



Approximations that matter: virtual particles as carriers of interactions

Nicolò Cangioti¹ · Gianni Arioli¹ · Giovanni Valente¹

Received: 6 March 2024 / Accepted: 13 December 2024
© The Author(s) 2025

Abstract

In this paper we aim to develop an indispensability argument in support of the existence of virtual particles in scattering processes. In order to avoid the Paradox of Infinite Limits, which allegedly poses a challenge to scientific realism, one needs to de-idealize the fictitious systems introduced by the two limiting procedures employed in the perturbation scheme, namely the infinite expansion in Dyson series and the limits for negative and positive infinite times associated with the assumption of free particles. We show that these limits do not introduce essential idealizations, in agreement with scientific realism. What is more, according to our argument, unobservable virtual particles arise as essential approximations and they should be interpreted as propagators of the interaction responsible for subatomic scattering. As such, their existence is based on the use of approximations that matter.

Keywords Virtual particles · Infinite limits · Approximations · Idealizations · Feynman diagrams

1 Introduction

Virtual particles are an essential element in the description of the sub-atomic world within perturbative quantum field theory. In scattering phenomena, they appear as internal lines of Feynman diagrams, wherein the external lines are meant to depict the (allegedly) real

✉ Nicolò Cangioti
nicolo.cangioti@polimi.it

Gianni Arioli
gianni.arioli@polimi.it

Giovanni Valente
giovanni.valente@polimi.it

¹ Department of Mathematics, Politecnico di Milano, piazza Leonardo da Vinci 32, 20133 Milan, Italy

particles undergoing mutual interactions. Far from having a mere diagrammatic function, virtual particles play a fundamental role in physical calculations, since their contributions enable one to recover experimental data with arbitrary precision. Yet, it is also true that they are never directly observed. Moreover, while they conserve energy, they do not satisfy the relativistic energy-momentum relation, thereby failing to be on-shell. Hence, there arises the philosophical question whether or not virtual particles exist and, depending on the answer to this, what are they supposed to be exactly. Decades after Richard Feynman's original work (Feynman, 1949), the issue concerning the ontological nature of virtual particles is still open. The literature is indeed divided between two main camps, namely the realist and the anti-realist: according to the latter (e.g. Fox (2008), Redhead (1980), Redhead (1988), Weingard (1982), Weingard (1988)), virtual particles do not exist, and in fact they are nothing more than useful computational tools; instead, according to the former (e.g. Jaeger (2019), Valente (2011)), virtual particles do exist, either as actual particles or as mediators of the interaction between the particles undergoing scattering. This debate is closely tied to the issue whether, and in what sense, Feynman diagrams have representational power. Again, on this matter one can identify two opposite positions: for some authors (e.g. Brown (1996), Dorato and Rossanese (2018), Passon (2019), Redhead (1980), Redhead (1988)), Feynman diagrams are merely book-keeping devices, which are just conveniently employed for calculations; for others (e.g. Meynell (2008), Stöltzner (2017), Stöltzner (2018), Valente (2011)), they should be taken as adequate representations of sub-atomic processes. In the present paper, we wish to discuss these issues in the context of another lively debate that is currently taking place in the philosophical community, regarding the alleged distinction between idealizations and approximations (e.g. Batterman (2004), Butterfield (2013), Norton (2012), Palacios and Valente (2021), Ruiz de Olano et al. (2022)), which is germane to the use of infinite limits in physics.

We offer an analysis of sub-atomic scattering processes as being described by Feynman diagrams, wherein we distinguish two kinds of limiting procedures. For one, when applying perturbative methods in the interaction picture one expands the time evolution operator into an asymptotic Dyson series, whose individual terms correspond to single Feynman diagrams of increasing order: as this series is typically divergent, it can only provide a usable approximation when it is truncated. In addition, when constructing the so-called S -matrix, which yields the transition probabilities to be compared with the experimental data, one introduces the unrealistic assumption of free particles, whereby the incoming and the outgoing states are defined in the limit of minus and plus infinite time, respectively. In our view, a closer scrutiny of both limiting procedures allows us to elucidate the physical significance of virtual particles. On the one hand, truncating the Dyson series enacts an approximation that is constitutive of the idealization describing the target scattering process, which corresponds to a finite set of Feynman diagrams. As we claim, this can be taken to exemplify the criterion of content determination proposed by Ruiz de Olano et al. (2022). On the other hand, the infinite-time limit is indispensable to treat asymptotic states as free, thereby enabling one to apply perturbative methods. In principle, that is justified by the intuitive idea that the particles undergoing the scattering do not interact in their remote past and remote future. However, the resulting idealization, namely a limit system of incoming and outgoing free particles, cannot in itself adequately represent

the real scattering processes: indeed, in reality the latter takes place due to the underlying interaction, which is instead encoded in the intermediate states associated with virtual particles. We submit that characterizing virtual particles in terms of the notion of *essential approximations*, a new category which we introduce here, paves the way for their physical interpretation as carriers of the interaction of the underlying fields.

