Highlights in High-Energy Physics
https://www.sciltp.com/journals/hihep

Opinion

The Rise of Gauge Theories: From Many Models to One Theory

Jean Iliopoulos

Laboratoire de Physique de l'Ecole Normale Supérieure, ENS-PSL, CNRS, Sorbonne Université, Université Paris Cité, 75005 Paris, France; ilio@phys.ens.fr

How To Cite: Iliopoulos, J. The Rise of Gauge Theories: From Many Models to One Theory. *Highlights in High-Energy Physics* **2025**, *1*(1), 5. https://doi.org/10.53941/hihep.2025.100005.

Received: 5 April 2025
Revised: 16 June 2025
Accepted: 20 June 2025
Published: 30 June 2025

Abstract: By a worth noting coincidence, the year 2024 marks two anniversaries: The 70th anniversary of CERN and the 50th anniversary of the J/Ψ discovery. In this talk I will present very briefly the significance of these anniversaries [1] . CERN's founding fathers were among Europe's leading scientists. CERN became the most successful International Institution. Scientists built the Scientific European Union long before politicians thought of the Economic European Union. If the Scientific Institution appears to be more solid than the Economic one, it is because its foundations are made with ideas. Regarding the second anniversary, I want to argue that J/Ψ was not just a new resonance; we knew already a large number of them. It was not even only the first indirect evidence for the existence of a new quark flavor. It was all that, but it was also much more. It was the final proof which convinced the large majority of our community that we were witnessing a radical change of paradigm in our understanding of microscopic physics. From phenomenological models and specific theories, each one applied to a restricted set of experimental data, we had to think in terms of a fundamental theory of universal validity. From many models to the STANDARD THEORY of particle physics. For most of us this transition was a revelation, for some others it was a painful experience. It is appropriate to combine it with CERN's anniversary because the experimental verification of the Standard Model started at CERN with the neutral currents and was completed also at CERN with the BEH scalar boson. It is this story that I will attempt to narrate in this note, although I do not consider myself to be an unbiased observer.

Keywords: gauge theories; charm; standard model

1. The Origins

If I had to assign a year to the birth of experimental particle physics I would choose 1950, the discovery of the neutral pion at the Berkeley electron synchrotron. It is the first "elementary" particle discovered with an accelerator (the charged pion was discovered in 1947 in cosmic rays). The existence of π^0 was predicted in 1938 by N. Kemmer who wrote the first isospin invariant theory for the nuclear forces (Kemmer has not received the appropriate recognition for his groundbreaking work. His equations for the pion-nucleon interaction can be found in every textbook, but his name is rarely mentioned). So, π^0 is the first particle whose existence was predicted by an argument based on an internal symmetry and also the first to be discovered in an accelerator. Since that time accelerators became the main engines of discovery. With their use the field expanded very rapidly and today the "Table of Elementary Particles" has several hundred entries, although we know that very few among them are "elementary".

During the second half of the 20th century experimental particle physics has followed a monotonically rising trajectory, in contrast, that of its theoretical counterpart was more circuitous. Modern theoretical high energy physics has a precise date of birth: 2 June 1947, the date of the Shelter Island Conference. The most important contributions which were presented in that conference were not theoretical breakthroughs but two experimental results: the non-zero values of the Lamb shift and of the electron anomalous magnetic moment. Both these results were important because, for the first time, they showed, beyond any possible doubt, that the predictions based on the Dirac equation for the electron are only approximately correct. This in turn motivated a serious study of quantum field theory. In the following months R.P. Feynman and J.S. Schwinger, independently, using apparently different formulations, set up the program for the renormalised perturbation expansion of Quantum Electrodynamics and Schwinger gave the first calculation of $g-2$. As it turned out, similar results were obtained independently in Japan by Sin-Itiro Tomonaga, who obtained also the first complete calculation of the Lamb shift. The equivalence of all these approaches was formally shown by F. Dyson in 1948. Rare are the examples in physics in which so much progress was accomplished in such a short time. The agreement between theory and experiment in the phenomena of quantum electrodynamics is impressive.

The natural next step was to apply this approach to the other two interactions, to wit the strong and the weak ones. The former were represented by the isospin invariant pion-nucleon interaction and it was shown that the renormalisation program applied to it. However, the results were useless because the effective pion-nucleon coupling constant turned out to be very large, on the order of 10, making the power series expansion meaningless. The dynamics appears to be dominated by phenomena such as resonance production, which cannot be described by ordinary perturbation theory. There remained the weak interactions. Naturally, they are much weaker, but now a new problem appeared: the renormalisation program is a quite complex process and applies only to a small number of quantum field theories. The Fermi theory, which describes very well the low energy weak interactions, is not one of them.

