Perturbation theory

Giovanni Gallavotti
Fisica, Roma 1 and I.N.F.N.

Article Outline

1. Glossary
2. Definition
3. Introduction
4. Poincaré's theorem and quanta
5. Mathematics and Physics. Renormalization
6. Need of convergence proofs
7. Multiscale analysis
8. A paradigmatic example of PT problem
9. Lindstedt series
10. Convergence. Scales. Multiscale analysis
11. Non convergent cases
12. Conclusion and Outlook
13. Bibliography

1 Definition of the subject and its importance

Perturbation Theory: Computation of a quantity depending on a parameter ε starting from the knowledge of is value for $\varepsilon = 0$ by deriving a power series expansion in ε, under the assumption of its existence, and if possible discussing the interpretation of the series. Perturbation theory is very often the only way to get a glimpse of the properties of systems whose equations cannot be "explicitly solved" in computable form. Its importance is witnessed by its applications in Astronomy, where it lead not only to the discovery of new planets (Neptune) but also to the disovery of Chaotic motions, with the completion of the Copernican revolution and the full understanding of the role of Aristotelian Physics formalized into uniform rotations of deferents and epicycles (nowadays Fourier representation of quasi periodic motions). It also played an essential role in the development of Quantum Mechanics and the understanding of the periodic table. The successes of Quantum Field Theory in Electrodynamics first, then in Strong interactions and finally in the unification of the elementary forces (stong, electromagnetic and weak) are also due to perturbation theory, which has also been essential in the theoretical understanding of the critical point universality. The latter two themes concern the new methods that have been developed in the last fifty years, marking a kind of new era for perturbation theory; namely dealing with singular problems, via the techniques called, in Physics, "Renormalization Group" and, in Mathematics, "Multiscale Analysis".

2 Glossary

- **Formal power series**: a power series, giving the value of a function $f(\varepsilon)$ of a parameter ε, that is derived assuming that f is analytic in ε.
- **Renormalization group**: method for multiscale analysis and resummation of formal power series. Usually applied to define a systematic collection of terms to organize a formal power series into a convergent one.
- **Lindstedt Series**: an algorithm to develop formal power series for computing the parametric equations of invariant tori in systems close to integrable.
- **Multiscale Problem**: any problem in which an infinite number of scales play a role.

3 Introduction

Perturbation theory, henceforth PT, arises when the value a function of interest is associated with a problem depending on a parameter, here called ε. The value has to be a simple, or at least explicit and rigorous, computation for $\varepsilon = 0$ while its computation for $\varepsilon \neq 0$, small, is attempted by expressing it as the sum of a power series in ε which will be called here the "solution".

It is important to say since the beginning that a real PT solution of a problem involves two distinct steps: the first is to show that assuming that there is a convergent power series solving the problem then the *coefficients* of the n-th power of ε exist and can be computed via finite computation. The resulting series will be called *formal solution* or *formal series* for the problem. The second step, that will be called *convergence theory*, is to prove that the formal series converges for ε small enough, or at least find a "summation rule" that gives a meaning to the formal series thus providing a real solution to the problem.

None of the two problems is trivial, in the interesting cases, although the second is certainly the key and a difficult one.

Once Newton's law of universal gravitation was established it became necessary to develop methods to find its implications. Laplace's "Mécanique Céleste", [1], provided a detailed and meticulous exposition of a general method that has become a classic, if not the first, example of perturbation theory, quite different from the parallel analysis of Gauss which can be more appropriately considered a "non perturbative" development.

Since Laplace one can say that many applications along his lines followed. In the XIX century wide attention was dedicated to extend Laplace's work to cover various astronomical problems: tables of the coefficients were dressed and published, and algorithms for their construction were devised, and planets were discovered (Neptune, 1846). Well known is the "Lindstedt algorithm" for the computation of the n-th order coefficients of the PT series for the non resonant quasi periodic motions. The algorithm provides a power series representation for the quasi periodic motions with non resonant frequencies which is extremely simple: however it represents the n-th coefficient as a sum of many terms, some of size of the order of a power on $n!$. Which of course is a serious problem for the convergence.

It became a central issue, known as the "small denominators problem" after Poincaré's deep critique of the PT method, generated by his analysis of the three body problem. It led to his "non-integrability theorem" about the generic nonexistence of convergent power series in the perturbation parameter ε whose sum would be a constant of motion for a Hamiltonian H_ε, member of a family of Hamiltonians parameterized by ε and reducing to an integrable system for $\varepsilon = 0$. The theorem suggested (to some) that even the PT series of Lindstedt (to which Poincaré's theorem does not apply) could be meaningless even though formally well defined, [2].

A posteriori, it should be recognized that PT was involved also in the early developments of Statistical Mechanics in the XIX century: the virial theorem application to obtain the Van der Waals equation of state can be considered a first order calculation in PT (although this became clear only a century later with the identification of ε as the inverse of the space dimension).

4 Poincaré's theorem and quanta

With Poincaré begins a new phase: the question of convergence of series in ε becomes a central one in the Mathematics literature. Much less, however, in the Physics literature where the new discoveries in the atomic phenomena attracted the attention. It seems that in the Physics research it was taken for granted that convergence was not an issue: atomic spectra were studied via PT and early authoritative warnings were simply disregarded (for instance, explicit by Einstein, in [3], and clear, in [4], but "timid" being too far against the mainstream, for his young age). In this way quantum theory could grow from the original formulations of Bohr, Sommerfeld, Eherenfest relying on PT to the final formulations of Heisenberg and Schrödinger quite far from it. Nevertheless the triumph of quantum theory was quite substantially based on the technical development and refinement of the methods of formal PT: the calculation of the Compton scattering, the Lamb shift, Fermi's weak interactions model and other spectacular successes came in spite of the parallel recognition that the some of the series that were being laboriously computed not only could not possibly be convergent but their very existence, to all orders n, was in doubt.

The later Feynman graphs representation of PT was a great new tool which superseded and improved earlier graphical representations of the calculations. Its simplicity allowed a careful analysis and understanding of cases in which even *formal* PT seemed puzzlingly failing.

Renormalization theory was developed to show that the convergence problems that seemed to plague even the computation of the individual coefficients of the series, hence the formal PT series at fixed order, were, in reality, often absent, in great generality, as suspected by the earlier treatments of special (important) cases, like the higher order evaluations of the Compton scattering, and other quantum electrodynamics cross sections or anomalous characteristic constants (*e.g.* the magnetic moment of the muon).

5 Mathematics and Physics. Renormalization

In 1943 the first important result on the convergence of the series of the Lindstedt kind was obtained by Siegel, [5]: a formal PT series, of interest in the theory of complex maps, was shown to be convergent. Siegel's work was certainly a stimulus for the later work of Kolmogorov who solved, [6], a problem that had been considered not soluble by many: to find the convergence conditions and the convergence proof of the Lindstedt series for the quasi periodic motions of a generic analytic Hamiltonian system, in spite of Poincaré's theorem and actually avoiding contradiction with it. Thus showing the soundness of the comments about the unsatisfactory aspects of Poincaré's analysis that had been raised almost immediately by Weierstrass, Hadamard and others.