The structure of the paper reflects the dialectics of our argument. We begin in Sect. 2 by recalling the account of scattering phenomena within the framework of perturbative quantum field theory, together with the construction of Feynman diagrams. Then, in Sect. 2, the limiting procedures involved therein are framed in the context of the philosophical debate between idealizations and approximations (Sect. 3). In particular, we explain how a Paradox of Infinite Limits, which poses a challenge to scientific realism, can arise if the fictitious limit systems constitute essential idealizations, and we proceed to contrast the latter with the newly defined notion of essential approximations. In the following section, we discuss some critical objections to the existence of virtual particles: as we argue, the most powerful arguments just show that virtual particles cannot be interpreted in terms of a particle ontology. Finally, in Sect. 5, we develop a refined indispensability argument in support of the reality of unobservable virtual particles. The upshot is that, understood in terms of essential approximations, the perturbative scheme for sub-atomic scattering does not run against the Paradox of Infinite Limits, in compliance with scientific realism. Moreover, one can give a physical interpretation of virtual particles as propagators of the interaction responsible for real sub-atomic scattering, without being committed to a particle ontology. That is the sense in which, as the title of our paper suggests, virtual particles arise as approximations that matters.

2 Subatomic scattering between idealizations and approximations

Quantum field theory is based on a physical approach that combines two of the most successful theories of the 20th century, namely quantum mechanics and special relativity, with the classical theory of fields. From an historical perspective, the first bricks of this theory were laid in the 1920 s, thanks to the work of important physicists, such as Erwin Schrödinger and Paul Dirac, that already contributed to the development of non-relativistic quantum mechanics. In fact, the need for a second quantization manifests itself immediately when we try to describe the behaviour of relativistic particles by means of the usual tools of non-relativistic quantum mechanics. In particular, if one tries to apply the standard treatment to a single-particle relativistic wave equation, such as the Klein–Gordon equation, there arise negative-energy states and other puzzling issues, like the violation of causality. However, already the earlier attempts to formulate quantum field theory encountered computational difficulties due to the perturbative methods used for solving the interaction model. Yet, at the same time, the practical schemes to account for sub-atomic phenomena, such as scattering processes, showed surprisingly good agreement with the experimental data, thereby prompting further developments of the theory. In particular, the pictorial diagrams introduced in 1948 by Richard Feynman, which were subsequently tied to his own name, purported to offer a useful representation of the mutual interactions between sub-atomic particles underlying such phenomena. The goal of theoretical physicists

thus became to settle the foundations of perturbative quantum field theory on a rigorous basis.

The basic principle of the formal techniques to handle interactions is to consider the Hamiltonian describing the theory as the sum of two parts, an interaction free part, for which we know how to derive solutions of the dynamical equations of relativistic quantum fields, and a (small) interaction term, which is less amenable to treatment. The problem can be rigorously addressed with perturbation methods, by employing two distinct limiting procedures. To begin with, one expands the time-evolution operator in the Dyson series. As the infinite series tends to diverge, it is truncated to finite orders for the sake of producing calculations that match the experimental data. To achieve this, the terms in the abstract Dyson series must be connected to physically meaningful terms. However, since we do not have full empirical access to the quantum scale, we cannot provide a fine-grained description of sub-atomic matter. In particular, scattering processes involve two (or more) particles in mutual interaction, but we do not know exactly the details of how their collisions occur. As it is sometimes tellingly put, there is a “bubble of ignorance” surrounding the interaction region. All we can do is to treat scattering as a sort of black-box, and perform calculations that relate to the quantities being measured in high energy experiments, like cross sections or decay rates. For such computational purposes, the perturbative scheme resorts to the so-called *S-matrix*, whereby one determines the probability amplitudes of the transition from the pre-scattering multi-particle state to the post-scattering multi-particle state. At this step, though, a controversial assumption is made: for, given that one can find rigorous solutions just of the unperturbed Hamiltonian, one ought to suppose that particles be free before and after their collision. This assumption in turn presupposes another limiting procedure, namely that one takes the limits for infinite negative and positive times, respectively. Yet, subatomic scattering is due to the underlying interaction between particles, which takes place within extremely short intervals of time. As such, the assumption of free particles can hardly give insight onto the dynamics of the process. The above-described perturbation scheme thus rests upon the idealizations introduced by the expansion in Dyson series as well as by the infinite-time limits. Nevertheless, some additional ingredient is needed to provide an explanation of the target process, namely subatomic scattering. Indeed, in order to make predictions in agreement with the experimental data, one still needs to take into account what happens inside the unknown interaction region.