This double failure soon tarnished the glory of quantum field theory. The disappointment was such that the subject was not even taught in many universities. By the late fifties the theoretical high energy physics landscape was fragmented in many disconnected domains, having no common trends and often ignoring each other. For strong interaction processes the main approach was based on the assumed analytic properties of the S-matrix elements, but we had also several simple models, none with any solid theoretical basis, each one applied to a particular corner of phase space. For weak interactions the Fermi theory proved to be a very good phenomenological model, but it had no logical justification and no obvious connection with anything else. The most important progress in our understanding of Nature's fundamental laws came from the application of symmetry principles and not from dynamical calculations. Quantum field theory was noticeable essentially by its absence. A totally marginal subject confined to very few precision calculations in quantum electrodynamics. Many physicists had only vague and often erroneous ideas about it and, to a certain extent, this misunderstanding has survived even today. The "Dark Ages" would last two decades.

2. The Secret Road to the New Theory

It is usually said that progress in science occurs when an unexpected experimental result contradicts the current theoretical beliefs. This forces scientists to change their ideas and leads to a new theory. This has often been the case in the past, but the revolution we are going to describe here had a theoretical, rather an aestethic motivation. It was a triumph of abstract theoretical thought which brought geometry into physics. This "secret road" has been long and circuitous with no discernible well defined direction. Many a time it gave the impression of leading to a dead end. In fact it was a long series of isolated and mostly confidential contributions and many important ideas had to be rediscovered again and again. Several milestones did not seem to point to a single path and most of them went unnoticed when they were first proposed. The pioneers were often unaware of each other's work and it is only now that we can see a coherent picture. A strict chronological order is not possible and I choose instead to group together a few important contributions which became the main ingredients of the Standard Model. Notice that most of them had motivations unrelated to their final application. The construction of the Standard Model used many concepts and techniques of quantum field theory. A short list includes:

- Gauge invariance.
- Renormalisation, in particular for gauge invariant field theories.
- Renormalisation and symmetry, including the phenomenon of anomalies in the conservation of symmetry currents.
- The renormalisation group and the property of asymptotic freedom.
- Higher internal symmetries and the quark model.

The phenomenon of spontaneous symmetry breaking, in particular in the presence of long range gauge forces. I had no time to present the full story of all these discoveries in Rome and a short account can be found in reference [1]. I gave only a discussion of the concept of gauge invariance which goes back to classical electrodynamics. I do not know who was the first to remark that the dynamical system described by the components of the electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$ – that is, six degrees of freedom in our counting—was in fact redundant because some of the equations do not involve any time derivatives and should be considered as constraints. It seems that the first person who attempted to reduce the redundancy was C.F. Gauss who, in some manuscript notes in 1835, introduced the concept of the "vector potential" $\boldsymbol{A}$. It was further developed by several authors and was fully written by G. Kirchoff in 1857, following earlier work, in particular by F. Neumann. The components of the electric and magnetic fields could be expressed in terms of the vector and scalar potentials, thus reducing the number of degrees of freedom from six to four. It was soon noticed that it still carried redundant variables and several "gauge conditions" were used. The condition, which in modern notation is written as $\partial_\mu A^\mu = 0$, was proposed by the Danish mathematical physicist L.V. Lorenz in 1867. It seems that Maxwell favored the condition $\boldsymbol{\nabla} \cdot \boldsymbol{A}(x) = 0$ which today we call *"the Coulomb gauge"*. Using Maxwell's equations we can immediately see the redundancy of the system $(\boldsymbol{\Phi}, \boldsymbol{A})$ because the equation for $\boldsymbol{\Phi}$ does not involve any time derivative. Lorenz arrived to this conclusion because he had a formulation of classical electrodynamics equivalent to Maxwell's.

