The convergence argument for theory assessment

Elena Castellani

Department of Humanities and Philosophy,
University of Florence

QSPACE '19, Bratislava, 14.2.19

Theory assessment in fundamental physics

- Fundamental physics and empirical confirmation: most of the theoretical developments concern regimes (energy scales or length scales) far away current possibility of empirical access.
- Trust in fundamental physics: many of these developments - such as, first of all, those included in the string theoretical framework - have attained a high degree of trust among part of the scientific community.
- Debate: is this trust is justified?
- Criticism from some people: the extent to which empirically still unconfirmed theories are trusted today indicates a substantial *change in scientific methodology*.

Scientific methodology: issues

Philosophy of science's traditional methodological issues:

- the modalities followed in building scientific theories → questions about heuristic ('discovery')
- the modalities followed in assessing scientific theories (on the grounds of **empirical** and **extra-empirical** support) → questions about 'justification'
- the inter-relations between building and assessing; the influence of 'external' aspects (sociological, economical, ..); ...

Scientific methodology: issues

Philosophy of science's traditional methodological issues:

- the modalities followed in building scientific theories $\rightarrow$ questions about heuristic ('discovery')
- the modalities followed in assessing scientific theories (on the grounds of **empirical** and **extra-empirical** support) $\rightarrow$ questions about 'justification'
- the inter-relations between building and assessing; the influence of 'external' aspects (sociological, economical, ..); ...

The specific question at stake, in the debate on the status of fundamental physics:

- whether the methodology followed in fundamental physics in building/assessing theories (such as string theory), is 'scientific'

  (in the same sense, say, that the methodology followed in building/assessing the Standard Model is scientific)

Some points

- The problematic aspects regard the **assessment methodology** (not the criteria followed in building the theory)

Some points

- The problematic aspects regard the **assessment methodology** (not the criteria followed in building the theory)

  NOTE: this relies on a controversial/artificial separation between building and assessing (the *discovery/justification contexts* of logical empiricism)

Some points

- The problematic aspects regard the **assessment methodology** (not the criteria followed in building the theory)

  NOTE: this relies on a controversial/artificial separation between building and assessing (the *discovery/justification contexts* of logical empiricism)

- In general $\rightarrow$ **two sorts of theory assessment**:

Some points

- The problematic aspects regard the **assessment methodology** (not the criteria followed in building the theory)

  NOTE: this relies on a controversial/artificial separation between building and assessing (the *discovery/justification contexts* of logical empiricism)

- In general → **two sorts of theory assessment**:

  a) assessment based on empirical support

  b) assessment based on 'extra-empirical' support

Some points

- The problematic aspects regard the **assessment methodology** (not the criteria followed in building the theory)

  NOTE: this relies on a controversial/artificial separation between building and assessing (the *discovery/justification contexts* of logical empiricism)

- In general → **two sorts of theory assessment**:

  a) assessment based on empirical support → in the case considered here, problematic for scientific reasons, both technical and theoretical (the technology is not advanced enough; the theory is not finished/not developed enough)

  b) assessment based on 'extra-empirical' support

Some points

- The problematic aspects regard the **assessment methodology** (not the criteria followed in building the theory)

  NOTE: this relies on a controversial/artificial separation between building and assessing (the *discovery/justification contexts* of logical empiricism)

- In general → **two sorts of theory assessment**:

  a) assessment based on empirical support → in the case considered here, problematic for scientific reasons, both technical and theoretical (the technology is not advanced enough; the theory is not finished/not developed enough)

  b) assessment based on 'extra-empirical' support → philosophical issue: what is the admissible tipology, role and extent of extra-empirical support in science

The real issue

- How to obtain a reasonable balance between *empirical* and *non-empirical* criteria for theory assessment, given the particular physical context considered.

- NOT: whether non-empirical theory assessment could or should substitute the traditional way of confirming physical theories, i.e., by confrontation with empirical data.

$\rightarrow$ On the fact that a physical theory must receive, sooner or later, an empirical confirmation there can be no real disagreement.

Extra-empirical support: criteria/strategies (1)

- General **meta-level** criteria:
  for ex., the "non-empirical-confirmation" arguments individuated by Dawid (2013)
  - No Alternatives Argument: despite extensive search, an alternative theory has not been found.
  - Meta-inductive Argument: other comparable theories in the research field were empirically successful later on.
  - Unexpected Explanation Argument: the theory explains something without having been developed to that end.