In 1956 not only Kolmogorov theorem appeared but also convergence of another well known and widely used formal series, the virial series, was achieved in an unnoticed work by Morrey, [7], and independently rediscovered in the early 1960's.

At this time it seems that all series with well defined terms were thought to be either convergent or at least asymptotic: for most Physicists convergence or asymptoticity were considered of little interest and matters to be left to Mathematicians.

However with the understanding of the formal aspects of renormalization theory the interest in the convergence properties of the formal PT series came back to the center of attention.

On the one hand mathematical proofs of the existence of the PT series, for interesting quantum fields models, to all orders were investigated settling the question once and for all (Hepp's theorem, [8]); on the other hand it was obvious that even if convergent (like in the virial or Meyer expansions, or in the Kolmogorov theory) it was well understood that the radius of convergence would not be large enough to cover all the physically interesting cases. The sum of the series would in general become singular in the sense of analytic functions and, even if admitting analytic continuation beyond the radius of convergence, a singularity in ε would be eventually hit. The singularity was supposed to correspond to very important phenomena, like the critical point in statistical mechanics or the onset of chaotic motions (already foreseen by Poincaré in connection with his non convergence theorem). Thus research developed in two direction.

The first aimed at understanding the nature of the singularities from the formal series coefficients: in the 1960's many works achieved the understanding of the scaling laws (*i.e.* some properties of the divergences appearing at the singularities of the PT series or of its analytic continuation, for instance in the work of M.Fisher, Kadanoff, Widom and may others). This led to trying to find *resummations*, *i.e.* to collect terms of the formal series to transform them into *convergent series* in terms of *new parameters*, the *running couplings*.

The latter would be singular functions of the original ε thus possibly reducing the study of the singularity to the singularities of the running couplings. The latter could be studied by independent methods, typically by studying the iterations of an auxiliary dynamical system (called the *beta function flow*). This was the *approach* or *renormalization group method* of Wilson, [9, 10].

The second direction was dedicated to finding out the real meaning of the PT series in the cases in which convergence was doubtful or *a priori* excluded: in fact already Landau had advanced the idea that the series could be just illusions in important problems like the paradigmatic quantum field theory of a scalar field or the fundamental quantum electrodynamics, [11, 12].

In a rigorous treatment the function that the series were supposed to represent would be in fact a trivial function with a dependence on ε unrelated to the coefficients of the well defined and non trivial but formal series. It was therefore important to show that there were at least cases in which the perturbation series of a nontrivial problem had a meaning determined by its coefficients. This was studied in the scalar model of quantum field theory and a proof of "non triviality" was achieved after the ground breaking work of Nelson on two dimensional models, [13, 14]: soon followed by the similar results in two dimensions and the difficult extension to three dimensional models by Glimm and Jaffe, [15], and generating many works and results on the subject which took the name of "constructive field theory", [16].

But Landau's *triviality conjecture* was actually dealing with the "real problem", *i.e.* the 4-dimensional quantum fields. The conjecture remains such at the moment, in spite of very intensive work and attempts at its proof. The problem had relevance because it could have meant that not only the simple scalar models of constructive field theory were trivial but also the QED series which had received strong experimental support with the correct prediction of fine structure phenomena could be illusions, in spite of their well defined PT series: which would be remain as mirages of a non existing reality.

The work of Wilson made clear that the "triviality conjecture" of Landau could be applied only to theories which, after the mentioned resummations, would be controlled by a beta function flow that could not be studied perturbatively, and introduced the new notion of *asymptotic freedom*. This is a property of the beta function flow, implying that the running couplings are bounded and small so that the resummed series are more likely to have a meaning, [9].

This work revived the interest in PT for quantum fields with attention devoted to new models that had been believed to be non renormalizable. Once more the apparently preliminary problem of developing a formal PT series plaid a key role: and it was discovered that many Yang-Mills quantum field theories were in fact renormalizable in the ultraviolet region, [17, 18], and an exciting period followed with attempts at using Wilson's methods to give a meaning to the Yang-Mills theory with the hope of building a theory of the strong interactions. Thus it was discovered that several Yang Mills theories were asymptotically free as a consequence of the high symmetry of the model, proving that what seemed to be strong evidence that no renormalizable model would have asymptotic freedom was an ill founded believe (that in a sense slowed down the process of understanding, and not only of the strong interactions).

Suddenly understanding the strong interactions, until then considered an impossible problem became possible, [12], as solutions could be written and *effectively computed* in terms of PT which, although not proved to be convergent or asymptotic (still an open problem in dimension $d = 4$) were immune to the argument of Landau. The impact of the new developments lead a little later to the unification of all interactions into the *standard model* for the theory of elementary particles (including the electromagnetic and weak interactions). The standard model was shown be asymptotically free *even in presence of symmetry breaking*, at least if a few other interactions in the model (for instance the Higgs particle self interaction) were treated heuristically while waiting for the discovery of the "Higgs particle" and for a better understanding of the structure of the elementary particles at length scales intermediate between the Fermi scale ($\sim 10^{-15}\, cm$ (the weak interactions scale) and the Planck scale (the gravitational interaction scale, 15 orders of magnitude below).

Given that the very discovery of renormalizability of Yang-Mills fields and the birth of a strong interactions theory had been firmly grounded on experimental results, [12], the latter "missing step" was, and still is, considered an acceptable gap.

6 Need of convergence proofs

The story of the standard model is paradigmatic of the power of PT: it should convince anyone that the analysis of formal series, including their representation by diagrams, which plays an essential part, is to be taken seriously. PT is certainly responsible for the revival and solution of problems considered by many as hopeless.

In a sense PT in the elementary particles domain can only, so far, partially be considered a success. Different is the situation in the developments that followed the works of Siegel and Kolmogorov. Their relevance for Celestial Mechanics and for several problems in applied physics (particle accelerator design, nuclear fusion machines for instance) and for statistical mechanics made them too the object of a large amount of research work.

The problems are simpler to formulate and often very well posed but the possibility of existence of chaotic motions, always looming, made it imperative not to be content with heuristic analysis and imposed the quest of mathematically complete studies. The lead were the works of Siegel and Kolmogorov. They had established convergence of certain PT series, but there were other series which would certainly be not convergent even though formally well defined and the question was, therefore, which would be their meaning.

More precisely it was clear that the series could be used to find approximate solutions to the equations, representing the motion for very long times under the assumption of "small enough" ε. But this could hardly be considered an understanding of the PT series in Mechanics: the estimated values of ε would have to be too small to be of interest, with the exception of a few special cases. The real question was what could be done to give the PT series the status of exact solution.

As we shall see the problem is deeply connected with the above mentioned asymptotic freedom: this is perhaps not surprising because the link between the two is to be found in the "multiscale analysis" problems, which in the last half century have been the core of the studies in many areas of Analysis and in Physics, when theoretical developments and experimental techniques became finer and able to explore nature at smaller and smaller scales.