In this context, an interesting story is the following: Around the years 1840 F.E. Neumann and, independently, W.E. Weber, studied the interaction between two closed electric circuits carrying currents I and I', respectively. Their methods and physical assumptions were different and so were the expressions they obtained for the magnetic interaction energy between the two circuits. However, they both seemed to fit Ampère's measurements. In the years after 1870, H.L.F. von Helmholtz criticised and compared the two expressions. In particular, he noticed that Neumann's and Weber's formulae for the elementary magnetic interaction energy, differ by a quantity which can be expressed as a multiple of a perfect differential, so, they give the same result when integrated over the closed circuits. In our present terminology he showed that the two expressions are *gauge equivalent.* He even went a step further: he generalised the two results by exhibiting a one-parameter family of expressions for the magnetic energy which interpolate between those of Neumann and Weber. It was the first example of a *family of gauges.* By the end of the century H.A. Lorentz published a book and some encyclopedia articles with the full classical electromagnetic theory. The invariance under gauge transformations of the vector and scalar potentials is an integral part of it.

At the beginning of the 20th century the development of the general theory of relativity offered a new paradigm for a gauge theory and triggered several attempts to unify electromagnetism and gravitation. In reference [1] I mention the work of G. Nordström, as well as that of T. Kaluza and O. B. Klein, which is used today in supergravity and superstring theories. An independent approach is due to H.K.H. Weyl in 1919. It is mostly known for his unsuccessful attempt to enlarge the gauge symmetry of general relativity to include invariance under local scale transformations of the metric: $g \to e^{2\lambda(x)} g$, with $\lambda(x)$ an arbitrary function of the space-time point x. He called the resulting invariance *eichinvarianz* in German, which was translated in English as *gauge invariance*. Weyl was the first to understand that the classical electromagnetic theory was mathematically incomplete. People had discovered by trial and error its property of gauge invariance, but the underlying global symmetry was missing. Weyl proposed to identify it with scale transformations. It was a wrong physical answer to a correct mathematical question. The laws of physics are in no way invariant under global scale transformations. As it turned out, the correct answer was found in the framework of quantum mechanics.

The first person who understood that the invariance of quantum mechanics under phase transformations of the wave function could be the missing global symmetry of electromagnetism, was V.A. Fock. In a paper published in 1926, just after Schrödinger wrote his equation, he showed that promoting the invariance from global to local, is equivalent to writing the equation for an electron in an external electromagnetic field. The extension of this simple idea to non-abelian internal symmetries – in particular the $SU(2)$ isospin transformations developed by Heisenberg and Kemmer between 1932 and 1938 – took many years and followed a very complicated and counter-intuitive way [1]. It was finally achieved by C.N. Yang and R.L. Mills in 1954. Since that time gauge theories became part of high energy physics.

3. The Standard Model–A Theoretical Speculation

The entire theoretical scheme, which became the Standard Theory of particle physics, was fully written in 1973. It is a gauge theory based on the group $SU(3) \times SU(2) \times U(1)$, spontaneously broken to $SU(3) \times U(1)_{em}$. It describes all phenomena we observe in our experiments. Yet, it was not generally accepted. For most physicists it was a wild theoretical speculation with no connection to the real world. The main reason for this negative attitude was the mistrust towards quantum field theory, but it is also true that the model seemed to make many strange predictions with little, if any, experimental support. Let me mention some of them.

- The model predicted the existence of 12 vector bosons. But only one, the photon, was known! Three, $(W^\pm,\ Z^0)$ were predicted to be very heavy (For the 1973 physicists a mass of 100 GeV was essentially infinite!) and the other eight, the gluons, were declared *unobservable* by a strange property of *confinement.*
- The model predicted the existence of a scalar boson, the BEH, with unknown mass. For many physicists it was

a heresy coming after the assumed triumph of the $V - A$ theory of weak interactions.

- Neutral currents were predicted but the obvious ones, the $K^0 \to \mu^+ + \mu^-$ decay, were excluded. Gargamelle had established the existence of strangeness conserving neutral currents but not everybody was convinced (I won some bottles of very fine wine by betting for neutral currents, in particular against Jack Steinberger). Furthermore, the possible existence of weak neutral currents had been envisaged before the formulation of gauge theories, so their existence was not considered as a decisive proof of the new theory.
- Probably the most "extravagant" prediction was that of the charmed quark, implying the existence of an entire new hadronic world of charmed particles. For most people the arguments were not considered serious. I still remember the objections: some obscure higher order effects – triangle diagrams for the anomalies, or square diagrams for the absence of flavor violating neutral currents – would dictate the structure of the world? Totally absurd! In retrospect, I think that the large majority of physicists rejected this particular prediction because it went against the prevailing philosophy of compartmentalisation of high energy physics. Theoretical arguments motivated by properties of weak interactions were not admissible to make predictions on hadronic physics, a domain reserved exclusively to strong interactions (*Sutor, ne supra crepidam or, Let the cobbler stick to his last.)*
- The QCD prediction for the ratio $R(Q^2) = \frac{\sigma(e^+ + e^- \to \mathbf{hadrons})}{\sigma(e^+ + e^- \to \mu^+ + \mu^-)}$ seemed to be in violent contradiction with experiment.