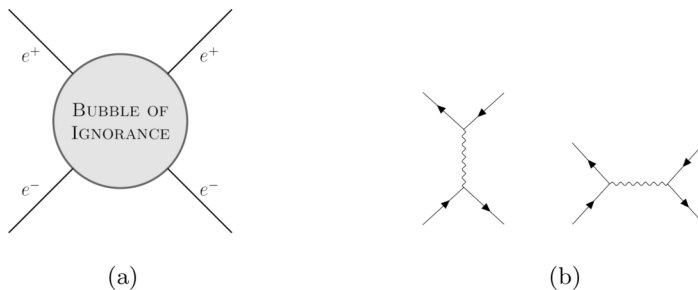


Fig. 1 **a** Target process representing the scattering between electron and positron. **b** Idealization provided by the two second-order Feynman diagrams

The perturbation scheme for sub-atomic scattering can be depicted by means of Feynman diagrams, which are graphical devices meant to summarize in a simple visual form the contributions to the scattering amplitudes calculated with the S-matrix. As such, these diagrams have proven to be a rather useful computational tool in perturbative quantum field theory. They are connected to the underlying physical processes by a set of rules, sometimes referred to the Feynman–Dyson rules, whose precise formulation tends to vary from author to author. In particular, through the application of Wick’s theorem (to which we will return in greater detail in Sect. 3), Feynman diagrams generate the so-called *virtual particles*, which appear as the internal lines connected at the vertices with the external lines representing real particles undergoing scattering. They are therefore posited inside the interaction region. As an illustrative example, we choose the Bhabha scattering between a pair of electron-positron. That is a phenomenon first studied by Bhabha (1935), which involves scattering between an electron e^- and a positron e^+ , with the production of some virtual photons γ . The Feynman diagrams of second order are displayed in the figure here above. Empirical adequacy requires calculations of scattering amplitudes to take into account the whole set of generated Feynman diagrams, up to a given order. Such diagrams incorporate the idealizations introduced in the perturbative scheme in that they correspond to different terms of the Dyson series, and each of them presupposes the assumption of free particles holding in the infinite past and infinite future for the pre-scattering and post-scattering states, respectively. Virtual particles are unobservable entities occupying the intermediate states, which are added to match the experimental data.

What is the exact physical meaning of virtual particles, though, remains a controversial matter in philosophy of quantum field theory. On the one hand, the label "virtual" seems to suggest that they are not real entities, which in fact are never directly observed; but, on the other one, they must be posited in order to make empirically adequate predictions for scattering amplitudes, and as such they play an indispensable role in the explanation of subatomic phenomena. That poses the ontological question whether, and in what sense, virtual particles are real or not. More to the point, over and above the other well-known objections that have been leveled against their existence, since virtual particles arise within the perturbative scheme after one takes the two limiting procedures employed therein, namely the infinite expansion in Dyson series and the infinite-time limits for the pre- and post-scattering states, one may contend that they constitute mere mathematical idealizations corresponding to fictitious, rather than physically real, systems. To address this critical contention, here below we review in greater details the two infinite limits and evaluate whether they are indispensable for the description of sub-atomic scattering.

2.1 The infinite expansion in Dyson series

In the interaction picture introduced above, the time-evolution from the initial instant t_0 can be directly defined as

$$\psi(t) = U(t, t_0)\psi(t_0),$$

where the Dyson operator $U(t, t_0)$ satisfies the following equation

$$i \frac{\partial}{\partial t} U(t, t_0) = V(t)U(t, t_0) \quad (1)$$

The latter can in turn be written as the Volterra equation

$$U(t, t_0) = 1 - i \int_{t_0}^t dt_1 V(t_1)U(t_1, t_0). \quad (2)$$

Solving Eq. (2) by performing a Picard iteration yields the following Neumann series

$$U(t, t_0) = 1 - i \int_{t_0}^t dt_1 V(t_1) + (-i)^2 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 V(t_1)V(t_2) + \dots \quad (3)$$