7 Multiscale analysis

To illustrate the multiscale analysis in PT it is convenient to present it in the context of Hamiltonian mechanics, because in this field it provides us with nontrivial cases of almost complete success.

We begin by contrasting the work of Siegel and that of Kolmogorov: which are based on radically different methods. The first being much closer in spirit to the developments of renormalization theory and to the Feynman graphs.

Most interesting formal PT series have a common feature: namely their n-th order coefficients are constructed as sums of many "terms" and the first attempt to a complete analysis is to recognize that their sum, which gives the uniquely defined n-th coefficient is much smaller than the sum of the absolute values of the constituent terms. This is a property usually referred as a "cancellation" and, as a rule, it reflects some symmetry property of the problem: hence one possible approach is to look for expressions of the coefficients and for cancellations which would reduce the estimate of the n-th order coefficients, very often of the order of a power of $n!$, to an exponential estimate $O(\varrho^{-n})$ for some $\varrho > 0$ yielding convergence (parenthetically in the mentioned case of Yang-Mills theories the reduction is even more dramatic as it leads from divergent expressions to finite ones, yet of order n!)

The multiscale aspect becomes clear also in Kolmogorov's method because the implicit functions theorem has to be applied over and over again and deals with functions implicitly defined on smaller and smaller domains, [19, 20]. But the method purposedly avoids facing the combinatorial aspects behind the cancellations so much followed, and cherished, [17], in the Physics works.

Siegel's method was developed to study a problem in which no grouping of terms was eventually needed, even though this was by no means clear *a priori*, [21]; and to realize that no cancellations were needed forced to consider the problem as a multiscale one because the absence of rapid growth of the n-th order coefficients became manifest after a suitable "hierarchical ordering" of the terms generating the coefficients. The approach establishes a strong connection with the Physics literature because the technique to study such cases was independently developed in quantum field theory with renormalization, as shown by Hepp in [8], relying strongly on it. This is very natural and, in case of failure, it can be improved by looking for "resummations" turning the power series into a convergent series in terms of functions of ε which are singular but controllably so. For details see below and [22].

What is "natural", however, is a very personal notion and it is not surprising that what some consider natural is considered unnatural or clumsy or difficult (or the three qualifications together) by others.

Conflict arises when the same problem can be solved by two different "natural" methods. and in the case of PT for Hamiltonian systems close to integrable ones (closeness depending on the size of a parameter ε), the so called "small denominators" problem, the methods of Siegel and Kolmogorov are antithetic and an example of the just mentioned dualism.

The first method, that will be called here "Siegel's method" (see below for details), is based on a careful analysis of the structure of the various terms that occur at a given PT order achieving a proof that the n-th order coefficient which is represented as the sum of many terms some of which *might* have size of order of a power of n! has in fact a size of $O(\varrho^{-n})$ so that the PT series is convergent for $|\varepsilon| < \varrho$. Although strictly speaking original work of Siegel does not immediately apply to the Hamiltonian Mechanics problems (see below), it can nevertheless be adapted and yields a solution, as made manifest much later in [21, 23, 22].

The second method, called here "Kolmogorov's method", instead does not consider the individual coefficients of the various orders but just regards the sum of the series as a solution of an implicit function equation (a "Hamilton-Jacobi" equation) and devises a recursive algorithm approximating the unknown sum of the PT series by functions analytic in a disk of fixed radius ϱ in the complex ε-plane, [19, 20].

Of course the latter approach implies that no matter how we achieve the construction of the n-th order PT series coefficient there will have to be enough cancellations, if at all needed, so that it turns out bounded by $O(\varrho^{-n})$. And in the problem studied by Kolmogorov cancellations would be necessarily present if the n-th order coefficient was represented by the sum of the terms in the

Lindstedt series.

That this is not obvious is supported by the fact that it was considered an open problem, for about thirty years, to find a way to exhibit explicitly the cancellation mechanism in the Lindstedt series implied by Kolmogorov's work. This was done by Eliasson, [23], who proved that the coefficients of the PT of a given order n as expressed by the construction known as the "Lindstedt algorithm" yielded coefficients of size of $O(\varrho^{-n})$: his argument, however, did not identify in general which term of the Lindstedt sum for the n-th order coefficient was compensated by which other term or terms. It proved that the sum had to satisfy suitable relations, which in turn implied a total size of $O(\varrho^{-n})$. And it took a few more years for the complete identification, [22], of the rules to follow in collecting the terms of the Lindstedt series which would imply the needed cancellations.

It is interesting to remark that, aside from the example of Hamiltonian PT, multiscale problems have dominated the development of analysis and Physics in recent time: for instance they appear in harmonic analysis (Carleson, Fefferman), in PDE's (DeGiorgi, Moser, Caffarelli-Kohn-Ninberg) in relativistic quantum mechanics (Glimm, Jaffe, Wilson) in Hamiltonian Mechanics (Siegel, Kolmogorov, Arnold, Moser) in statistical mechanics and condensed matter (Fisher, Wilson, Widom)... Sometimes, although not always, studied by PT techniques, [10].

8 A paradigmatic example of PT problem

It is useful to keep in mind an example illustrating technically what it mean to perform a multiscale analysis in PT. And the case of quasi periodic motions in Hamiltonian mechanics will be selected here, being perhaps the simplest.

Consider the motion of ℓ unit masses on a unit circle and let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_\ell)$ be their positions on the circle, *i.e.* $\boldsymbol{\alpha}$ is a point on the on the torus $\mathcal{T}^\ell = [0, 2\pi]^\ell$. The points interact with a potential energy $\varepsilon f(\boldsymbol{\alpha})$ where ε is a strength parameter and f is a trigonometric even polynomial, of degree N: $f(\boldsymbol{\alpha}) = \sum_{\boldsymbol{\nu}\in\mathcal{Z}^\ell,\, |\boldsymbol{\nu}|\le N} f_{\boldsymbol{\nu}} e^{i\boldsymbol{\nu}\cdot\boldsymbol{\alpha}}$, $f_{\boldsymbol{\nu}} = f_{-\boldsymbol{\nu}} \in \mathcal{R}$, where $\mathcal{Z}^\ell$ denotes the lattice of the points with integer components in $\mathcal{R}^\ell$ and $|\boldsymbol{\nu}| = \sum_j |\nu_j|$.

Let $t \to \boldsymbol{\alpha}_0 + \boldsymbol{\omega}_0 t$ be the motion with initial data, at time $t = 0$, $\boldsymbol{\alpha}(0) = \boldsymbol{\alpha}_0, \dot{\boldsymbol{\alpha}}(0) = \boldsymbol{\omega}_0$, in which all particles rotate at constant speed with rotation velocity $\boldsymbol{\omega}_0 = (\omega_{01}, \ldots, \omega_{0\ell}) \in \mathcal{R}^\ell$. This is a solution for the equations of motion for $\varepsilon = 0$ and it is a quasi periodic solution, *i.e.* each of the angles α_j rotates periodically at constant speed ω_{0j}, $j = 1, \ldots, \ell$.

