arguments supporting the belief in string theory:

- *NAA*: string theory is the only candidate as a 'theory of all interactions' so far. String theorists claim that it is the only candidate because it is the only one possible unification theory, so there are no alternatives. This argument, alone, is not enough to justify their strong belief. In fact, the lack of alternatives may be due to the limited creativity of scientists, so that they have not yet tried all the ways, or to sociological factor s [1], or other contingent factors. In particular, trying to appraise a theory's status without any experimental data leading the way may lead us to overlook other explanations, other theories. This is what Kyle Stanford called 'the problem of unconceived alternatives', to which such an appraisal should give a solution. In any case, NAA alone is not able to justify the belief in string theory, so this argument have to be empowered.

- **UEA**: this argument strictly links to the Zahar's extension of the methodology of scientific research programmes. When a theory shows to give as a consequence theoretical results not expected at the time of its first formulation, the trust in the theory increases. The explanation of very different phenomena from the original hypothesis would look like a miracle if the theory itself completely fails to describe reality. In any case, this explanatory power may be due to other reasons, for example there may be an unknown principle behind string theory responsible for all these theoretical consequences, a principle independent of the general framework, so that the rest of the theory may be false. In order to empower both NAA and UEA, we need another argument.

- **MIA**: this argument aims at the demonstration of the validity of non-empirical arguments based on an indirect kind of empirical data; this character makes it a 'kind of empirical test on a meta-level'. Concisely, we can state that if looking at history of science we see that the most of theories satisfying NAA and UEA have been in the end verified, it is likely that our theory (in this case string theory) will also be verified when empirical tests will be possible, so that we can trust the theory viability. One example quoted by Dawid is the prediction of the Higgs boson by the Standard Model. He claims that SM was the only one satisfactory model explaining all the experimental data at the time, and it also predicted many novel unexpected facts starting from the attempt to resolve internal theoretical problems (such as the renormalizability of nuclear interactions). Thus, the discovery in the summer of 2012 of the Higgs boson shows that the trust in a research programme satisfying NAA and UEA is justified. In this sense MIA gives empirical evidence to unverified theories based on the corroboration of other theories in a similar research field. As any empirical test, MIA can empower or lower the trust in those

[1] this concerns what we called 'external history', and we will analyze it further.

theories. For example, if the Higgs boson would not be discovered, MIA would have decreased the trust in the methodology of non-empirical assessment, and as a consequence the trust on the theory trusted on its basis. MIA then resembles the role of empirical data for theories not yet verifiable.

These three arguments are strictly related to the fundamental concept on which Dawid's discussion is based, that is **limitation to scientific underdeterminantion.** This concept is also related to the 'problem of unconceived alternatives'. When a theoretician constructs a theory fitting the available data and predicting novel facts, we cannot be sure that he found the correct theory, because there may be other theories fitting data and predicting different facts. This means that "scientific theory building is expected to be significantly underdetermined by the currently available empirical data" [11], and represents the so called 'scientific underdetermination'.
In particular, Dawid considers a precise kind of underdetermination, that is underdetermination based on the available empirical data and some other basic principles typical of scientific theory building, such as the validity of the induction principle, the lack of ad-hoc assumptions, the coherence of the theory, its predictive power and so on, constituting the so called 'ampliative rules'. This kind of underdetermination was first named by Stanford 'transient underdetermination', but it was conceptually different form the one adopted by Dawid, so for this reason he calls it differently. 'Transient underdetermination' was embedded in the canonical paradigm, that is the empirical theory assessment; alternative theories have to be found and empirical tests allow to decide which one of them is viable. In the new sense given by Dawid, 'the degree of underdetermination can be assessed without knowing the alternative theories. Therefore it becomes important to understand the degree of underdetermination in terms of the number of *possible* alternatives, irrespective of the question whether those alternatives are known or not.' This distinction is important because of the contemporary scenario in theoretical physics, where string theory represents the only candidate, so we need to estimate the underdetermination even if we not know any alternative.
Limitation to scientific underdetermination plays a crucial role in the scientific process. If the number of alternative theories fitting a certain set of data and giving different predictions were infinite, it would be very unlikely to find a successful theory, and so to discover new phenomena. On the other hand, looking at history of science, we observe a succession of different theories fitting the available data, so the theories in this succession represent alternative theories. It is then evident that "theoretical progress without scientific underdetermination, to the contrary, would have to be entirely cumulative" (ibid.). For this reasons we can say that scientific underdetermination is present, but it is strongly limited, because otherwise scientific success would be a miracle. Furthermore, scientific underdetermination prevents scientists to believe in research programmes founded on purely theoretical reasonings, so its limitation can increase the trust in such research programmes.

Dawid claims that the assessment of this limitation, while neglected by the canonical paradigm of theory assessment, is of importance both in traditional empirical methodology and especially for a non-empirical theory assessment, which he claims to belong to scientific reasoning.

The three arguments NAA, UEA and MIA are used by Dawid as an instrument capable of limiting scientific underdetermination:

- **NAA:** when scientists are not able to find alternative theories, there can be two explanations. First, it can happen because of some difficulties, for example the objective is very challenging so they need much time to find good theories, or they overlooked some options, or even it is due to facts related to 'external' history. In this case NAA does not support the trust in the research programme. Second, scientists may have not yet discovered alternatives because there are very few alternatives, that is scientific underdetermination is strongly limited. In this second case, NAA supports the theory in question, because if there are very few theories fitting available data it is likely that scientists are developing a viable theory. It is of major importance to understand that the network of arguments we are outlining is founded on this second assumption.

- **UEA:** following Dawid, suppose we have different problems to solve. If a scientists constructs a theory solving one of these problems, he cannot be sure of the viability of the theory because of scientific underdetermination. Also, many theories may exist solving all these problems. It is now that NAA can be used to support the theory. If the scientist in question strongly believes in NAA, he thinks that the theory he found is the only viable one. For this reason, if the theory firstly solves only one or few problems, it is expected anyway to solve all the problems. From this viewpoint, the scientist supporting NAA is not surprised by UEA. If a theory is the only viable one, it is expected to solve all problems, so UEA is a direct consequence of NAA in such a scenario. Then, by reverse reasoning, the scientist observing UEA can state that the NAA he hypothesised increases its probability. In this sense "finding unexpected explanatory power therefor supports the conjecture of limitations to underdetermination. UEA can strenghten the case for NAA" (ibid.).