Given that the fields are time-ordered, namely $t_1 > t_2 > \dots > t_n$, we can write the n -th term of the series by means of the familiar time-ordering operator T :

$$\begin{aligned} U_n(t, t_0) &= \int_{t_0}^t \int_{t_0}^{t_1} \dots \int_{t_0}^{t_{n-1}} dt_1 \dots dt_n V(t_1) \dots V(t_n) \\ &= \frac{1}{n!} \int_0^t dt_1 \dots dt_n T\{V(t_1) \dots V(t_n)\}. \end{aligned} \quad (4)$$

Then, by taking into account all terms we can finally expand the time-evolution operator $U(t, t_0)$ in Eq. (3) into the full *Dyson series*:

$$\begin{aligned} U(t, t_0) &= 1 - i \int_{t_0}^t V(t_1)dt_1 + \frac{(-i)^2}{n^2} \int_{t_0}^t T\{V(t_1)V(t_2)\}dt_1dt_2 + \dots \\ &= \sum_{k=0}^{\infty} U_k(t, t_0) = \lim_{n \rightarrow +\infty} \sum_{k=0}^n U_k(t, t_0). \end{aligned} \quad (5)$$

Clearly, this is the sum of infinitely many terms. Since each term of the expansion is associated with a Feynman diagram, the full Dyson series corresponds to a superposition of diagrams of increasing orders. As a matter of fact, the above series (5) is typically divergent. That raises the question whether or not taking the infinite limit is really indispensable.

Such a question is closely related to the debated issue concerning the representational power of Feynman diagrams, that is whether or not the set of diagrams associated with the terms of the Dyson series should be taken to represent the target subatomic scattering process. As Kaiser (2005) pointed out, the diagrams were originally intended as representing scattering processes observed in bubble chambers, although not in a literal space-time sense: indeed, since only time is a physical parameter in Feynman diagrams, the lines connecting spatially separated vertices cannot coincide with trajectories in Minkowski spacetime. So, the sense in which the diagrams could represent the target ought to be understood in a less pictorial manner. Feynman's own interpretation subsequently shifted towards a mechanistic approach (cfr. Wüt-

hrich (2018) for a historical reconstruction), whereas later authors advocated a more abstract understanding of representation. For instance, Meynell (2008) insists on the denotational function of the diagrams and maintains that they allow us to build an image of the sub-atomic world by playing a kind of make-believe game, along the lines of Walton's theory of representation (Walton, 1990). By borrowing Morgan and Morrison's notion of models as mediators between physics and mathematics (Morgan & Morrison, 1999; Stöltzner, 2017, 2018) even goes on to argue that, in order to vindicate the successful applications of the theory, one must take the infinite set of Feynman diagrams as representative of scattering processes. According to this realist position, the limit $n \rightarrow \infty$ leading to the full Dyson series (5) appears indispensable.

On the negative side, though, other authors advocate a deflationary view. In particular, Redhead (1988) contends that, in analogy with the motion of a violin string, the infinite sum "is just a mathematical expansion with no direct physical significance for the component[s]... To invest them with physical significance is like asking whether the harmonics really exists on the violin string". That is actually the root of the argument for superposition Weingard (1988) formulated to the rejection the existence of virtual particles, which we will address in Sect. 3. In addition, based on the fact that the infinite series (5) diverges, Brown (1996) denies that the set of corresponding Feynman diagrams can adequately represent any subatomic process. What is more, Mario Bacelar Valente points out that, for the pragmatic purpose of recovering empirical results, in practice physicists truncate the full expansion to just a few finite terms. It follows that, in contrast to Stöltzner, according to the anti-realist position towards Feynman diagrams, the infinite Dyson series (5) can, and actually should, be dispensed from the explanation of subatomic scattering. So, although the expansion in Dyson series is part and parcel of the perturbation scheme, whether or not the infinite limit $n \rightarrow \infty$ is indispensable remains a controversial issue.