The motion will be called *non resonant* if the components of the rotation speed $\boldsymbol{\omega}_0$ are rationally independent: this means that $\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu} = 0$ with $\boldsymbol{\nu} \in \mathcal{Z}^\ell$ is possible only if $\boldsymbol{\nu} = \mathbf{0}$. In this case the motion $t \to \boldsymbol{\alpha}_0 + \boldsymbol{\omega}_0 t$ covers, $\forall\, \boldsymbol{\alpha}_0$, densely the torus $\mathcal{T}^\ell$ as t varies. The PT problem that we consider is to find whether there is a family of motions "of the same kind" for each ε, small enough, solving the equations of motion; more precisely whether there exists a function $\mathbf{a}_\varepsilon(\boldsymbol{\varphi})$, $\boldsymbol{\varphi} \in \mathcal{T}^\ell$, such that setting

$$\boldsymbol{\alpha}(t) = \boldsymbol{\varphi} + \boldsymbol{\omega}_0 t + \mathbf{a}_\varepsilon(\boldsymbol{\varphi} + \boldsymbol{\omega}_0 t), \qquad \text{for } \boldsymbol{\varphi} \in \mathcal{T}^\ell \tag{8.1}$$

one obtains, $\forall\, \boldsymbol{\varphi} \in \mathcal{T}^\ell$ and for ε small enough, a solution of the equations of motion for a force $-\varepsilon\partial_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha})$: *i.e.*

$$\ddot{\boldsymbol{\alpha}}(t) = -\varepsilon\partial_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha}(t)). \tag{8.2}$$

By substitution of Eq. (8.1) in Eq. (8.2), the condition becomes $(\boldsymbol{\omega}_0 \cdot \partial_{\boldsymbol{\varphi}})^2 \mathbf{a}(\boldsymbol{\varphi} + \boldsymbol{\omega}_0 t) = -\partial_{\boldsymbol{\alpha}} f(\boldsymbol{\varphi} + \boldsymbol{\omega}_0 t)$. Since $\boldsymbol{\omega}_0$ is assumed rationally independent $\boldsymbol{\varphi} + \boldsymbol{\omega}_0 t$ covers densely the torus $\mathcal{T}^\ell$ as t varies: hence the equation for $\mathbf{a}_\varepsilon$ is

$$(\boldsymbol{\omega}_0 \cdot \partial_{\boldsymbol{\varphi}})^2 \mathbf{a}_\varepsilon(\boldsymbol{\varphi}) = -\varepsilon\, \partial_{\boldsymbol{\alpha}} f(\boldsymbol{\varphi} + \mathbf{a}_\varepsilon(\boldsymbol{\varphi})) \tag{8.3}$$

Applying PT to this equation means to look for a solution $\mathbf{a}_\varepsilon$ which is analytic in ε small enough and in $\boldsymbol{\varphi} \in \mathcal{T}^\ell$. In colorful language one says that the perturbation effect is of slightly deforming a nonresonant torus with given frequency spectrum (*i.e.* given $\boldsymbol{\omega}_0$) on which the motion develops, without destroying it and keeping the quasi periodic motion on it with the same frequency spectrum.

9 Lindstedt series

As it follows from a very simple special case of Poincaré's work, Eq. (2) cannot be solved if also $\boldsymbol{\omega}_0$ is considered variable and the dependence on $\varepsilon, \boldsymbol{\omega}_0$ analytic. Nevertheless if $\boldsymbol{\omega}_0$ is *fixed and non resonant* and if $\mathbf{a}_\varepsilon$ is supposed analytic in ε small enough and in $\boldsymbol{\varphi} \in \mathcal{T}^\ell$, then there can be at most one solution to the Eq. (2) with $\mathbf{a}_\varepsilon(\mathbf{0}) = \mathbf{0}$ (which is not a real restriction because if $\mathbf{a}_\varepsilon(\boldsymbol{\varphi})$ is a solution also $\mathbf{a}_\varepsilon(\boldsymbol{\varphi} + \mathbf{c}_\varepsilon) + \mathbf{c}_\varepsilon$ is a solution for any constant $\mathbf{c}_\varepsilon$). This so because the coefficients of the power series in ε, $\sum_{n=1}^\infty \varepsilon^n \mathbf{a}_n(\boldsymbol{\varphi})$, are uniquely determined if the series is convergent. In fact they are trigonometric polynomials of order $\le nN$ which will be written as

$$\boldsymbol{\alpha}_n(\boldsymbol{\varphi}) = \sum_{0<|\boldsymbol{\nu}|\le Nn} \boldsymbol{\alpha}_{n,\boldsymbol{\nu}} e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}} \tag{9.1}$$

It is convenient to express them in terms of graphs. The graphs to use to express the value $\mathbf{a}_{n,\boldsymbol{\nu}}$ are

(i) trees with n *nodes* $v_1, \ldots, v_n$,
(ii) one root r,
(iii) n lines joining pairs (v_i', v_i) of nodes or the root and one node, *always one and not more*, (r, v_i);
(iv) the lines will be different from each other and distinguished by a mark label, $1, \ldots, n$ attached to them. The connections between the nodes that the lines generate have to be *loopless*, *i.e.* the graph formed by the lines must be a tree.
(v) The tree lines will be imagined oriented towards the root: hence a partial order is generated on the tree and the line joining v to v' will be denoted $\lambda_{v'v}$

and v' will be the node closer to the root immediately following v, hence such that $v' > v$ in the partial order of tree.

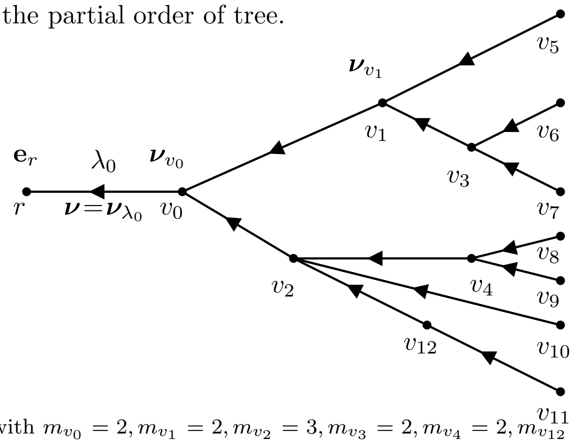

Fig.1: A tree θ with $m_{v_0} = 2, m_{v_1} = 2, m_{v_2} = 3, m_{v_3} = 2, m_{v_4} = 2, m_{v_{12}} = 1$ lines entering the nodes v_i, $k = 13$. Some labels or decorations explicitly marked (on the lines λ_0, λ_1 and on the nodes v_1, v_2); the number labels, distinguishing the branches, are not shown. The arrows represent the partial ordering on the tree.

The number of such trees is large and exactly equal to n^{n-1}, as an application of Cayley's formula implies: their collection will be denoted T_n^0.