- **MIA:** first, Dawid argues that scientific success implies limitation to scientific underdetermination. If scientists find a certain theory explaining the available data and predicting novel facts, they cannot believe with certainty in the viability of their theory because of scientific underdetermination. In any case, if the theory is corroborated by experiments, we should answer how is it possible that we found a viable theory among the many possible theories fitting data and giving different predictions. In other words, we must explain why science is so successful. The

best answer we can give is that scientific underdetermination is strongly limited, so scientists are likely to find viable theories in the theory building process.
Second, we use MIA to argue that "regular predictive success in a research field justifies the assumption that future predictions of similar kind will be correct as well" (ibid.). We have just claimed that limitation to scientific underdetermination is supported by NAA and UEA, and it explains scientific success. Thus, we can make an assumption based on MIA, a 'meta-inductive inference': if theories satisfying NAA and UEA are regularly verified, we can expect that this will happen for another theory satisfying NAA and UEA as well.

The web of reasonings is complete: UEA supports NAA supporting limitation to scientific underdetermination and the latter explains the success of science, so that MIA makes the inductive step generalizing scientific success to all theories satisfying NAA and MIA.
This strategy of theory assessment, as I have already asserted, can be considered to belong to scientific reasoning, because it contains an obvious 'empirical basis', that is it presents potential falsifiers able to decrease the trust in the strategy itself. This cases are not difficult to identify.
For example, if we trust the viability of a theory because of NAA and then we find an alternative, both the theory and NAA - and the whole strategy as a consequence - will be weakened.
The same occurs if we find that UEA is due to a subtle principle underling the theory under consideration, so that UEA was misguiding and does not supports the theory anymore.
MIA, on the other hand, can be weakened if theories expected to achieve predictive success based because they satisfy NAA and UEA are finally falsified. In this case the inductive inference results to be wrong and trust in the non-empirical strategy would decrease importantly.
In general, we can say that every time a theory is expected to be probable because of the above arguments and it finally results to be wrong, the trust in the strategy of non-empirical assessment decreases, while it increases when such theories are finally verified. It resembles quite well the scientific process, because it makes predictions that can be verified or falsified. The difference is that this is made at a meta-level, that is empirical data are nothing else that research programs and their development. For this reason we can say that non-empirical theory assessment belongs to scientific reasoning.
At this point it is important to specify further the role of scientific underdetermination in high energy particle physics. Experimental particle physics is characterized by the energy scale accessible by experiments. Also, theories in this research field are characterized by a certain range of energy where they are assumed to represent a good description of the microscopic world. For this reasons, we can say that the success of a certain theory considered viable in a certain range of energy can be viewed as a consequence of scientific

underdetermination *at that energy scale*. This success can then justify the assumption of a limitation of alternatives at higher energy scales through the meta-inductive inference. Dawid underlines that "attributing a low probability to the existence of alternatives giving different predictions at the next stage of empirical testing is fully consistent with the expectation that an infinite sequence of ever higher energy scales lies beyond the next empirical step". This is because many theories might exist that have the theory under examination as their effective theory at the scale it is considered to be valid. For this reason, predictive success can justify the claim of a limitation of theories *giving different predictions at the next step of empirical testing*, but not in general the limitation of alternatives at every energy scale. We can see different theories at higher energy scales as the same theory at lower energy scales if they do not predict different facts at the next generation of empirical tests. Then, scientific underdetermination is not limited *globally*, but only *locally*, that is we can assume that few theories are predictively different at the next stage of empirical testing, but not that few theories exist in general, this because more generic theories containing local theories as their effective low-energy descriptions can exist, and we cannot recognize them as different theories at the next stage of empirical tests.

For theories valid only in a certain range of energy, NAA does not imply that the theory under consideration is the 'final theory', because limitation to scientific underdetermination is local in such a case; it might be the only viable theory at its characteristic energy scale, but there may be different theories also valid at higher energy scales having that theory as their low-energy limit. But if the theory under consideration is universal, that is if it is assumed to be valid at any energy scales, we see that NAA directly implies that this theory should be the 'true theory', that is it supports a 'final theory claim'. In fact, no other more universal theories exist in this case, so NAA implies that it would be the only possible candidate. Now, the problem is that "the empirical corroboration of NAA must be based on MIA, which in turn can only establish local limitations to scientific underdetermination" (ibid.). In other words, in order to explain scientific success one must assume that at each step of empirical testing there are few theories giving different predictions, so NAA can be used at each step. This does not allow to assume NAA in general, because while at a certain step there might be only few different theories, many theories might exist having it as an effective theory and being predictively different at a certain higher energy.

Dawid shows that limitation of scientific underdetermination can be used to extend the validity of the 'inference to the best explanation' (IBE) to non-testable theories. In fact, IBE historically belongs to the canonically paradigm of theory assessment. IBE represents a selection process where scientists firstly collects a set of possible explanations for a certain problem, and then choose among them the one they consider is best. In order to believe in the correctness of this procedure, scientists assume that 'the true theory does not belong to the set of unconceived alternatives', a threat that can be

disregarded if the best theory they selected seems very convincing. The assumption of having not overlooked the true theory evidently contains an element of assessment of scientific underdetermination. The selected theory, anyway, can be trusted only if empirically verified in the canonical paradigm.
Dawid claims that IBE can be extended to a situation where we only have one candidate and we cannot yet empirically verify it. The line of reasoning implies a double utilization of IBE.
First of all, having only one theory, referring to Alexander Bird one may talk of 'inference to the only explanation', that is we trivially choose the only one option as the best theory explaining data. But this theory cannot be trusted because its prediction cannot be tested. So we need another step to increase trust in the theory. If we were able to assess a strong limitation to scientific underdetermination, we may increase the trust in the viability to the theory. In order to assess this limitation, we consider the three arguments already mentioned. If we cannot explain NAA, UEA and MIA in any other way than through assuming a limitation to scientific underdetermination, we can use IBE in order to infer the viability of the argument. Once we used IBE to justify the assumed limitation, we can use the latter to justify the trust in the theory itself.
This procedure extends the validity of IBE also to the non-canonical paradigm, that is the non-empirical theory assessment paradigm. Furthermore, it underlines the importance of the second-level IBE also in cases where empirical confirmation is available. In fact, whenever scientists use IBE to select a theory, whether verifiable or not, they implicitly use the second-level IBE to assess limitation to scientific underdetermination to justify the choice they made in the first IBE. This argument shows the general importance of LSU for theory assessments, not only for the non-empirical methodology.
Here a deep discussion about the 'marginalization of phenomena' would take us too far, so we only give a hint. Using this term Dawid indicates the marginal role assumed by phenomena in modern physics. He recognizes five different kind of marginalization:

- *Marginalization of the micro-physical phenomena in an observational context:* today the objects we study can be analyzed only via complex experiments and appear as lines or signatures when extracted by accelerators and detectors, losing the possibility of a direct observation.