2.2 The infinite-time limits

The other limiting procedure requires one to take the infinite-time limits $t \rightarrow -\infty$ and $t \rightarrow +\infty$ for the infinite past and the infinite future of the particles undergoing scattering. Let us see how it is introduced in connection with the assumption of free particles. In scattering experiments, the measurable quantities are encoded in the probability amplitudes for cross sections (or decay rates), which are computed by means of the S -matrix as the collective sum of *all* the possible results of the collision. In general, we can denote pre- and post-scattering configurations of a system composed by N particles by means of multiparticle states, labeled 'in' and 'out', respectively, that are specified in terms of their own momentum p : that is,

$$|\alpha\rangle_{\text{in/out}} = |p_1, \dots, p_N\rangle_{\text{in/out}}. \quad (6)$$

These are supposed to be states of the full interacting theory, yet they do not lend themselves to a direct treatment in the computation of scattering amplitudes. In fact, in perturbative quantum field theory, they must be connected to the states of the free theory. To put it technically, if $|\alpha\rangle$ represents the eigenstate of the free part of the Hamiltonian H_0 , the incoming and outgoing states can be schematically written as

$$|\alpha\rangle_{\text{in}} = \Omega_- |\alpha\rangle \quad \text{and} \quad |\alpha\rangle_{\text{out}} = \Omega_+ |\alpha\rangle. \quad (7)$$

where the Møller wave operators Ω_{\pm} are defined as the infinite-time limits $\lim_{t \rightarrow \pm\infty} \mathcal{U}(t)$ of the unitary operator $\mathcal{U}(t) = e^{iHt} e^{-iH_0 t}$. In this framework, the S -matrix, whose elements yield the probability amplitudes for the transition from the in-state $|\alpha\rangle_{\text{in}}$ onto the out-state $|\beta\rangle_{\text{out}}$ states, is then built as follows:

$$S_{\beta\alpha} = {}_{\text{out}} \langle \beta | \alpha \rangle_{\text{in}} = \langle \beta | \Omega_+^\dagger \Omega_- | \alpha \rangle = \langle \beta | U(+\infty, -\infty) | \alpha \rangle. \quad (8)$$

The term $U(+\infty, -\infty)$ is constructed in the limits of infinitely distant past and future computed for the evolution operator $U(t_0, t)$ expressed by the Dyson series (5) in the Heisenberg picture. As such, the asymptotic free states $|\alpha\rangle$ and $|\beta\rangle$ appearing in the S -matrix (8) correspond to the external lines of Feynman diagrams. Incoming and outgoing particles are thus idealized as being free in the infinite-time limits $t \rightarrow \pm\infty$.

In the attempt to justify the assumption of free particles, in his review of the *Philosophy of Particle Physics* Williams (2023) argues that “this particular idealization has a straightforward unpacking” [p. 12], as particles can be regarded as strictly non-interacting when their spatial separation becomes infinite. Here, though, there arises the worry that the assumption be unrealistic given that scattering experiments take place just over small distances (e.g. the size of the tunnels of particle accelerators is of the order of a few meters). Yet, Williams rebuts that it is sufficient that the distance between the particles be large with respect to the interaction region. In particular, when the strength of the interaction decreases exponentially with the spatial separation, like in the example of the Yukawa potential, the particles may be regarded as free (at least to a very good approximation) already after an extremely short distance. Granted, Williams’ justification seems intuitively reasonable. However, in the construction of the S -matrix (8) the assumption of free particles is formalized in terms of the limits $t \rightarrow \pm\infty$ for negative and positive times. So, for the purported justification to hold, the infinite-time limits should be thought of as assuring infinite distance between the particles. A rigorous connection between time and spatial location is then required.

More to the point, to be associated with sub-atomic particles undergoing scattering within some region of spacetime, the states of the interacting theory must be given spatio-temporal localization. The method for localizing states of the form (6) is to construct them out of single-particle wave-packet states as follows:

$$e^{-iHt} \int d\alpha g(\alpha) |\alpha\rangle_{\text{in/out}} \quad (9)$$

The effect of integrating over the Gaussian function $g(\alpha)$ is precisely to localize the state as a wave-packet. By applying the time-evolution operator, we then get

$$e^{-iHt} \int d\alpha g(\alpha) |\alpha\rangle_{\text{in/out}} = \int d\alpha e^{-iE_\alpha t} g(\alpha) |\alpha\rangle_{\text{in/out}}. \quad (10)$$

The next step is to split the full Hamiltonian H as

$$H = H_0 + V(t),$$

with the requirement that the eigenvalues E_α of free Hamiltonian H_0 are the same as the in/out states: that is,

$$H|\alpha\rangle_{\text{in/out}} = E_\alpha|\alpha\rangle_{\text{in/out}} \quad \text{and} \quad H_0|\alpha\rangle = E_\alpha|\alpha\rangle. \quad (11)$$

Under these conditions, one can finally show that the localized in/out states become asymptotically free, in the sense that in the limit for $t \rightarrow \pm\infty$ one obtains

$$e^{-iHt} \int d\alpha g(\alpha) |\alpha\rangle_{\text{in/out}} \approx e^{-iH_0t} \int d\alpha g(\alpha) |\alpha\rangle. \quad (12)$$

A unique association between free and interacting states is thereby established in the limit.¹ Since it is necessary to prove the convergence required to define the S -matrix (5), taking the infinite-time limit seems to be indispensable to the perturbation scheme for subatomic scattering. But is it really so? And what does it entail for the physical significance of virtual particles?