To compute $\mathbf{a}_{n,\boldsymbol{\nu}}$ consider all trees in T_n^0 and attach to each node v a vector $\boldsymbol{\nu}_v \in \mathcal{Z}^\ell$, called "mode label", such that $f_{\boldsymbol{\nu}_v} \neq 0$, hence $|\boldsymbol{\nu}_v| \le N$. To the root we associate one of the coordinate unit vectors $\boldsymbol{\nu}_r \equiv \mathbf{e}_r$. We obtain a set T_n of *decorated trees* (with $\le (2N+1)^{\ell n} n^{n-1}$ elements, by the above counting analysis).

Given $\theta \in T_n$ and $\lambda = \lambda(v', v) \in \theta$ we define the *current* on the line λ to be the vector $\boldsymbol{\nu}(\lambda) \equiv \boldsymbol{\nu}(v', v) \stackrel{def}{=} = \sum_{w \le v} \boldsymbol{\nu}_w$: *i.e.* we imagine that the node vectors $\boldsymbol{\nu}_{v_i}$ represent currents entering the node v_i and flowing towards the root. Then $\boldsymbol{\nu}(\lambda)$ is, for each $|l$, the sum of the currents which entered all the nodes not following v, *i.e.* current accumulated after passing the node v.

The current flowing in the root line $\boldsymbol{\nu} = \sum_v \boldsymbol{\nu}_v$ will be denoted $\boldsymbol{\nu}(\theta)$.

Let T_n^* be the set trees in T_n in which *all lines* carry a *non zero* current $\boldsymbol{\nu}(\lambda) \neq \mathbf{0}$. A *value* $\text{Val}(\theta)$ will be defined, for $\theta \in T_n^*$, by a product of node factors and of line factors over all nodes and

$$\text{Val}(\theta) = \frac{i(-1)^n}{n!} \prod_{v\in\theta} f_{\boldsymbol{\nu}_v} \prod_{\lambda=(v',v)} \frac{\boldsymbol{\nu}_{v'} \cdot \boldsymbol{\nu}_v}{(\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}(v',v))^2} \tag{9.2}$$

The coefficient $\mathbf{a}_{n,\boldsymbol{\nu}}$ will then be

$$\mathbf{a}_{n,\boldsymbol{\nu}} = \sum_{\substack{\theta\in T_n^* \\ \boldsymbol{\nu}(\theta)=\boldsymbol{\nu}}} \text{Val}(\theta) \tag{9.3}$$

and, when the coefficients are imagined to be constructed in this way, the formal power series $\sum_{n=1}^\infty \varepsilon^n \sum_{|\boldsymbol{\nu}|\le Nn} \mathbf{a}_{n,\boldsymbol{\nu}}$ is called the "Lindstedt series". Eq. (9.2) and its graphical interpretation in Fig.1 should be considered the "Feynman rules" and the "Feynman diagrams" of the PT for Eq. (8.3), [24, 10].

10 Convergence. Scales. Multiscale analysis.

The Lindstedt series is well defined because of the non resonance condition and the n-th term is not even a sum of too many terms: if $F \stackrel{def}{=} \max_{\boldsymbol{\nu}} |f_{\boldsymbol{\nu}}|$, each of them can be bounded by $\frac{F^n}{n!} \prod_{\lambda\in\theta} \frac{N^2}{\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}(\lambda)^2}$; hence their sum can be bounded, if G is such that $\frac{(2N+1)^{\ell n} n^{n-1} F^n}{n!} \le G^n$, by $G^n \prod_{\lambda\in\theta} \frac{N^2}{\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}(\lambda)^2}$.

Thus all $\mathbf{a}_n$ are well defined and finite *but* the problem is that $|\boldsymbol{\nu}(\lambda)|$ can be large (up to Nn at given order n) and therefore $\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}(\lambda)$ although never zero can become very small as n grows. For this reason the problem of convergence of the series is an example of what is called a *small denominators problem*. And it is necessary to assume more than just non resonance of $\boldsymbol{\omega}_0$ in order to solve it in the present case: a simple condition is the *Diophantine* condition, namely the existence of $C, \tau > 0$ such that

$$|\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}| \ge \frac{1}{C\,|\boldsymbol{\nu}|^\tau}, \qquad \forall\, \mathbf{0} \ne \boldsymbol{\nu} \in \mathcal{Z}^\ell \tag{10.1}$$

But this condition is not sufficient in an obvious way: because it only allows us to bound individual tree-values by $n!^a$ for some $a > 0$ related to τ; furthermore it is not difficult to check that there are single graphs whose value is actually of "factorial" size in n. Although non trivial to see (as mentioned above) this was only apparently so in the earlier case of Siegel's problem but it is the new essential feature of the terms generating the n-th order coefficient in Eq. (9.3).

A resummation is necessary to show that the tree-values can be grouped so that the sum of the values of each group can be bounded by ϱ^{-n} for some $\varrho > 0$ and $\forall\, n$, although the group may contain (several) terms of factorial size. The terms to be grouped have to be ordered hierarchically according to the sizes of the line factors $\frac{1}{(\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}(\lambda))^2}$, which are called *propagators* in [22, 25].

A similar problem is met in quantum field theory where the graphs are the *Feynman graphs*: such graphs can only have a small number of lines that converge into a node but they can have loops, and to show that the perturbation series is well defined to all orders it is also necessary to collect terms hierarchically according to the propagators sizes. The systematic way was developed by Hepp, [8, 26] for the PT expansion of the Schwinger functions in quantum field theory of scalar fields, [16]. It has been used in many occasions later and it plays a key role in the renormalization group methods in Statistical Mechanics (for instance in theory of the ground state of Fermi systems), [27, 10].

However it is in the Lindstedt series that the method is perhaps best illustrated. Essentially because it ends up in a convergence proof, while often in the field theory or statistical mechanics problems the PT series can be only proved to be well defined to all orders, but they are seldom, if ever, convergent so that one has to have recourse to other supplementary analytic means to show that the PT series are asymptotic (in the cases in which they are such).

The path of the proof is the following.

(1) consider only trees in which no two lines λ_+ and λ_-, with λ_+ following

λ_- in the partial order of the tree, have the *same* current $\boldsymbol{\nu}_0$. In this case the maximum of the $\prod_\lambda \frac{1}{(\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}(\lambda))^2}$ over all tress $\theta \in T^*_n$ can be bounded by G_1^n for some G_1.
This is an immediate consequence and the main result in the original Siegel's work, [5], which dealt with a different problem with small denominators in its formal PT solution: the coefficients of the series could also be represented by tree graphs, very similar too the ones above: but the only allowed $\boldsymbol{\nu} \in \mathcal{Z}^\ell$ were the non zero vectors with all components ≥ 0.
The latter property automatically guarantees that the graphs contain no pair of lines λ_+, λ_- following each other as above in the tree partial order and having the same current. Siegel's proof also implies a multiscale analysis, [21]: but it requires no grouping of the terms unlike the analogue Lindstedt series, Eq. (9.3).