- *Marginalization of the phenomena in connecting theory to experiment:* experiments are even more imbued with theory, needed in order to interpret experimental results. Here we can identify an element of LSU: experimental results are interpreted using well-established theoretical concepts, and we choose the best explanation implicitly assessing a limitation to possible alternative explanations.

- *Marginalization of the phenomena in concept formation:* in the last century we discovered that to achieve a good description of the microscopic world we should abandon naive ontology, adopting abstract mathematical concepts like the wave function to get correct predictions. In general, physical concepts are increasingly 'theoretical'.

- *Marginalization of the phenomena in theory dynamics:* until the end of the last century, theory building was driven by the huge amount of experimental data, experiments having a crucial role in the process. Today theoretical physics seems to be much more independent of phenomena. Theory building is carried on by trying to solve conceptual and technical problems and not to account for data. The necessity to go beyond the Standard Model is not due to some counter-evidence, but mainly to conceptual problems.

- *The trust in theoretical conceptions is increasingly based on theoretical considerations:* the role played by non-empirical theory assessment in the evaluation of research programs indicates the important role that theoretical considerations have received at the expense of the increasingly difficult experiments as a means to assess the viability of theories.

This scenario shows that theoretical considerations, and so non-empirical theory assessment, has always been made also in the process of empirical confirmation, so it already belongs to the traditional paradigm of theory assessment.
Dawid also presents paleonthology as an example of a special science where LSU and non-empirical theory assessement are fundamental because of the lack of new empirical data; he argues that contemporary physics is facing a similar situation to special sciences and maybe the role of non-empirical theory assessment is going to be increasingly important also in this context.
Dawid states that string theorists are only strengthening this element because of the actual situation in contemporary physics, while its opponents do not recognize it as scientific reasoning, but evaluates string theory from their different traditional paradigm.
To conclude, the role played by MIA shows that non-empirical theory assessment depends indirectly on empirical evidence. For this reason Dawid does not claim that his strategy can be the final judge of a theory on its own, or that it can replace empirical confirmation, but it is always secondary. Even if we cannot attribute the same status to both strategies, we can recognize that not so much difference lags between the two. Scientists often base their beliefs on theoretical reasonings rather than empirical data, and furthermore empirical data contains a certain amount of theory. For these reasons, while accepting empirical confirmation as the final judge because of its important heuristic role in the scientific process, one can return to theoretical considerations the value they deserve.

4.2 A methodological debate

Before discussing and criticizing Dawid's application of his own philosophy to string theory, I would like to deepen his analysis of the debate around this topic. Dawid argues that this debate is due to physicists having different methodological backgrounds, failing to find a common ground to frame and make constructive the debate. In particular, he argues that string theory opponents are stuck to the canonical scientific methodology based on theory building and experimental proofs, while string theorists recognizes that a theoretically based appraisal is needed. Dawid also shows that theoretical considerations are always at work in experimental physics (calling this fact 'marginalization of phenomena') and that, while experiments are always primary, non-empirical evaluations play an important role that should be recognized. A new methodology of this kind seems to be necessary due to the lack of new experimental data, but this topic is a subtle one. One might argue that string theorists are trying to save their theory from its inability to predict novel verifiable phenomena by asking for a review of the scientific method. Its opponents, on the other hand, claim that such a review would mine the functioning of the scientific process. In order to analyse this situation, one has to understand how much is true that theoretical assessments are almost as important as empirical tests, and how much Dawid's philosophy in particular represents the correct tool to give such theoretical assessments. To this aim, I will analyze Dawid's philosophy from my own perspective and also in the light of Lakatos methodology. Before, I give a brief account of the main arguments in favour and against string theory we can find in the popular literature, explicitly looking to the methodological assumptions made by the authors.
The main argument against string theory, that we can find in both [63] and [52], is the impossibility to falsify string theory. But - and this is crucial - 'falsify' in the naive sense. They understand falsification as it was firstly proposed by naive falsificationists, but I have already discussed that it was not a good falsification criterion. They fail to understand Lakatos' teaching, remaining stuck to an obsolete and inexplicably still popular concept. As an example, in [52] we can find "It does not seem that world has twenty-five spatial dimensions. Why the theory was not immediately abandoned is one of the great mysteries of science". I think this quotation is representative of a naive line of reasoning based on a popular but outdated philosophy of science. Lakatos taught us that all research programmes develop in an 'ocean of anomalies', trying to resolve anomalies following the positive heuristic. The above mentioned statement makes evident a methodological ignorance which permeates these books[2]. 'Falsification' does not mean disprove by experiments, but rather a research programme being superseded by a more heuristically powerful one, and this process might happen also in absence of experimental

[2]and also the presence of a pinch of hubris, because string theorists were well aware that our world has not 25 dimensions

proofs. Naive falsification is a useless demarcation criterion, and especially in the actual situation, where experimental physics cannot investigate the Planck scale in order to test quantum gravity theories. Another criticism related to experimental testability regards the risk to lose the right track that one runs when developing a research programme on theoretical grounds only. Testable predictions are demanded because they represents necessary constrains guiding the theory building process. In any case, testability is a sistematic problem in today quantum gravity research, not only with respect to string theory.