4. The London Conference

This brings me to the 17th International Conference on High Energy Physics held in London in July 1974. It can be viewed as the last Conference of the Dark Ages and the first Conference of the New Era. Several old and new results were announced in this Conference. I will mention only those which are important for our story.

- D.C. Cundy gave the plenary report on neutrino physics. Obviously, the Gargamelle results were the central point. I remember him saying: *"Those who have bet on neutral currents, now is the time to pay and collect."*
- B. Richter gave the plenary report on $e^+e^- \to$ hadrons. In Figure 1 I show the graph he presented.

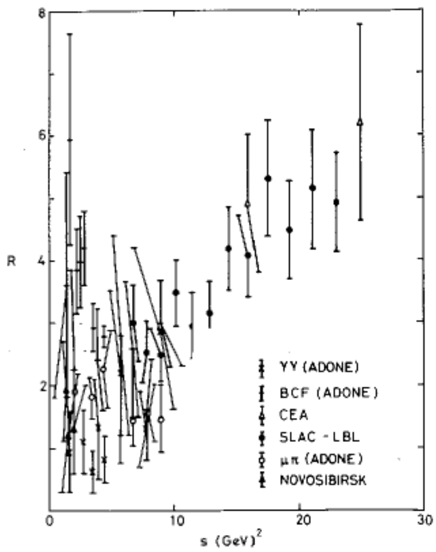


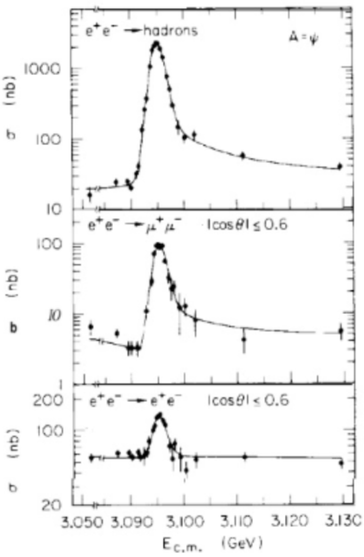
Figure 1. A compilation of all early measurements of the ratio R, as presented in the 1974 London International Conference on High Energy Physics by Burton Richter.

I remind you that the QCD prediction was that R should approach the value $R = 2$ (the sum of the electric charges squared of u, d and s quarks) from above. No such behavior is visible in Figure 1.

- In my plenary report on gauge theories at the same Conference I said: *"... the hadron production cross section, which absolutely refuses to fall, creates a serious problem. The best explanation may be that we are observing the opening of the charmed thresholds, in which case everything fits together very nicely."* The addition of a charmed quark would add an extra 4/3 to R (By a numerical accident, the data of Figure 1 contain also the production of the τ lepton which was not known at the time). A. Salam, who was the Chairman of the Conference, had ready a bottle of a fine Bordeaux red wine to reward speakers who finish on time. Naturally I was late and I said something like: *"I know I am about to loose my bottle, but I am ready to bet now a* *whole case that, if the weak interaction sessions of this Conference were dominated by the discovery of the neutral currents, the entire next Conference will be dominated by the discovery of the charmed particles."* This convinced Salam to give me the bottle which I opened immediately, poured myself a generous libation, offered the rest to those sitting in the first row, and drank "to charm!".

5. The Charming Theory of the New Particles

In November 1974 both Brookhaven and Stanford published their results. SPEAR decided to sweep the region above 3 GeV in fine steps of 1 MeV. To their great surprise they obtained a totally different picture, Figure 2.



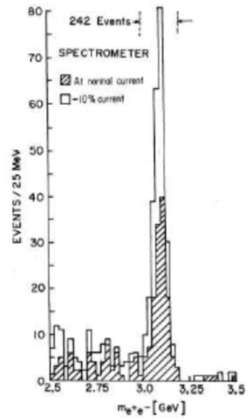


Figure 2. The discovery of the J/Ψ meson in Nov. 1974 independently by SPEAR (**left**) and AGS (**right**). Both exhibit peaks in the oppositely charged dielectron mass spectrum consistent with the J/Ψ mass at 3.1 GeV. This result was also confirmed by the Frascati group.