$\rightarrow$ "Each of the three arguments remains weak and questionable in isolation. But the arguments gain strength and significance in conjunction." (Dawid, 2017)

Extra-empirical support: criteria/strategies (2)

- More specific ('internal') criteria:
  - Traditional 'extra-empirical virtues': consistency, unifying power/generality, simplicity, fertility, explanatory power, elegance, beauty, ... (e.g., Kuhn's 'values' for choosing a theory)
  - Other criteria/arguments, such as the **convergence argument**: that is, the argument for extra-empirical support which is based on the fact that some of the crucial new results are obtained in *alternative*, *independent* ways, and even from *different starting points*.

    $\rightarrow$ This sort of argument has been particularly effective in providing a motivation for accepting apparently very unphysical features emerging from theoretical developments.

    A paradigmatic example is given by the story of the discovery of extra dimensions (22 extra space dimensions) in the framework of *early string theory*.

Early string theory (EST)

The origin and first developments of string theory, from **Veneziano**'s 1968 formulation of his famous *scattering amplitude*, to the so-called *first string revolution* in **1984**.

Early string theory (EST)

The origin and first developments of string theory, from **Veneziano**'s 1968 formulation of his famous *scattering amplitude*, to the so-called *first string revolution* in **1984**.

Why look at EST's history:

Early string theory (EST)

The origin and first developments of string theory, from **Veneziano**'s 1968 formulation of his famous *scattering amplitude*, to the so-called *first string revolution* in **1984**.

Why look at EST's history:

- it provides light on the origin of many *ideas* (string, duality, supersymmetry, extra dimensions, ...) and *mathematical techniques* that are basic ingredients in today's fundamental physics;

Early string theory (EST)

The origin and first developments of string theory, from **Veneziano**'s 1968 formulation of his famous *scattering amplitude*, to the so-called *first string revolution* in **1984**.

Why look at EST's history:

- it provides light on the origin of many *ideas* (string, duality, supersymmetry, extra dimensions, ...) and *mathematical techniques* that are basic ingredients in today's fundamental physics;
- it provides light on the *rationale of a scientific progress* characterized by the close interplay of mathematically driven creativity and physical constraints (both theoretical and experimental).

Early string theory (EST)

The origin and first developments of string theory, from **Veneziano**'s 1968 formulation of his famous *scattering amplitude*, to the so-called *first string revolution* in **1984**.

Why look at EST's history:

- it provides light on the origin of many *ideas* (string, duality, supersymmetry, extra dimensions, ...) and *mathematical techniques* that are basic ingredients in today's fundamental physics;
- it provides light on the *rationale of a scientific progress* characterized by the close interplay of mathematically driven creativity and physical constraints (both theoretical and experimental).

⇒ novel 'data' (case studies) for discussing the building/assessing issue: in particular, the role and characteristics of the *evidential support* (empirical as well as **extra-empirical**) in the construction process of a scientific theory.

Early string theory (EST)

The origin and first developments of string theory, from **Veneziano**'s 1968 formulation of his famous *scattering amplitude*, to the so-called *first string revolution* in **1984**.

Why look at EST's history:

- it provides light on the origin of many *ideas* (string, duality, supersymmetry, extra dimensions, ...) and *mathematical techniques* that are basic ingredients in today's fundamental physics;
- it provides light on the *rationale of a scientific progress* characterized by the close interplay of mathematically driven creativity and physical constraints (both theoretical and experimental).

⇒ novel 'data' (case studies) for discussing the building/assessing issue: in particular, the role and characteristics of the *evidential support* (empirical as well as **extra-empirical**) in the construction process of a scientific theory.

Reference: A. Cappelli, E. Castellani, F. Colomo, and P. Di Vecchia (eds.) (2012), *The Birth of String Theory*, Cambridge University Press.

Early string theory

• First phase (1968-1973/4): the so-called **dual theory of strong interactions** ('falsified' as such).

• Second phase (1974-1984): those very features that were drawbacks of the theory for describing hadronic physics (*spin-one* and *spin-two massless particles* and *extra dimensions*) were taken to reveal the true nature of string theory, leading to re-interpret it as a **unified quantum theory of all fundamental interactions**.

However, it remained a side issue for a decade until, towards the end of 1984 (thanks to crucial results obtained in the context of supersymmetric string theory), the interest in string theory exploded again.