Other criticisms are related to the theory itself, and I think they are not respectful of the 'scientific expert principle' (the same claim is made by string theorists, as accounted by Dawid, [11], p.24). For example, Smolin (and Carlo Rovelli as well) always remind that string theory is not background independent, so it fails to incorporate Einstein's teaching about the dynamicity of spacetime. But it is quite obvious that string theorists know about this. Rickles reminds how already in 1988 Gary Horowitz underlined the crucial importance of a background independent formulation of string theory [23], also stating that "the problem of background independence was understood early on by the string theory community in a way that matches the way it was understood in the canonical quantum gravity community and imbued in just as much importance" ([49], p.198-199). I think arguments of this kind are not constructive, because underline something already well known by the experts in that field, so they have the only scope of doing propaganda against the enemy. String theory opponents, if they really want to respect the 'scientific expert principle', should not criticize the research programme itself, but only the unhealthy dynamics characterizing the scientific community. The really interesting parts of Woit's and Smolin's book are the last, where they both criticize the monopoly of scientific research by string theorists. This is a constructive criticism. In observance of the 'scientific expert principle', one should not criticize what other experts are doing in a different research programme, but all the scientists have the right to ask for the preservation of a democratic community, with no unfair treatments.

Finally, string theory opponents usually move criticisms against the anthropic principle, saying that it is not a scientific principle because it does not give testable predictions, and the existence of a Megaverse, a non testable assumption.

I remind you that the weak anthropic principle simply states that our existence is an evidence that must be taken into account.

Susskind [58] calls critics like Smolin 'Popperazi', underlying their insistence on the concept of (naive) falsification and complaining of strict philosophical rules given by philosophy, such as a demarcation criterion defining what is science and what is not. I have already discussed that Smolin (whom I have selected as a paradigmatic example, as well as Susskind) is effectively a 'Popperazo', and I might understand what Susskind means referring to strictly philosophical rules. Popper's naive falsificationism not only fails to give a rational reconstruction of history of science, but also shows a normative character which might appear as hubris to scientists. In fact, this demarcation criterion

imposes to scientists to construct theories proposing an empirical basis capable to falsify it, claiming that a theory avoiding this process should be called pseudo-scientific. But philosophy of science should not order to scientists what to do. Lakatos methodology, in addition to propose a workable concept of 'falsification', has not a normative character, only suggesting a definition of scientific honesty.
In any case, Susskind defends the anthropic principle pointing to the successful predictions it was able to make, such as the calculation of the cosmological constant's value by Weinberg and the existence of an excited state of Carbon-12 in the process of stellar nucleosynthesis. I think that the anthropic principle, as these cases show, is nothing more than an estimation or application of inductive bayesian logic supported by evidence, this evidence being our own existence. In my opinion, there is no principle, nothing mysterious, only ourselves used as evidence to support a certain statement. The anthropic principle plays an interesting role only when it is used to support the existence of the Megaverse. In fact, Landscape researchers argue that the existence of many other universes is the only fact able to explain why our Universe is so life-friendly. They argue that the incredible fine-tuning of parameters we observe is necessary for our existence, and if one does not assume the existence of the Magaverse in order to explain such a situation as a statistical result one is forced to believe in the project of a Superior Intelligence, an Intelligent Design.
Smolin counter-argues by considering the example of planets: knowing that many other planets exist, we can rightly say that our existence is a matter of statistics, because it is not surprising that at least one among a huge number of planets satisfy the conditions to support life, but if we talk about the Universe, we have no evidence of other universes, so we cannot say that our universe being able to support life is a matter of statistics as well, because we cannot assign any probability having a sample of one, and such a fact might be due to other reasons. Rickles ([49], p.232, note 56) notices that this line of reasoning might be inverted to argue in favour of the existence of other universes: such as the mystery of Earth being life-friendly can be solved assuming the existence of many other planets, our Universe being life friendly might be solved through analog reasoning. Now, we can appreciate once again the great advantages given by Lakatos methodology. This methodology, as already discussed, as a great virtue: it allows to appraise a metaphysical statement in the same way one can appraise a theoretical statement. It is evident that the difficulties to solve the debate around the anthropic principle and the existence of a Multiverse are due to their metaphysical character, something very difficult to handle by other methodologies. In a lakatossian optics, the line of reasoning is simple: one is allowed to assume the validity of a metaphysical assumption and retain it until it is superseded by a better hypothesis, giving an excess of empirical content. Then, if neither the assumption of the existence of a single Universe nor the assumption of the existence of many universes give a progressive problemshift, they should be considered in equal foots. Landscape researchers claim that the existence of a Multiverse is the best explanation of our existence, so the anthropic principle is used to support such

an assumption. But now this usage is very different from the above examples: it is not used to explain an observable phenomena, but to predict an unprovable fact with the only aim of explaining our existence. Life is not used as an evidence constraining something, as a proof in favour of a certain phenomena, but as something to be explained. But as Smolin notices, we have not only two options, the Megaverse and the Intelligent Design, but three, because there might exist an unknown reason explaining why the Universe is life-friendly, allowing for our existence. If the existence of a Megaverse does not constitute a progressive problemshift, it cannot be considered more probable than the existence of only one Universe. And it really does not represent a progressive problemshift: assuming the existence of a Megaverse does not lead to the prediction of novel facts or to the explanation of novel facts, being our own existence not a really surprising unexpected fact. It only raises the probability of a Universe being able to support life, a probability which we cannot assess because of the lack of more samples than one.
As mentioned, the existence of many universes is considered also a solution to problems related by naturalness. For example, 'hierarchical naturalness' would be explained claiming that the values we observe are necessary to support life, and we live in a Universe - among infinitely many others - that supports life, so it is explained why they have such values. In any case, naturalness does not constitute a real problem, and is a debated and ambiguous issue, as it is discussed by Hossenfelder in her book [24].
To conclude the discussion of this particular topic, I would like to frame the entire situation in the context of the methodology of scientific research programmes. Briefly speaking, the Landscape programme represents a degenerative research programme, so the assumption of a Multiverse based on string theory, eternal inflation and the anthropic principle cannot be considered as acceptable; it cannot be considered as valid not even by its own, because it failed to give any substantial result. If one day the hypothesis of the existence of a Multiverse will lead to a progressive research programme superseding the dominating one, the Multiverse should be considered as possible. This is how the methodology of scientific research programmes appraises not only theoretical statements, but also metaphysical ones. The instrumental character of this methodology might be very useful to solve such ambiguous debates around metaphysical questions; concisely, one might say that scientists adopt those assumptions giving better results than others, of whatever kind they are.
This brief discussion should have stressed the main issues around which the debate is constructed. I had argued how the debate around the Landscape and anthropic principle will end automatically when it will be found to be successful or when it will be superseded by a better research programme. But, obviously, this process will occur only if proliferation of theories is encouraged, leading to new research programmes able to make science proceed forward. In my opinion, string theory's opponents should concentrate their criticisms on the fundamental issue represented by a democratic scientific community, when no research programmes are allowed to dominate and other approaches are stimulated. As already mentioned, Dawid's philosophy is considered (mainly by string theorists) to