3 The paradox of infinite limits and the notion of essential approximations

Scientific realism is the doctrine according to which even the unobservable entities postulated by our successful theories should be regarded as real, especially when they play an indispensable role in the explanation of physical phenomena. The empirical indispensability of virtual particles for the S -matrix calculations of sub-atomic scattering thus gives a positive argument for their existence. Nevertheless, as we discuss in this section, the use of infinite limits in physics poses a challenge to scientific realism if such limits prove indispensable for the explanation of empirical phenomena. In fact, when the variable that goes to infinity have physical significance, such limits introduce fictitious infinite systems that cannot, strictly speaking, correspond to real finite systems. If that happens to be the case within the perturbative scheme for sub-atomic scattering processes, then one would face a *prima facie* tension with the indispensability argument in favor of virtual particles. In order to dispel confusion concerning indispensability, we now frame the issue of infinite limits into the context of the current debate over idealizations and approximations, wherein there arises the so-called Paradox of Infinite Limits, which allegedly challenges scientific realism.

Part of the vexed debate about idealizations and approximations in the philosophical literature hinges upon disagreements concerning the ways in which the relevant nomenclature is introduced. However, the issue should not be reduced to a mere mat-

¹Brenig and Haag (1959) even prove that, under precise assumptions, the effect of interaction potentials that falls off sufficiently fast at long distances becomes negligible when $t \rightarrow \pm\infty$.

ter of definitions, since after all one adopts genuine conceptual distinctions to shed light onto foundational problems in physics. For the sake of fixing the meaning of the terms, here we concentrate on the distinction between idealizations and approximations as it arises in connection with the use of infinite limits, which appear to be indispensable to the explanation of physical phenomena. A relatively recent proposal elaborated by Norton (2012) characterizes an idealization as a fictitious system, possibly constructed in the mathematical limit whereby a given parameter n goes to infinity, whose properties are meant to provide inexact descriptions of the properties of the target system, for which n is only finite. Such inexact descriptions are in turn regarded as approximations. The latter may be taken as good as one wishes to an arbitrary degree, depending on how close they are to the quantitative values of the target's properties. Yet, according to Norton, not all limit systems can be elevated to the status of idealizations: for, if the mathematical limit retains contradictory properties, or it fails to share some crucial properties with the target, then it should just be demoted to a mere approximation. In fact, while an idealization ought to provide approximations of the relevant properties of the target, approximations do not need to refer to any idealized system in order to be introduced. Norton illustrates this proposal by means of several examples of infinite limits in statistical mechanics, wherein taking the limit $n \rightarrow \infty$ for the number of particles leads one to construct badly-behaving fictitious systems, which should therefore be de-idealized.

Norton's purported distinction, though, stands in opposition to the proposals advocated by other authors. For one, Redhead (1980) argues that there is equivalence between idealizations and approximations, in that it is always possible to embed inexact descriptions of some properties of the target into a new idealized system, even though the latter may as well not capture all properties of the target (as it can be shown in the example of the harmonic oscillator). In a somewhat different vein, Batterman (2004) observes that it is not always possible to de-idealize the limit system to derive mere approximations. In fact, there are examples of so-called singular limits, wherein the description of a real-world phenomenon cannot be given for finite n , and hence one is bound to take the limit $n \rightarrow \infty$ in order to recover empirically correct results for certain crucial properties (e.g., in classical phase transitions the derivative of the free energy exhibits the required discontinuity only in the thermodynamical limit). According to Batterman, in such cases the infinite system must be granted physical significance: in fact, since the limit is indispensable it constitutes an *essential idealization*, and as such it cannot be de-idealized. This claim, however, has been criticized. For instance, along the lines of Norton's proposal, Butterfield (2013) submits that even in the case of singular limits it must be possible to recover the desired behaviour "on the way to the limit", that is by considering approximations for very large but still finite values of the parameter n , without having to reify the limit system. How this can be done, though, must be evaluated on a case-by-case basis.