Early string theory

• First phase (1968-1973/4): the so-called **dual theory of strong interactions** ('falsified' as such).

• Second phase (1974-1984): those very features that were drawbacks of the theory for describing hadronic physics (*spin-one* and *spin-two massless particles* and *extra dimensions*) were taken to reveal the true nature of string theory, leading to re-interpret it as a **unified quantum theory of all fundamental interactions**.

However, it remained a side issue for a decade until, towards the end of 1984 (thanks to crucial results obtained in the context of supersymmetric string theory), the interest in string theory exploded again.

NOTE: the theory had such a *compelling mathematical structure*, obtained in agreement with *consistency conditions* and deep *physical principles*, that the intuition was that it had to be somehow related to the physical world $\rightarrow$ this surely was a strong *motivation* for pursuing it further.

EST: First phase (1968-1973/4)

Aim: To find a viable theory of hadrons.

Context: The so-called *S-matrix approach* to describing the physics of strong interactions in the 1960s.

Motivation: The difficulties arising in a field theoretic description of strong interactions.

EST: First phase (1968-1973/4)

Aim: To find a viable theory of hadrons.

Context: The so-called *S-matrix approach* to describing the physics of strong interactions in the 1960s.

Motivation: The difficulties arising in a field theoretic description of strong interactions.

Program (inspired by earlier work of **Heisenberg**): To determine the relevant observable physical quantities, namely, the *scattering amplitudes* (the elements of the S-matrix) on the basis of general principles such as *unitarity*, *analyticity* and *crossing symmetry*, and a minimal number of additional assumptions.

EST: First phase (1968-1973/4)

Aim: To find a viable theory of hadrons.

Context: The so-called *S-matrix approach* to describing the physics of strong interactions in the 1960s.

Motivation: The difficulties arising in a field theoretic description of strong interactions.

Program (inspired by earlier work of **Heisenberg**): To determine the relevant observable physical quantities, namely, the *scattering amplitudes* (the elements of the *S*-matrix) on the basis of general principles such as *unitarity*, *analyticity* and *crossing symmetry*, and a minimal number of additional assumptions.

In particular: the assumption of the DHS (Dolen-Horn-Schmid) duality, also known as the *dual bootstrap*.

DHS duality

• The assumption ('duality principle') by Dolen, Horn and Schmid (**1967**), suggested by the **experimental data**, that the contributions from *resonance intermediate states* and from *particle exchange* each formed a complete representation of the scattering process $\rightarrow$ so that they should not be added to one another in order to obtain the total amplitude.

DHS duality

• The assumption ('duality principle') by Dolen, Horn and Schmid (**1967**), suggested by the **experimental data**, that the contributions from *resonance intermediate states* and from *particle exchange* each formed a complete representation of the scattering process $\rightarrow$ so that they should not be added to one another in order to obtain the total amplitude.

• In terms of the *Mandelstam's variables* and in the framework of the so-called *Regge theory*:

the duality principle (as initially stated) established *direct relations between a low-energy and a high-energy description* of the hadronic scattering amplitude $A(s, t)$, namely:
the low-energy description in terms of direct-channel resonance poles (the contributions from resonance intermediate states), and
the high-energy description in terms of the exchange of so-called Regge poles in the crossed-channel (the contributions from particle exchange)

could each be obtained from the other by *analytic continuation* (and, thus, each formed a complete representation of the scattering process).

The ‘dual boostrap’

The duality principle was seen to represent an effective implementation of two (connected) ideas defended, in particular, by *Geoffrey Chew* and his school:

• the idea of nuclear democracy — no hadron is more fundamental than the others),

• the bootstrap idea — the idea of a self-consistent hadronic structure in which the entire ensemble of hadrons provided the forces (by hadron exchange) making their own existence (as intermediate states) possible.

EST: first phase

The 'official' starting point:
the 1968 discovery by **Gabriele Veneziano** of his "dual amplitude" for the scattering of four mesons.

It provided a first, brilliant solution to the problem of finding a scattering amplitude that, in the framework of the *S-matrix approach*, obeyed also the *duality principle* (DHS duality) ($\rightarrow$ hence the name *dual amplitude*).

EST: first phase

The ‘official’ starting point:
the 1968 discovery by **Gabriele Veneziano** of his “dual amplitude” for the scattering of four mesons.

It provided a first, brilliant solution to the problem of finding a scattering amplitude that, in the framework of the *S-matrix approach*, obeyed also the *duality principle* (DHS duality) ($\rightarrow$ hence the name *dual amplitude*).