I was in Paris at the end of 1974 when I received a telephone call from A. Lagarrigue inviting me to an informal meeting to discuss important results from SPEAR. G. 't Hooft was visiting Ecole Normale and I took him along. In those days news did not travel with the speed of internet. B. Jean-Marie had come from Stanford with the results. They were indeed very impressive. A 3 GeV hadron, decaying into pions, with a width of less than 100 keV? Incredible! Lagarrigue asked me what I thought of it and I confess I was bewildered. It was 't Hooft who first gave me the explanation. It is simple to understand in QCD. Let us consider the series of 1^- mesons: ρ is a bound state of a quark-antiquark pair of the first family. Its mass lays well above the 2π mass and its width is very large $\sim$147 MeV. The ϕ is an $s\bar{s}$ bound state with a mass of 1020 MeV. It lays barely above the threshold of a $K\bar{K}$ pair, yet its branching ratio to $K\bar{K}$ is 83% despite the fact that the phase space is tiny. The pure pionic partial width is only 650 keV. In the old days we had a rule, called the OZI (Okubo-Zweig-Iizuka) rule, one of those empirical rules of the dark ages with no real theoretical justification. It stated that in a quark-antiquark bound state, the decay modes requiring the annihilation of the initial $q\bar{q}$ pair, were highly suppressed (It was called "the re-arrangement model" by H. Rubinstein). Let us come now to J/Ψ and assume it is a $c\bar{c}$ bound state. The 0^- mesons are supposed to be the pseudo-Goldstone bosons of spontaneously broken chiral symmetry. The latter is very good for the first family, questionable for the s quark, and very poor for charm. Therefore we expect the charmed 0^- mesons D to be quite heavy and the mass of J/Ψ to lay below the $D\bar{D}$ threshold. As a result J/Ψ decays mainly into pions. The decay amplitude for a 1^- meson goes through three gluons, so the width is proportional to α_s^3. Between 1 and 3 GeV α_s has dropped by a factor of two, so we expect the J/Ψ width to be 8 times smaller than the ϕ pionic width. It is precisely what is found experimentally. As I said, I first heard this argument from 't Hooft, but later I found it in papers by Appelquist and Politzer as well as De Rujula and Glashow. The first has a submission date of November 19. This makes me believe it was found before the experimental discovery.

Within a year the entire region between 3 and 5 GeV was studied in detail, see Figure 3.

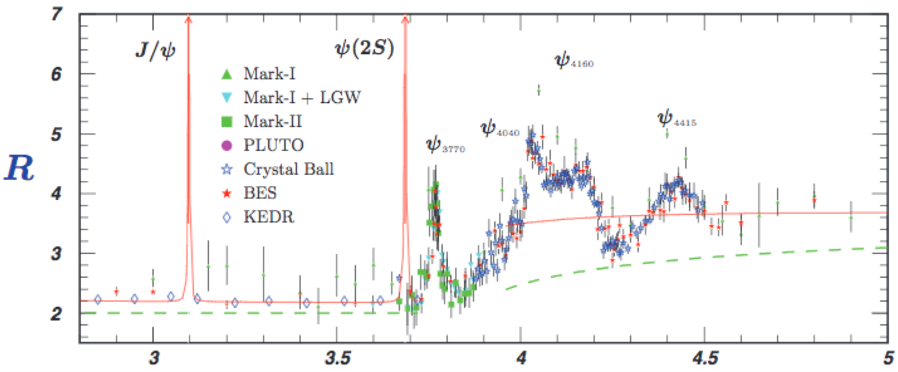


Figure 3. The value of $\boldsymbol{R}$ for energies between 3 and 5 GeV.

As expected, there are broader resonances with masses $\geq$ 4 GeV which are those lying above the $D\bar{D}$ threshold. Lo and behold, the particles with naked charm were found among the decay products of these resonances. It was in 1976. In the meantime a rich charmonium spectroscopy (I believe that the term 'charmonium" was coined by A. De Rujula) was discovered in full agreement with the theoretical predictions. Although nobody paid the bet I offered in the 1974 Conference, the entire 1976 one was indeed dominated by charmed particles and gauge theories. The report in this Conference by A. De Rujula had the title: *"Theoretical basis of the new particles"*. The first sentence is *"I review the four-quark standard gauge field theory of weak, electromagnetic and strong interactions"*. Now we talk about "the four-quark standard gauge field theory". The phase transition from Many Models to One Theory was complete. The order parameter has been the fraction of physicists who changed their views: a small minority before 1974 to the large majority after 1976. The complete verification of the theory took many more years and many great discoveries, but the mood of the community had changed. The following discoveries of the vector bosons, the b and t quarks which complete the third family of the τ lepton, the gluon jets and the BEH scalar as well as the very good general fit using all available data, were no more great surprises, they were expected. THE STANDARD MODEL had become THE STANDARD THEORY.