Essential idealizations are suspected to pose a threat to scientific realism, in that they give rise to what has been dubbed the Paradox of Infinite Limits (cfr. Palacios and Valente (2021)). For, it seems that, inasmuch as the infinite system is indispensable for the explanation of a physical phenomenon, one would have to believe in its existence even though it is not, strictly speaking, a real system, but rather a fictitious

one. In light of the purpose of the present paper, the paradox can be formulated as the combination of the following three statements.

- (1) *Finiteness of Real Systems*. If a variable n characterizing a real system corresponds to some physical parameter, it cannot take on infinite values.
- (2) *Indispensability of the Limit System*. The explanation of phenomenon P can only be given by means of claims about an infinite system constructed in the limit $n \rightarrow \infty$.
- (3) *Enhanced Indispensability Argument (EIA)*. If a claim plays an indispensable role in the explanation of the phenomenon P we ought to believe in its existence. The first condition is a basic desideratum of scientific realism. It is meant to assure that relevant physical parameters are finite for real target systems. For instance, real gas systems have a very large but finite number n of constituents. Likewise, time t can be a relevant parameter, and so statement (1) assures that real physical processes take place within a finite interval of time. Statement (2) challenges statement (1), for it asserts that a finite system cannot explain a phenomenon of interest, while an infinite system can. Taking the limit thus appears as indispensable to the explanation of the phenomenon. Statement (3) then advances a challenge to scientific realism. In fact, according to Baker's (Baker, 2009) Enhanced Indispensability Argument, we ought to believe in the existence of mathematical entities, such as infinite-limit systems, without which we cannot provide an explanation of physical phenomena. This entails that a scientific realist should be committed to the reality of essential idealizations arising in the indispensable limit $n \rightarrow \infty$. But that is in stark contrast with the finiteness of real system asserted by statement (1).

Granted that a scientific realist cannot naively reject statement (1), there are two alternative solutions to the Paradox of Infinite Limits. Baron (2019) suggests that one should give up statement (3). In his critique of EIA, he distinguishes between constructive indispensability, whereby there is reason to believe that a mathematical entity can be dispensed but at the present stage we do not yet know how, and substantive indispensability, whereby there is no reason to suppose that such an entity can ever be dispensed. To ground such a distinction, Baron proposes the test of coherence, stating that, if a claim is not consistent with other currently accepted theories, it is only constructively indispensable. Then, if it is constructively indispensable, it does not support EIA in statement (3), and hence it does not mandate any ontological commitment. However, whether or not indispensability is only constructive is a merely contingent fact, depending on our current state of knowledge. As such, moreover, Baron's solution does not apply to cases of substantive indispensability, which are still subject to the paradox. The other solution advocated by Palacios and Valente (2021) rests on the possibility of de-idealizing the fictitious systems arising in the infinite-limit, so as to block statement (2). To this extent, one should show that the limit system is not an essential idealization, thereby disarming the alleged challenge to scientific realism. The infinite idealization can be dispensed if one can still give approximations of the properties of the real target system for finite values of the parameter n , e.g. through Butterfield's de-idealizing strategy to recover empirically

correct values "on the way to the limit". In the particular case of singular limits, one should not expect to recover empirically correct values for all properties, but it is sufficient to give good approximations of some physically meaningful properties that are relevant for the phenomenon to be explained.

In the perturbation scheme for subatomic scattering described by Feynman diagrams, a challenge for scientific realism could arise if the infinite expansion in Dyson series or the infinite-time limits introduced with the assumption of free particles are essential idealizations. Then, as virtual particles are internal lines of the relevant set of Feynman diagrams, one would face a potential threat for their existence too. To disarm this threat, one needs to show that the infinite systems introduced by the two limiting procedures in the perturbation scheme can, and in fact should, be de-idealized, so as to block the Paradox of Infinite Limits, along the lines of Palacios and Valente's solution. We will do this in Sect. 5, where we offer a physical interpretation of virtual particles compatible with scientific realism. Before doing so, here we wish to introduce a new category in the context of the philosophical literature on idealizations and approximations, namely the notion of *essential approximations*, as we will argue that virtual particles just fall into such a category.