The result by Veneziano immediately gave rise to the very intense theoretical activity that is known, in general, as the *dual theory of strong interactions*:

from the first two models proposed – the **Dual Resonance Model** and the **Shapiro-Virasoro Model**, respectively – to all the subsequent endeavours to extend, complete and refine the theory, including its string interpretation and the addition of *fermions*.

Conjectures/Discoveries

Some decisive *conjectures* or ‘discoveries’ characterize EST’s first phase.

In particular:

- The *string conjecture* in 1969: in independent attempts to gain a deeper understanding of the physics described by dual amplitudes, **Nambu**, **Nielsen** and **Susskind** each arrived at the conjecture that the underlying dynamics of the dual resonance model was that of a *quantum-relativistic oscillating string*.

Conjectures/Discoveries

Some decisive *conjectures* or ‘discoveries’ characterize EST’s first phase.

In particular:

- The *string conjecture* in 1969: in independent attempts to gain a deeper understanding of the physics described by dual amplitudes, **Nambu**, **Nielsen** and **Susskind** each arrived at the conjecture that the underlying dynamics of the dual resonance model was that of a *quantum-relativistic oscillating string*.
- The ‘discovery’ of the so-called *critical dimension*: that is, the discovery that consistency conditions of the *Dual Resonance Model* required the value $d = 26$ for the spacetime dimension (reducing to $d = 10$ dimensions when including fermions $\rightarrow$ in the *Ramond-Schwarz-Neveu model*).

The case study: the discovery of the critical dimension

A significant discovery case, illustrative of both the *rationale* leading to apparently bold guesses and the kind of *evidential support* motivating a theory's progress.

The case study: the discovery of the critical dimension

A significant discovery case, illustrative of both the *rationale* leading to apparently bold guesses and the kind of *evidential support* motivating a theory's progress.

Originally, the critical value $d = 26$ was obtained in two *independent ways* $\rightarrow$ *alternative approaches*:

• The first $\rightarrow$ the one followed by **Claud Lovelace**, in a work published in **1971**, where he addressed a problematic *singularity case* arising in the construction of the nonplanar *one-loop amplitude* in the framework of the *(DRM) unitarization program*.

• The second $\rightarrow$ the same result issued, soon after, through another route: namely, from the examination of the *DRM physical spectrum of states* in the context of the *ghost elimination program*.

Alternative ways to the critical dimension: (1)

1. Lovelace's way (1971)

• **The context**: the "unitarization programme", to go beyond the initial narrow-resonance approximation.

The programme was to generalise the initial scattering amplitudes, considered as the lowest order or tree diagrams of a perturbative expansion, to include loops.

As a first step for restoring unitarity, one-loop diagrams were constructed:

in this construction process a *singularity* problem arising was solved by interpreting the singularity as being due to the propagation of a *new intermediate particle* (the particle that was later understood as the *graviton*).

• The decisive step was to consider the possibility that the spacetime dimension d might be different from 4 and treat it as a free parameter.

$\rightarrow$ the result was that the singularity became a pole only for $d = 26$.

• The decisive step was to consider the possibility that the spacetime dimension d might be different from 4 and treat it as a free parameter.

$\rightarrow$ the result was that the singularity became a pole only for $d = 26$.

• A spacetime of 26 dimensions was not easy to accept, especially in the context of the phenomenology of strong interactions where it was proposed.

However, almost at same time, Lovelace's "wild conjecture" (his own words) was vindicated through another completely independent route:

the very same number of spacetime dimensions made its appearance in the context of the ghost elimination programme.

Alternative ways to the critical dimension: (2)

2. The "no ghost" way (1972)

- The context: the **ghost elimination programme**.

In the framework of the generalisation of Veneziano's result to the scattering of an arbitrary number *N* of scalar particles

– in studying the properties of the resulting multi-particle Veneziano model, known as the Dual Resonance Model –

a serious problem was represented by the presence of negative-norm states ("ghosts")

$\rightarrow$ these states, leading to unphysical negative probabilities, had to be eliminated from the theory.

• By using the so-called *DDF states* (the infinite set of positive-norm states found by **Del Giudice**, **Di Vecchia** and **Fubini** in 1971),

the result was obtained in 1972 that these DDF states could indeed span the whole space of physical states **if the spacetime dimension d was equal to 26**

$\rightarrow$ the very same value as the one conjectured by Lovelace.