6. From Dream to Expectation

Feynman has said that progress in physics is to prove yourself wrong as soon as possible. For half a century now we have not been able to prove the Standard Theory is wrong. It has passed successfully all tests and all its predictions have been brilliantly verified. What comes next?

In 2011 the European Physical Society awarded to Glashow, Maiani and myself the High Energy Physics Prize. At this occasion we were invited to speak at the European Conference in Grenoble. The title of my talk was "Following the Path of Charm" and I tried to argue that precision measurements at a certain energy scale allow us to make predictions at some higher scale. Let me start from an expanded version of a plot showing the value of $\boldsymbol{R}$ from low energies up to and above the Z^0 mass, Figure 4.

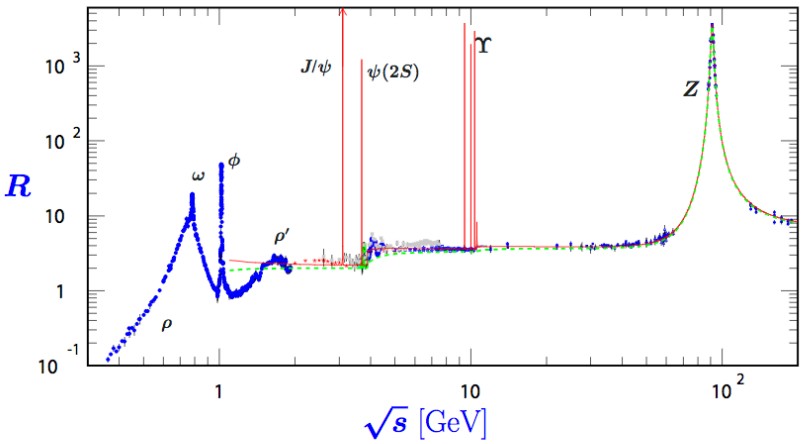


Figure 4. The ratio $\boldsymbol{R}$ from low energies, up to and above the $\boldsymbol{Z}$ mass. The green curve is the parton model prediction and the red one includes QCD corrections. Remarkable agreement.

What is most remarkable is the precision with which theoretical predictions fit the experimental data in most of the energy interval. The only regions in which there is no agreement are the one at low energy below 1 GeV, in which the QCD is in the strong coupling regime, and small, very localised regions signaling the thresholds for the production of new species of hadrons, charm or b. Outside those regions, our theory gives an excellent fit, although it is just the result of one loop perturbation expansion. I concluded that, for reasons that are not fully understood, perturbation theory is reliable outside the regions of strong interactions.

I tried to use this fact in order to make predictions regarding the multi-hundred GeV scale which was expected to be explored by LHC. The argument was based on the low energy data which favored a Higgs particle of relatively low mass, $\leq$ 200 GeV. It went roughly as follows: The existing data are compatible with the Standard Theory only if the Higgs is light. Therefore, if we find a very heavy Higgs we must also find new interactions which invalidate the Standard Theory calculations. If on the other hand we find a light Higgs, we must find new interactions which stabilise its mass at this low value. I thought that both possibilities were good news for LHC. In my talk at EPS in 2011 I said:

> *"I want to exploit this experimental fact* [the validity of perturbation theory] *and argue that the available precision tests of the Standard Model allow us to claim with confidence that new physics is present at the TeV scale and the LHC can, probably, discover it. The argument assumes the validity of perturbation theory and it will fail if the latter fails. But, as we just saw, perturbation theory breaks down only when strong interactions become important. But new strong interactions imply new physics".*

My conclusion was that, for LHC, which was about to start operating, new physics was around the corner!
Today we know that LHC found no corner!
But I secretly believe the argument is correct, only the corner is a bit further down.
Although I will not see it, I am confident some among our young colleagues will find it.

Conflicts of Interest

The author declares no conflict of interest.

Reference

1. Iliopoulos J. From Many Models to ONE THEORY. *arXiv* **2025**, http://arxiv.org/abs/2501.10233.