(2) Trees which contain lines λ_+ and λ_-, with λ_+ following λ_-, in the partial order of the tree, and having the *same* current $\boldsymbol{\nu}_0$ can have values which have size of order $O(n!^a)$ with some $a > 0$. Collecting terms is therefore essential.
A line λ of a tree is said to have scale k if $2^{-k-1} \leq \frac{1}{C|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}|} < 2^{-k}$. The lines of a tree $\theta \in T^*_n$ can then be collected in *clusters*.[1]
A cluster of scale p is a maximal connected set of lines of scale $k \geq p$ with at least one line of scale p. Clusters are connected to the rest of the tree by lines of lower scale which can be *incoming* or *outgoing* with respect to the partial ordering. Clusters also contain nodes: a node is in a cluster if it is an extreme of a line contained in a cluster; such nodes are said *internal* to the cluster.

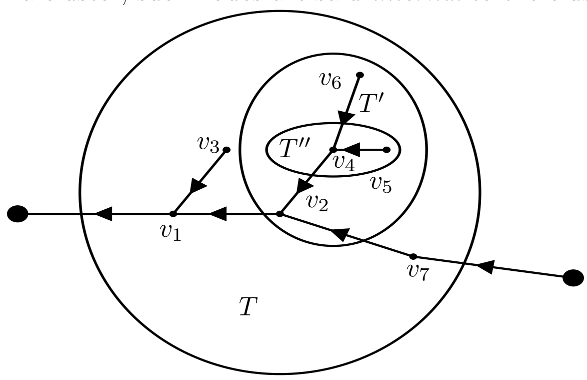

Fig.2: An example of three clusters symbolically delimited by circles, as visual aids, inside a tree (whose remaining branches and clusters are not drawn and are indicated by the bullets); not all labels are explicitly shown. The scales (not marked) of the branches increase as one crosses inward the circles boundaries: recall, however, that the scale labels are integers ≤ 1 (hence typically ≤ 0). The $\boldsymbol{\nu}$ labels are not drawn (but must be imagined). If the $\boldsymbol{\nu}$ labels of (v_4, v_5) add up to $\mathbf{0}$ the cluster T'' is a self-energy graph. If the $\boldsymbol{\nu}$ labels of (v_2, v_4, v_5, v_6) add up to $\mathbf{0}$ the cluster T' is a self-energy graph and such is T if the $\boldsymbol{\nu}$ labels of $(v_1, v_2, v_3, v_4, v_5, v_6, v_7)$ add up to $\mathbf{0}$. The cluster T' is maximal in T.

Of particular interest are the *self energy* clusters. These are clusters with only

[1] The scaling factor 2 is arbitrary: any scale factor > 1 could be used.

one incoming line and only one outgoing line which *furthermore* have the same current $\boldsymbol{\nu}_0$. To simplify the analysis the Diophantine condition can be strengthened to insure that if in a tree graph the line incoming into a self energy cluster and ending in an internal node v is detached from the node v and reattached to another node internal to the same cluster which is not in a self-energy subcluster (if any) then the new tree nodes are still enclosed in the same clusters. Alternatively the definition of scale of a line can be modified slightly to achieve the same goal.

(3) Then it makes sense to sum together all the values of the trees whose nodes are collected into the same families of clusters and differ only because the lines entering the self energy clusters are attached to a different node internal to the cluster, but external to the inner self energy subclusters (if any). Furthermore the value of the trees obtained by changing simultaneously sign to the $\boldsymbol{\nu}_v$ of the nodes inside the self energy clusters have also to be added together.

After collecting the terms in the described way it is possible to check that each sum of terms so collected is bounded by ϱ_0^{-n} for some ϱ_0 (which can also be estimated explicitly). Since the number of addends left in not larger than the original one the bound on $\sum_{\boldsymbol{\nu}} |\mathbf{a}_{n,\boldsymbol{\nu}}|$ becomes $\le \frac{F^n(2N+1)^\ell n^n N^{2n}}{n!}\varrho_0^{-n} \le \varrho^{-n}$, for suitable ϱ_0, ϱ, so that convergence of the formal series for $\mathbf{a}_\varepsilon(\boldsymbol{\varphi})$ is achieved for $|\varepsilon| < \varrho$, see [22].

11 Non convergent cases

Convergence is *not* the rule: very interesting problems arise in which the PT series is, or is believed to be, only asymptotic. For instance in quantum field theory the PT series are well defined but they are not convergent: they can be proved, in the scalar φ^4 theories in dimension 2 and 3 to be asymptotic series for a function of ε which is *Borel summable*: this means in particular that the solution can be in principle recovered, for $\varepsilon > 0$ and small, just from the coefficients of its formal expansion.

Other non convergent expansions occur in statistical mechanics, for example in the theory of the ground state of a Fermi gas of particles on a lattice of obstacles. This is still an open problem, and a rather important one. Or occur in quantum field theory where sometimes they can be proved to be Borel summable.

The simplest instances again arise in Mechanics in studying *resonant quasi periodic motions*. A paradigmatic case is provided by Eqs. (8.1),(8.2) when $\boldsymbol{\omega}_0$ has some vanishing components: $\boldsymbol{\omega}_0 = (\omega_1, \ldots, \omega_r, 0, \ldots, 0) = (\widetilde{\boldsymbol{\omega}}_0, \mathbf{0})$ with $1 < r < \ell$. If one writes $\boldsymbol{\alpha} = (\widetilde{\boldsymbol{\alpha}}, \widetilde{\boldsymbol{\beta}}) \in \mathcal{T}^r \times \mathcal{T}^{\ell-r}$ and looks motions like Eq. (8.1) of the form

$$\begin{aligned}\widetilde{\boldsymbol{\alpha}}(t) =& \widetilde{\boldsymbol{\varphi}} + \widetilde{\boldsymbol{\omega}}_0 t + \widetilde{\mathbf{a}}_\varepsilon(\widetilde{\boldsymbol{\varphi}} + \widetilde{\boldsymbol{\omega}}_0 t)\\ \widetilde{\boldsymbol{\beta}}(t) =& \boldsymbol{\beta}_0 + \widetilde{\mathbf{b}}_\varepsilon(\widetilde{\boldsymbol{\varphi}} + \widetilde{\boldsymbol{\omega}}_0 t)\end{aligned} \tag{11.1}$$

where $\widetilde{\mathbf{a}}_\varepsilon(\widetilde{\boldsymbol{\varphi}}), \widetilde{\mathbf{b}}_\varepsilon(\widetilde{\boldsymbol{\varphi}})$ are functions of $\widetilde{\boldsymbol{\varphi}} \in \mathcal{T}^r$, analytic in ε and $\widetilde{\boldsymbol{\varphi}}$.

In this case the analogue of the Lindstedt series can be devised provided $\boldsymbol{\beta}_0$ is chosen to be a stationary point for the function $\widetilde{f}(\widetilde{\boldsymbol{\beta}}) = \int f(\widetilde{\boldsymbol{\alpha}}, \widetilde{\boldsymbol{\beta}}) \frac{d\widetilde{\boldsymbol{\alpha}}}{(2\pi)^r}$, and provided $\widetilde{\boldsymbol{\omega}}_0$ satisfies a Diophantine property $|\widetilde{\boldsymbol{\omega}}_0 \cdot \widetilde{\boldsymbol{\nu}}| > \frac{1}{C|\widetilde{\boldsymbol{\nu}}|^\tau}$ for all $\mathbf{0} \neq \widetilde{\boldsymbol{\nu}} \in \mathcal{Z}^r$ and for τ, C suitably chosen.