provide a useful tool for scientists to understand which research programmes are the most promising, and if there might be other unknown approaches to be investigated.
Finally, I am in agreement with Dawid claiming that the main problem is the failure to frame the debate on correct grounds. I have mentioned how string theory opponents are stuck to naive falsificationism, while string theorists recognizes the importance of a non-empirical theory assessment, viewed as a natural strengthening of an already used element in theory assessment due to the actual experimental situation.
In my opinion, this debate could be easily solved by a common understanding of Lakatos methodology, that is of the real functioning of science. In any case, Dawid proposes a new methodology which might help evaluating research programmes in absence of experimental data, and might also be integrated with Lakatos methodology. In any case, we should understand to what extent his methodology might represent a reliable instrument.

4.3 A string theory non-empirical assessment

In the first chapter I presented the web of reasonings holding up Dawid's philosophy. Summarizing, Dawid argues that NAA, UEA and MIA justify to assume that scientific underdetermination is strongly limited, this explaining why science is so successful. An assessment of limitation to scientific underdetermination (LSU) might also constitute a powerful tool in order to give a non-empirical appraisal of a certain research programme. Looking back to details in these arguments, it is evident that if the three arguments supporting LSU would be proved not to be reliable, the entire web of reasoning and its consequences would break down. For this reason, I think criticism to Dawid's philosophy may be addressed to criticism to these three arguments, which I am going to comment.
NAA, I remind you, is the 'No Alternative Arguments'. String theory is the only candidate as a unifying theory so far, meaning that no alternatives are available at the moment. One might explain this situation addressing NAA to 'external' factors or to the lack of imagination by scientists, while someone else might say that we found only one possibility because scientific underdetermination is strongly limited, so it is very difficult to find good theories. Moreover, if we have NAA because very few theories exist fitting data (so, because of LSU), the probability of our theory being a viable one would obviously increase. String theory represents a particular case to be discussed in these terms. Scientific underdetermination means that available data do not constrain completely theory building, so that different theories might exist fitting data and predicting different novel facts. Now, saying that string theory is the only possible candidate because of a strong LSU, means that one is assuming string theory can explain available data, because of the very definition of scientific underdetermination. But this assumption is wrong. String theory has not demonstrated to be able to explain the available

data, so scientific underdetermination loses its significance. Furthermore, we do not even know if a consistent string theory really exists. String theory, up to now, resembles more a mathematical framework based on unproved conjecture rather than a real physical theory. Obviously, a string theorists might assume that string theory will certainly be able, one day, to explain data, but this leads to two problems. First, the usefulness of such a methodology would lose reliability, because of its subjective character. Second, this kind of reasoning might be made by all other theoreticians working on research programmes not yet available to explain data, so that scientific underdetermination would lose its meaning, because many scientists would claim to have found the correct theory. The question to answer was: how is it possible that we only have one candidate as a unifying theory? I would say that string theory being a candidate, because of its inability to explain available data, is not linked to scientific underdetermination, and cannot gain support by LSU. One may also wonder how much 'external' factors contributed to this situation, where no alternative seems to be possible. I have already mentioned such factors, referring to Woit's and Smolin's books. I think these factors had (and still have today) a huge influence on the scientific process, and strongly contributed to string theory being 'the only game in town' ([63], chapter 16). In any case, a detailed discussion of this issue is not necessary in this context, because I was only concerned with showing that NAA does not support LSU, being scientific underdetermination inapplicable.
UEA represents the 'Argument of Unexpected Explanatory Coherence', similar to the progressive problemshifts in the Zahar sense. Briefly summarised, Dawid claims that UEA supports NAA because if one believes a theory being the only possibility and finds that such a theory is able to explain many different problems, one is led to believe that this theory is really the only possible one. Again, this line of reasoning fails to be applied correctly to string theory. In fact, there is a subtle difference between progressive problemshifts in the Zahar sense and UEA. Progressive problemshifts explains the rationality behind the scientists' choice to carry on a certain programme, even if it may be wrong. In string theory, such problemshifts simply suggest that string theory might be correct, pushing forward the research programme. For example, The anomaly cancellation result, being managed using a promising phenomenologically gauge group, represented a progressive problemshift, because of this consequence, but it only suggested that string theory might *potentially* be able to recover the Standard Model. UEA, on the other hand, is based on the consideration that a theory resolving different problems strengthen its reliability. The problem is that string theory, up to now, has not been able to solve any problem, it only representing a set of promises. The only one objection one may advance is the calculation of the Bekenstein entropy. I agree that this is an incredible result, but we should be careful, both because it has been performed only for very idealized cases and because it might be due to an underlying unknown principle, which might be independent of string theory. In fact, one of the biggest threats to UEA are represented by the existence of such principles, because the unexpected interconnection would be addressed to an underlying principle rather than to the theory, thus taking away its

merits of success; the apparent 'miracle' would be explained 'from the outside' of the theory, not anymore supporting NAA. This might constitute a real threat, and Dawid himself reminds that Strominger speculated that the calculation of black holes entropy might be performed by using general principles, with no use of string theory at all ([11], p.57). For these reasons, even if one wishes to take the Bekenstein entropy calculation as a clear unexpected success of string theory, one cannot immediately take it as a support for NAA (and for LSU).