Soon after, the proof of the so-called No-Ghost Theorem, establishing that the Dual Resonance Model has no ghosts if $d \leq 26$, was achieved by **Brower**, and by **Goddard** and **Thorn**.

In the critical dimension, *consistency* was thus satisfied, but at the high price of extra 22 space dimensions.

$\rightarrow$ a rather unrealistic feature!

Nonetheless, the extra dimensions became gradually accepted, owing to the fact that the critical dimension result received *further support* from successive *theoretical developments*.

In the critical dimension, *consistency* was thus satisfied, but at the high price of extra 22 space dimensions.

$\rightarrow$ a rather unrealistic feature!

Nonetheless, the extra dimensions became gradually accepted, owing to the fact that the critical dimension result received *further support* from successive *theoretical developments*.

In particular, a *third decisive evidence* was provided:

$\rightarrow$ through the **1973** work of **Goddard**, **Goldstone**, **Rebbi** and **Thorn** on the quantization of the string action.

Alternative ways to the critical dimension: (3)

3. The GGRT (1973) way

• The correct string action was proposed first by Nambu in 1970 and then by Goto, but the string interpretation became effectively applied only after the quantization of the string action obtained by Goddard, Goldstone, Rebbi and Thorn (GGRT) in 1973.

• With this result it became possible to derive, in a clear and unified way, all that had previously been discovered regarding the DRM spectrum by proceeding along various paths and according to a bottom-up approach.

In particular: the $d = 26$ condition was obtained from the canonical quantization of the string in the light-cone gauge: it resulted from the requirement of Lorentz invariance in the quantum theory,

Alternative ways to the critical dimension: (4)

4. The ‘conformal anomaly’ way

Further and more decisive support to the extra dimension conjecture came from successive developments of string theory, especially after it was re-interpreted as a *unified quantum theory of all fundamental interactions* including gravity.

In fact, with hindsight, the critical dimension is understood as a consequence of the *conformal symmetry* of string theory:

as shown by **Polyakov** in **1981**, the *conformal symmetry* of the classical string Lagrangian is ‘anomalous’, i.e. not conserved in the quantum theory, unless the value of the spacetime dimension is $d = 26$.

The critical dimension case study: summing up

• In the first phase, the *critical dimension condition* could be found only on the basis of 'side effects': for ex., as a condition required by *unitarity* of the theory, or by *Lorentz invariance* in the quantization of the string action.

• After, what had appeared to be a surprising convergence of different calculational procedures to one and the same result could be seen as a natural consequence of the theory in its fully-fledged form.

The critical dimension case study: summing up

• In the first phase, the *critical dimension condition* could be found only on the basis of ‘side effects’: for ex., as a condition required by *unitarity* of the theory, or by *Lorentz invariance* in the quantization of the string action.

• After, what had appeared to be a surprising convergence of different calculational procedures to one and the same result could be seen as a natural consequence of the theory in its fully-fledged form.

• Role of the convergence argument $\rightarrow$ illustrated in the case study:

the convergence to the same new result - the extra dimensions - obtained in *alternative ways* and from *different starting points* provided important *support*, that motivated persevering with the theory, notwithstanding the presence of unrealistic features such as extra space dimensions.

Historical note: Whewell's "Consilience of Inductions"

W. Whewell, *The Philosophy of the Inductive Sciences, Founded Upon Their History* (1840):

"The **Consilience of Inductions** from different and separate classes of facts"

$\rightarrow$ as both a test for theory's truth (besides *prediction* and *coherence*) and a "principal feature in the progress of science"

> "The **evidence** in favour of our induction is of a much higher and more forcible character when it enables us to explain and determine cases of a kind different from those which were contemplated in the formation of our hypothesis.
>
> The instances in which this have occurred, indeed, impress us with a **conviction that the truth of our hypothesis is certain**."

• Whewell's exemplary case of consilience → **Newton's law of universal gravitation**

Phenomena constituting different event kinds

"planetary motion", "satellite motion", "falling bodies"

were found to be member of a **unified, more general kind of phenomena**: "the phenomena caused to occur by an inverse-square attractive force of gravity"

• Consilience as feature of **progress** of science, for Whewell:

By seeing that an *inverse square attractive force* provided a cause for different event kinds, Newton was able to perform a **more general induction** → his universal gravitation law.

Thank You!