This time the series is likely to be, in general, non convergent (although there is not a proof yet). And the terms of the Lindstedt series can be suitably collected to improve the estimates. Nevertheless the estimates cannot be improved enough to obtain convergence. Deeper resummations are needed to show that in some cases the terms of the series can be collected and rearranged into a convergent series.

The resummation is deeper in the sense that it is not enough to collect terms contributing to a given order in ε but it is necessary to collect and sum terms of different order according to the following scheme.

(1) the terms of the Lindstedt series are first "regularized" so that the new series is manifestly analytic in ε with, however, a radius of convergence depending on the regularization. For instance one can consider only terms with lines of scale $\le M$.

(2) terms of different orders in ε are then summed together and the series becomes a series in powers of functions $\lambda_j(\varepsilon; M)$ of ε with very small radius of convergence in ε, *but* with an M-independent radius of convergence ϱ in the $\lambda_j(\varepsilon, M)$. The labels $j = 0, 1, \ldots, M$ are scale labels whose value is determined by the order in which they are generated in the hierarchical organization of the collection of the graphs according to their scales.

(3) one shows that the functions $\lambda_j(\varepsilon; M)$ ("running couplings") can be analytically continued in ε to an *M-independent domain $\mathcal{D}$* containing the origin in its closure and where they remain smaller than ϱ for all M. Furthermore $\lambda_j(\varepsilon; M) \xrightarrow[M\to\infty]{} \lambda_j(\varepsilon)$, for $\varepsilon \in \mathcal{D}$.

(4) the convergent power series in the running couplings admits an asymptotic series in ε at the origin which coincides with the formal Lindstedt series. Hence in the domain $\mathcal{D}$ a meaning is attributed to the sum of Lindstedt series.

(5) one checks that the functions $\widetilde{\mathbf{a}}_\varepsilon, \widetilde{\mathbf{b}}_\varepsilon$ thus defined are such that Eq. (11.1) satisfies the equations of motion Eq. (8.1).

The proof can be completed if the domain $\mathcal{D}$ contains real points ε.

If $\widetilde{\boldsymbol{\beta}}_0$ is a maximum point the domain $\mathcal{D}$ contains a circle tangent to the origin and centered on the positive real axis. So in this case the $\widetilde{\mathbf{a}}_\varepsilon, \widetilde{\mathbf{b}}_\varepsilon$ are constructed in $\mathcal{D} \cap \mathcal{R}_+$, $\mathcal{R}_+ \stackrel{def}{=} (0, +\infty)$.

If instead $\widetilde{\boldsymbol{\beta}}_0$ is a minimum point the domain $\mathcal{D}$ exists but $\mathcal{D} \cap \mathcal{R}_+$ touches the positive real axis on a set of points with positive measure and density 1 at the origin. So $\widetilde{\mathbf{a}}_\varepsilon, \widetilde{\mathbf{b}}_\varepsilon$ are constructed only for ε in this set which is a kind of "Cantor set",[28].

Again the multiscale analysis is necessary to identify the tree values which have to be collected to define $\lambda_j(\varepsilon; M)$. In this case it is an analysis which is much closer to the similar analysis that is encountered in quantum field theory in the "self energy resummations", which involve collecting and summing graph values of graphs contributing to different orders of perturbation.

The above scheme can also be applied when $r = \ell$, *i.e.* in the case of the classical Lindstedt series when it is actually convergent: this leads to an alternative proof of the Kolmogorov theorem which is interesting as it is even closer to the renormalization group methods because it expresses the solution in terms of a power series in running couplings.[25, Ch.8,9].

12 Conclusion and Outlook

Perturbation theory provides a general approach to the solution of problems "close" to well understood ones, "closeness" being measured by the size of a parameter ε. It naturally consists of two steps: the first is to find a formal solution, under the assumption that the quantities of interest are analytic in ε at $\varepsilon = 0$. If this results in a power series with well defined coefficients then it becomes necessary to find whether the series thus constructed, called *formal series*, converges.

In general the proof that the formal series exists (when it really does) is non-trivial: typically in quantum mechanics problems (quantum fields or statistical mechanics) this is an interesting and deep problem giving rise to renormalization theory. Even in classical mechanics PT of integrable systems it has been, historically, a problem to obtain (in wide generality) the Lindstedt series (of which a simple example is discussed above).

Once existence of a PT series is established, very often the series is not convergent and at best is an asymptotic series. It becomes challenging to find its meaning (if any, as there are cases, even interesting ones, on which conjectures exist claiming that the series have no meaning, like the quantum scalar field in dimension 4 with "φ^4-interaction" or quantum electrodynamics).

Convergence proofs, in most interesting cases, require a multiscale analysis: because the difficulty arises as a consequence of the behavior of singularities at infinitely many scales, as in the case of the Lindstedt series above exemplified.

When convergence is not possible to prove, the multiscale analysis often suggest "resummations", collecting the various terms whose sums yields the formal PT series (usually the algorithms generating the PT series give its terms at given order as sums of simple but many quantities, as in the discussed case of the Lindstedt series). The collection involves adding together terms of different order in ε and results in a new power series, the *resummed series*, in a family of parameters $\lambda_j(\varepsilon)$ which are functions of ε, called the "running couplings", depending on a "scale index" $j = 0, 1, \ldots$.

The running couplings are (in general) singular at $\varepsilon = 0$ as functions of ε but C^∞ there, and obey equations that allow to study and define them independently of a convergence proof. If the running couplings can be shown to be so small, as ε varies in a suitable domain $\mathcal{D}$ near 0, to guarantee convergence of the resummed series and therefore to give a meaning to the PT for $\varepsilon \in \mathcal{D}$ then the PT program can be completed.

The singularities in ε at $\varepsilon = 0$ are therefore all contained in the running couplings, usually very few and the same for various formal series of interest in a given problem.

The idea of expressing the sum of formal series as sum of convergent series in new parameters, the running couplings, determined by other means (a recursion relation denominated the beta function flow) is the key idea of the renormalization group methods: PT in mechanics is a typical and simple example.

On purpose attention has been devoted to PT in the analytic class: but it is possible to use PT techniques in problems in which the functions whose value is studied are not analytic; the techniques are somewhat different and new ideas are needed which would lead quite far away from the natural PT framework which is within the analytic class.