MIA represents the last element of the web of reasoning, allowing to assess LSU applying the above two concepts. Without MIA, in fact, they are not enough to show that NAA and UEA represent reliable and workable assumptions. MIA works on a meta-level, meaning that the the reliability of the theory at issue is assessed by looking at other similar theories' successes. If looking at history of science we can say that research programmes satisfying NAA and UEA are usually found to be successful, we might generalize this scenario claiming that a theory satisfying NAA and UEA has a great probability to be correct. It is evident why MIA is called a 'meta-inductive argument': it generalized a certain observation to other similar situations. The acceptability of this inductive logic is not trivial. Firs of all, there is the risk of a 'historical revisionism'. There have been few times in history of science when a research programme was considered to be the only one possibility before empirical testing. For example, I have already discussed how many people thought that the S-matrix theory was the only one approach able to investigate the strong interactions because QFT, being based on perturbative methods, seemed to fail to account for strong forces. Dual models also satisfied UEA, because they led to a microscopical picture represented by hadronic strings that was completely unexpected, going also against the very bootstrap philosophy, and potentially providing solutions for difficult problems, such as confinement. This assumption, later, was found to be wrong when QCD took the place. This example shows that not all the research programmes satisfying NAA and UEA are successful. One might argue that such research programmes are successful in the majority of cases, but I think this claim is not so trivial. In fact, there is the risk of confirmational bias, that is to look only at historical examples corroborating the claim itself, overlooking counter-examples. Also, history of science is often a 'teleological' one, so bankruptcy research programmes risk being kept out of the narrative. I would like to underline that I am not arguing against this assumption, but I only think that Dawid does not give a sufficient justification for such a strong assumption.

Another example, already mentioned, is given by the Standard Model and the prediction of the Higgs boson. One may say that the Higgs boson had to be found because the Standard Model satisfied NAA and UEA, but this example cannot be generalized to the string theory case. In fact, when the Higgs boson was predicted, the Standard Model was already a well established theory, corroborated by experimental results. Physicists believed in its prediction not because it represented the only possible explanation, but because experiments had already showed it to be a reliable and successful theory. The

same cannot be said for string theory, which has never been tested, and in fact it is not considered universally reliable.
Summarizing, MIA does not represent a 'legal' inductive reasoning, because of three reasons:

- there are examples in history when a research programme considered to be the only one possibility was found to be not really the only one possible explanation, so the assumption of NAA was not correct;
- not all theories satisfying NAA and UEA were found to be successful; assuming that they are successful in the majority of cases is not a trivial assumption, so it needs a better justification;
- many of the successful examples one can found cannot be used for string theory, because they refers to already well corroborated theories, so that their predictions were considered very probable only because they were considered reliable on empirical grounds.

In particular, the first argument, together with the example given by dual models and QCD, shows a possibility not accounted by LSU. In fact, S-matrix theory and QCD represent very different approaches, with the former strongly philosophically based. LSU is unable to account for the existence of such different approaches; in fact, QCD represented an 'unconceived alternative' when S-matrix theory was considered to be the only possible one. QFT was considered to be unable to describe hadronic physics, and the bootstrap philosophy represented in some sense the renouncement to explain some features of hadrons, such as the large number of particles that was found. The assumption that no possible explanation of these features could exist, was found to be completely wrong.
I think this situation might be considered analog to the one concerning the Landscape programme. Landscape researchers think that the incredible fine-tuning of parameters we observe in our Universe, necessary for life, cannot be explained if not assuming the existence of other Universes. The above example should suggest that the 'problem of unconceived alternatives' cannot be easily addressed by LSU. Maybe, Landscape researchers will end up like bootstrappers, thinking that no other solution was available when there was, on the contrary, an even better physical solution.

I have argued that NAA, UEA and MIA do not represent strong enough arguments for an evaluation of limitations to scientific underdetermination, and that they are not applicable to the string theory case. String theory has a very ambiguous epistemological status, making it difficult to give an 'absolute' assessment. I think that, given the situation, Lakatos methodology is the only one able to provide a correct interpretation. In fact, as already discussed, the great virtue of this methodology is its relativist and instrumental character: research programmes are evaluated only with respect to other research programmes, avoiding the difficulties related to an 'absolute' appraisal. In this optics, the estimation of LSU is not considered, but it is something that automatically emerges from the proliferation of theories. A research programme is indeed 'falsified' when another research programme is able to account for its successes and also predicts novel facts, that sooner or later are verified. This process implicitly contains the concept of scientific underdetermination, because a new research programme being able to explain another programme's results and also predicting novel facts means that theory building is underdeterminated, and new progressive theories emerge from this underdetermination. Lakatos methodology is not concerned with the evaluation of the degree of this underdetermination, but simply suggest the proliferation of research programmes as a means to make such theories emerge. In this viewpoint, one might explain scientific success not assuming the validity of LSU, but assuming a mechanism of 'natural selection'. In fact, if many research programmes can be developed starting by the same empirical data, competition would automatically select the best ones, leading to increasingly better theories. In any case, I think that in some cases LSU can describe correctly the situation, but only in retrospective. If we can identify in history of science a successful research programme developing with no competitors, we might address this success to LSU. This means that LSU might be a useful concept to be integrated in Lakatos methodology, but cannot work alone. I have already discussed that MIA cannot be used to argue in favour of a certain research programme, and that also the other arguments suggesting LSU are not based on solid grounds. Dawid's philosophy might be used together with Lakatos methodology to achieve more information for a rational reconstruction of historical facts, but using it for a non-empirical assessment of not yet corroborated research programmes represents a too risky game. I suggest, one more time, that encouraging the proliferation of theories is crucial to achieve scientific success. In fact, in addition to previous arguments favouring it, I also showed how it helps to avoid both the problem of unconceived alternatives, and an evaluation of LSU, through a mechanism of 'natural selection' by competition. In any case, looking at the above arguments, one may find some cases for which Dawid's philosophy may be applicable, so I do not want to say that it is a completely useless methodology. Furthermore, it may also help to shred lights on many problematic issues, so it deserves further reflections and investigations. Maybe, an empowered version of Dawid's philosophy would provide a useful tool for non-empirical theory assessment, something strongly demanded by the actual dramatic situation that theoretical physics is living.