Of course there are many problems of PT in which the formal series are simply convergent and the proof does not require any multiscale analysis. However here attention having been devoted to the novel aspect of PT that emerged in Physics and Mathematics in the last half century and the problems not requiring multiscale analysis have not been considered . It is worth, however, to mention that even in simple convergent PT cases it might be convenient to perform resummations. An example is Kepler's equation

$$\ell = \xi - \varepsilon \sin \xi, \qquad \xi, \ell \in T^1 = [0, 2\pi] \tag{12.1}$$

which can be (easily) solved by PT. The resulting series has a radius of convergence in ε rather small (Laplace's limit): however if a resummation of the series is performed transforming it into a power series in a "running coupling" $\lambda_0(\varepsilon)$ (only 1, because no multiscale analysis is needed, the PT series being convergent) given by, [29, Vol. 2, p. 321]

$$\lambda_0 \stackrel{def}{=} \frac{\varepsilon\, e^{\sqrt{1-\varepsilon^2}}}{1 + \sqrt{1-\varepsilon^2}}. \tag{12.2}$$

The resummed series is a power series in λ_0 with radius of convergence 1 and when ε varies between 0 and 1 the parameter λ_0 corresponding to it goes from 0 to 1. Hence in terms of λ_0 it is possible to invert *by power series* the Kepler equation for all $\varepsilon \in [0, 1)$, *i.e.* in the entire interval of physical interest (recall that ε has the interpretation of eccentricity of an elliptic orbit in the 2-body problem). Resummations can improve convergence properties.

13 Future directions

It is always hard to indicate future directions, which usually turn to different paths. Perturbation theory is an ever evolving subject: it is a continuous source of problems and its applications generate new ones. Examples of outstanding problems are understanding the triviality conjectures of models like quantum φ^4 field theory in dimension 4, [16]; or a development of the theory of the ground states of Fermionic systems in dimensions 2 and 3, [27]; a theory of weakly coupled Anosov flows to obtain information of the kind that it is possible to obtain for weakly coupled Anosov maps, [25]; uniqueness issues in cases in which PT series can be given a meaning, but in a priori non unique way like the resonant quasi periodic motions in nearly integrable Hamiltonian systems, [25].

Acknowledgements: Work partly supported by IHES.

References

[1] P.S. Laplace. *Mécanique Céleste.* reprinted by Chelsea, New York, Paris, 1799.

[2] H. Poincaré. *Les Méthodes nouvelles de la Mécanique céleste.* reprinted by Blanchard, 1987, Paris, 1892.

[3] A. Einstein. Zum quantensatz von Sommerfeld und Epstein. *Verhandlungen der Deutschen physikalischen Gesellshaft*, 19:82–02, 1917.

[4] E. Fermi. Il principio delle adiabatiche e i sistemi che non ammettono coordinate angolari. *Nuovo Cimento*, 25:171–175, 1923.

[5] K. Siegel. Iterations of analytic functions. *Annals of Mathematics*, 43:607–612, 1943.

[6] A.N. Kolmogorov. On the preservation of conditionally periodic motions. *Doklady Akademia Nauk SSSR and Lecture Notes in Physics, Stochastic behavior in classical and quantum Hamiltonians, ed. G. Casati, J. Ford, Vol. 93, 1979*, 96:527–530, 1954.

[7] C.B. Morrey. On the derivation of the equations of hydrodynamics from statistical mechanics. *Communications in Pure and Applied Mathematics*, 8:279–326, 1955.

[8] K. Hepp. Proof of the Bogoliubov-Parasiuk theorem on renmormalization. *Communications in Mathematical Physics*, 2:301–326, 1966.

[9] K. Wilson and J. Kogut. The renormalization group and the ε-expansion. *The renormalization group and the ε-expansion, Physics Reports*, 12:75–199, 1973.

[10] G. Gallavotti. Renormalization group in statistical mechanics and mechanics: gauge symmetries and vanishing beta functions. *Physics Reports*, 352:251–272, 2001.

[11] J. Frölich. On the triviality of $\lambda\varphi^4_d$ theories and the approach to the critical point in $d\,(\geq 4)$ dimensions. *On the triviality of $\lambda\varphi^4_d$ theories and the approach to the critical point in $d\,(\geq 4)$ dimensions, Nuclear Physics B*, 200:281–296, 1982.

[12] D.J. Gross. Twenty five years of asymptotic freedom. *Nuclear Physics B (Proceedings Supplements))*, 74:426–446, 1999.

[13] E. Nelson. A quartic interaction in two dimensions. *A quartic interaction in two dimensions, In Mathematical Theory of elementary particles, ed. R Goodman, I. Segal*, pages 69–73, 1966.

[14] B. Simon. *The $P(\varphi)_2$ Euclidean (quantum) field theory.* Princeton University PRess, Princeton, 1974.

[15] J. Glimm and A. Jaffe. *Quantum Physics. A functional integral point of view.* Springer-Verlag, New York, 1981.

[16] G. Gallavotti. Renormalization theory and ultraviolet stability for scalar fields via renormalization group methods. *Reviews of Modern Physics*, 57:471–562, 1985.

[17] G 't Hooft and MJG Veltman. Regularization and renormalization of gauge fields. *Nuclear Physics B*, 44:189–213, 1972.

[18] G. 't Hooft. When was asymptotic freedom discovered? or the rehabilitation of quantum field theory. *Nuclear Physics B (Proceedings Supplements)*, 74:413–425, 1999.

[19] G. Gallavotti. Perturbation theory for classical Hamiltonian systems. *in Scaling and self similarity in Physics*, Ed. ed. J. Fröhlich, Birkhauser, Boston:359–426, 1985.

[20] G. Gallavotti. Quasi integrable mechanical systems. *Phenomènes Critiques, Systèmes aleatories, Théories de jauge, Proceedings, Les Houches, XLIII (1984)*, II:539–624, 1986.

[21] J. Pöschel. Invariant manifolds of complex analytic mappings. *Phenomènes Critiques, Systèmes aleatories, Théories de jauge, Proceedings, Les Houches, XLIII (1984)*, II:949–964, 1986.

[22] G. Gallavotti. Twistless kam tori. *Communications in Mathematical Physics*, 164:145–156, 1994.

[23] L. H. Eliasson. Absolutely convergent series expansions for quasi–periodic motions. *Mathematical Physics Electronic Journal (MPEJ)*, 2:33 pages, 1996.

[24] G. Gallavotti. Invariant tori: a field theoretic point of view on eliasson's work. *Advances in Dynamical Systems and Quantum Physics,* Ed. R. Figari, World Scientific, 164:117–132, 1995.

[25] G. Gallavotti, F. Bonetto, and G. Gentile. *Aspects of the ergodic, qualitative and statistical theory of motion.* Springer Verlag, Berlin, 2004.

[26] K. Hepp. *Théorie de la rénormalization*, volume 2. Springer Verlag, Lecture notes in Physics, 1969.

[27] G. Gallavotti G. Benfatto. *Renormalization group.* Princeton University Press, Princeton, 1995.

[28] G. Gallavotti and G. Gentile. Degenerate elliptic resonances. *Communications in Mathematical Physics*, 257:319–362, 2005.

[29] T. Levi-Civita. *Opere Matematiche*, volume 2. Accademia Nazionale dei Lincei, Zanichelli,, Bologna, 1956.