Conclusions and future developments

In this work I used Lakatos' methodology to appraise the string theory research programme. Such an appraisal, given the ambiguous epistemological status of string theory, is not trivial, so I do not pretend my personal interpretation to be unquestionable. String theory led to a heated debate in the scientific community that has lasted for years, as accounted by many popular books. In my opinion, this debate has always failed to be a constructive one, often taking the form of mere propaganda or unproductive criticism. Even if many physicists reject the importance of philosophy of science for the scientific development, one can easily find philosophical arguments to be used in this context. In fact, this debate is mainly about methodology rather than physics, but different perspectives make scientists failing to frame it on correct grounds. In particular, it seems that scientists are stuck to outdated concepts; maybe, a better understanding of the real functioning of the scientific process might help to solve the debate or, at least, to make it productive. To this aim, I hope this work has been able to show the great virtues of the methodology of scientific research programmes, suggesting to the scientific community not to neglect meta-physical reflections. This analysis, in addition to provide an interpretation of a much discussed topic, also offers suggestions for breaking the stalemate theoretical physics has been undergoing for decades, such as encouraging the proliferation of research programmes. This methodology has its greatest virtue in its relativist character, able to avoid useless debates related to an absolute evaluation.
A major result of this work is the clear identification of two different string theory research programmes, born from two different ways of responding to the same problem, and their methodological evaluation. In particular, I have put forward many arguments explaining why the 'standard string theory research programme' cannot be considered neither a progressive nor a degenerative one, and why the 'landscape research programme' might be appraised by many scientists as a degenerative one.
These conclusions might lead scientists to accuse me of arrogance, but I would like to remind that my objective was to give experts food for thought, and not to take their place. I tried to carry on an analysis as objective as possible, trying to serve as a super-partes

councillor applying declared philosophical tools. Reminding an already given quotation:

> The statute law approach should become much more important when a tradition degenerates or a new bad tradition is founded. In such cases statute law may thwart the authority of the corrupted case law, and slow down or even reverse the process of degeneration. When a scientific school degenerates into pseudoscience, it may be worthwhile to force a methodological debate in the hope that working scientists will learn more from it than philosophers.

Regarding the future, I can say that philosophy of physics will be increasingly important. Richard Dawid's methodology, even if questionable, clearly shows that the actual situation has lead scientists to rely on philosophy to justify their work. Maybe, a better philosophy will be proposed able to solve the debate and making research programmes to develop on solid methodological foundations, also in absence of experiments. Up to now, in any case, such a methodology is not available, so we have to be content with the teachings of Lakatos, whose depth should not be underestimated.
My hope is that scientists will be able to restore an healthy scientific environment, suitable for scientific development. To emerge from this crisis, they are called upon to be open to confrontation, possibly mediated by philosophers of science, who can provide useful tools for a deeper meta-physical analysis.
Because, as Lakatos stated provocatively,

> *most scientists tend to understand little more about science than fish about hydrodynamics.*

Bibliography

[1] JOSEPH AGASSi. “How are Facts Discovered?” In: *Impulse* 3.10 (1959), pp. 2–4.

[2] Nathan Benjamin Agmon et al. “Lectures on the string landscape and the Swampland”. In: *arXiv preprint arXiv:2212.06187* (2022).

[3] Joseph J Atick and Edward Witten. “The Hagedorn transition and the number of degrees of freedom of string theory”. In: *Nuclear Physics B* 310.2 (1988), pp. 291–334.

[4] Tom Banks et al. “M theory as a matrix model: A conjecture”. In: *The World in Eleven Dimensions: Supergravity, Supermembranes and M-theory*. CRC Press, 1999, pp. 435–451.

[5] K Bardakci and S Mandelstam. “Analytic solution of the linear-trajectory bootstrap”. In: *Physical Review* 184.5 (1969), p. 1640.

[6] Curtis G Callan Jr and Juan M Maldacena. “D-brane approach to black hole quantum mechanics”. In: *Nuclear Physics B* 472.3 (1996), pp. 591–608.

[7] Philip Candelas et al. “Complete intersection calabi-yau manifolds”. In: *Nuclear Physics B* 298.3 (1988), pp. 493–525.

[8] Philip Candelas et al. “Vacuum configurations for superstrings”. In: *Nuclear Physics B* 258 (1985), pp. 46–74.

[9] E Cremmer and Joel Scherk. “Spontaneous compactification of extra space dimensions”. In: *Nuclear Physics B* 118.1-2 (1977), pp. 61–75.

[10] Jin Dai, Robert G Leigh, and Joseph Polchinski. “New connections between string theories”. In: *Modern Physics Letters A* 4.21 (1989), pp. 2073–2083.

[11] Richard Dawid. *String theory and the scientific method*. Cambridge University Press, 2013.

[12] Sebastian De Haro et al. “Conceptual analysis of black hole entropy in string theory”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 69 (2020), pp. 82–111.

[13] Lance J Dixon and Jeffrey A Harvey. “String theories in ten dimensions without spacetime supersymmetry”. In: *Nuclear Physics B* 274.1 (1986), pp. 93–105.

[14] A Font et al. “Strong-weak coupling duality and non-perturbative effects in string theory”. In: *Physics Letters B* 249.1 (1990), pp. 35–43.

[15] S Fubini and G Veneziano. *LEVEL STRUCTURE OF DUAL-RESONANCE MODELS.* Tech. rep. Massachusetts Inst. of Tech., Cambridge, 1969.

[16] Doron Gepner. “Exactly solvable string compactifications on manifolds of SU (N) holonomy”. In: *Physics Letters B* 199.3 (1987), pp. 380–388.

[17] Peter Goddard et al. “Quantum dynamics of a massless relativistic string”. In: *Nuclear Physics B* 56.1 (1973), pp. 109–135.

[18] Michael B Green. “Space-time duality and Dirichlet string theory”. In: *Physics Letters B* 266.3-4 (1991), pp. 325–336.

[19] Michael B Green and John H Schwarz. “Anomaly cancellations in supersymmetric D= 10 gauge theory and superstring theory”. In: *Physics Letters B* 149.1-3 (1984), pp. 117–122.

[20] Michael B Green, John H Schwarz, and Edward Witten. *Superstring theory: volume 2, loop amplitudes, anomalies and phenomenology.* Cambridge university press, 2012.

[21] David J Gross et al. “Heterotic string”. In: *Physical Review Letters* 54.6 (1985), p. 502.

[22] DJ Gross. “Strings at superPlanckian energies: In search of the string symmetry”. In: *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 329.1605 (1989), pp. 401–413.

[23] Gary T Horowitz. *String theory without space-time.* Tech. rep. 1988.

[24] Sabine Hossenfelder. *Lost in math: How beauty leads physics astray.* Hachette UK, 2018.

[25] James Hughes, Jun Liu, and Joseph Polchinski. “Supermembranes”. In: *Physics Letters B* 180.4 (1986), pp. 370–374.

[26] Lars-Göran Johansson and Keizo Matsubara. “String theory and general methodology: A mutual evaluation”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 42.3 (2011), pp. 199–210.

[27] Shamit Kachru et al. “De Sitter vacua in string theory”. In: *Physical Review D* 68.4 (2003), p. 046005.

[28] David Kaiser. “Drawing theories apart”. In: *Drawing Theories Apart.* University of Chicago Press, 2009.

[29] Keiji Kikkawa and Masami Yamasaki. “Casimir effects in superstring theories”. In: *Physics Letters B* 149.4-5 (1984), pp. 357–360.

[30] Ralf Krömer and David Corfield. “The duality of space and function, and category-theoretic dualities”. In: *Siegener Beiträge zur Geschichte und Philosophie der Mathematik* 1 (2013), pp. 125–144.

[31] Thomas S Kuhn. “Reflections on my critics”. In: (1970).

[32] Imre Lakatos. “History of science and its rational reconstructions”. In: *PSA: Proceedings of the biennial meeting of the philosophy of science association*. Vol. 1970. D. Reidel Publishing. 1970, pp. 91–136.

[33] Imre Lakatos. *The Methodology of Scientific Research Programmes: Ed by John Worrall and Gregory Currie*. Cambridge University Press, 1978.

[34] LD141405 Landau. “On analytic properties of vertex parts in quantum field theory”. In: *Nuclear Physics* 13.1 (1959), pp. 181–192.

[35] Larry Laudan. *Progress and its problems: Towards a theory of scientific growth*. Vol. 282. Univ of California Press, 1978.

[36] Juan Maldacena. “The large-N limit of superconformal field theories and supergravity”. In: *International journal of theoretical physics* 38.4 (1999), pp. 1113–1133.

[37] Stanley Mandelstam. “Dynamics based on rising Regge trajectories”. In: *Physical Review* 166.5 (1968), p. 1539.

[38] Edward B Manoukian. *Quantum Field Theory*. Vol. 2. Springer, 2016.

[39] Claus Montonen and David Olive. “Magnetic monopoles as gauge particles?” In: *Physics Letters B* 72.1 (1977), pp. 117–120.

[40] Y Nambu. *Quark model and factorization of the Veneziano amplitude II Lectures at the Copenhagen Symp. on Symmetries and Quark Models*. 1970.

[41] KS Narain, MH Sarmadi, and E Witten. “A note on toroidal compactification of heterotic string theory”. In: (1986).

[42] Holger Bech Nielsen and Poul Olesen. “Vortex-line models for dual strings”. In: *Nuclear Physics B* 61 (1973), pp. 45–61.

[43] DAVID I Olive. “From dual fermion to superstring”. In: *The birth of string theory* (2012), pp. 346–360.

[44] Jack E Paton and Chan Hong-Mo. “Generalized Veneziano model with isospin”. In: *Nuclear Physics B* 10.3 (1969), pp. 516–520.

[45] Joseph Polchinski. “Dirichlet branes and Ramond-Ramond charges”. In: *Physical Review Letters* 75.26 (1995), p. 4724.

[46] Joseph Polchinski. “Dualities of fields and strings”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 59 (2017), pp. 6–20.

[47] Karl R Popper and Mario Trinchero. *Logica della scoperta scientifica: il carattere autocorrettivo della scienza*. Fabbri, 1995.

[48] Tullio Regge. "Introduction to complex orbital momenta". In: *Il Nuovo Cimento (1955-1965)* 14 (1959), pp. 951–976.

[49] Dean Rickles. "A brief history of string theory". In: *From Dual Models to M-Theory, Berlin and Heidelberg, DE: Springer-Verlag* (2014).

[50] Adrian Norbert Schellekens. *Four dimensional strings*. Tech. rep. CM-P00055715, 1987.

[51] Matthew D Schwartz. *Quantum field theory and the standard model*. Cambridge university press, 2014.

[52] Lee Smolin. *The trouble with physics: the rise of string theory, the fall of a science, and what comes next*. HMH, 2007.

[53] Mark Srednicki. *Quantum field theory*. Cambridge University Press, 2007.

[54] Andrew Strominger. "Superstrings with torsion". In: *Nuclear Physics B* 274.2 (1986), pp. 253–284.

[55] Andrew Strominger and Cumrun Vafa. "Microscopic origin of the Bekenstein-Hawking entropy". In: *Physics Letters B* 379.1-4 (1996), pp. 99–104.

[56] Andrew Strominger, Shing-Tung Yau, and Eric Zaslow. "Mirror symmetry is T-duality". In: *Nuclear Physics B* 479.1-2 (1996), pp. 243–259.

[57] Leonard Susskind. *DUAL-SYMMETRIC THEORY OF HADRONS. I.* Tech. rep. Yeshiva Univ., New York, 1970.

[58] Leonard Susskind. *The cosmic landscape: String theory and the illusion of intelligent design*. Hachette UK, 2008.

[59] Gabriele Veneziano. "Construction of a crossing-simmetric, Regge-behaved amplitude for linearly rising trajectories". In: *Il Nuovo Cimento A (1965-1970)* 57 (1968), pp. 190–197.

[60] MA Virasoro. "Alternative constructions of crossing-symmetric amplitudes with Regge behavior". In: *Physical Review* 177.5 (1969), p. 2309.

[61] Edward Witten et al. "Magic, mystery, and matrix". In: *Notices of the AMS* 45.9 (1998), pp. 1124–1129.

[62] Edward Witten. "String theory dynamics in various dimensions". In: *Nuclear Physics B* 443.1-2 (1995), pp. 85–126.

[63] Peter Woit. *Not even wrong: The failure of string theory and the continuing challenge to unify the laws of physics*. Random House, 2011.

[64] Shing-Tung Yau and Steve Nadis. *The shape of inner space: String theory and the geometry of the universe's hidden dimensions*. Basic Books, 2010.