In a recent article about the distinction between idealizations and approximations, Ruiz de Olano et al. (2022) reaffirm the autonomy of approximate methods from the process of constructing idealization. Based on some historical examples from quantum field theory (e.g. the Nambu-Jona-Lasinio model, wherein scattering amplitudes are extracted by perturbing the system's Hamiltonian), these authors argue that there are good reasons to recognize approximations as being distinct from idealizations, both from a theoretical point of view and in terms of the practice of scientific modeling. In particular, according to their thesis of *Content determination*, in some cases it can happen that approximations provide physical and empirical content to the idealization constructed in the limit, instead of the other way around. This does not only counter Redhead's equivalence claim, but it also reverts the presupposition implicit in Norton's nomenclature that approximations accrue from the idealized limit system to describe the properties of the target system. We submit that the notion of essential approximations constitutes a special kind of category, which complies with the thesis of Content determination. For our purposes, it can be defined as follows:

Essential Approximations: Inexact descriptions of relevant properties of the target system, which prove indispensable to the explanation of some physical phenomenon of interest.

As we have already observed, approximations are compatible with scientific realism because they do not have to be defined in the limit $n \rightarrow \infty$: in fact, it is sufficient to take finite values of the parameter n . What is more, in the case of essential approximations, taking finite values of n is also necessary for the explanation of the physical phenomenon of interest. On this point, let us stress that there are (at least) two senses in which one could think of indispensability. First of all, taking an approximation can be necessary in order to obtain empirically correct values, within some accepted degrees of inexactness, of the relevant properties of the target. That is what one can refer to as empirical indispensability. Another sense in which one could think

of approximations to be indispensable is in terms of recovering the physical content of the target system undergoing the phenomenon of interest. In particular, such a kind of physical indispensability occurs when infinite idealizations are not even able to explain relevant physical properties of the target, whereas finite approximations can and are effectively used to do so.

As we will show in Sect. 5, the perturbative scheme for sub-atomic scattering wherein virtual particles appear as internal lines of Feynman diagrams ought to be understood in terms of essential approximations, in compliance with scientific realism. To make room for the physical interpretation of virtual particles as propagators of the underlying interaction responsible for real scattering processes, the next step in our argument is to review some critical aspects of the definition of virtual particles, as well as of their interpretation in terms of a particle ontology. We take this up in the following section.

4 On the ontology of virtual particles

In order to see how virtual particles arise within the perturbation scheme for subatomic scattering, let us refine the formalism presented thus far in Sect. 2. The fundamental mathematical tool for connecting the abstract Dyson series with the calculations of cross scattering is the n -points correlation function, which yields vacuum expectation values of the time-ordered product of n field operators $\phi(x_i)$, with $i = 1, \dots, n$. That is,

$$\langle \Omega | T \{ \phi(x_1) \cdots \phi(x_n) \} | \Omega \rangle, \quad (13)$$

where T is the time-ordering operator in Eq. (5) for the Dyson series and $|\Omega\rangle$ denotes the ground state of the interacting theory. As required in the construction of the S -matrix, state $|\Omega\rangle$ must converge in the infinite-time limit to the ground state $|0\rangle$ of the free theory with Hamiltonian H_0 . The next step to put the Dyson series in correspondence with a set of Feynman diagrams is to apply Wick's theorem, which enables one to rewrite each time ordered element of the expansion as a sum of normal ordered terms.² In particular, via a rule known as the Feynman prescription (whereby the Green's function must be chosen so that the contour fixed for computing the complex integral goes under the left pole and over the right pole), the propagator for the free field

$$\langle 0 | T \{ \phi(x_1) \phi(x_2) \cdots \phi(x_n) \} | 0 \rangle \quad (14)$$

becomes equal to a n -points Green functions $G(x_1, x_2, \dots, x_n)$, corresponding to the vanishing field vacuum expectation value. To make this general expression more explicit, let us refer to the simple case of two scalar fields $\phi(x)$ and $\phi(y)$, for which the correlation function (13) can be written as

²Recall that a normal order is defined by moving all creation operators to the left of all annihilation operators in the operator product.

$$\langle \Omega | T \{ \phi(x) \phi(y) \} | \Omega \rangle \tag{15}$$

If we consider the Klein–Gordon equation for free fields, under the requirement that the commutator $[\phi(x), \phi(y)]$ vanishes for spacelike separated points x and y , after taking the limit for infinite negative and positive times³, we derive the following Lorentz invariant expression

$$\langle 0 | T \{ \phi(x) \phi(y) | 0 \rangle = i D_F(x - y) = i \int \frac{d^4 p}{(2\pi)^4} \frac{e^{-ip(x-y)}}{(p^2 - m^2)}, \tag{16}$$

for momentum p and mass m . Here, the term $D_F(x - y)$ is the so-called *Feynman propagator*, which, courtesy of Wick’s theorem, denotes the internal lines of Feynman diagrams.

Since propagators are time-dependent probability functions associated with energy and momentum of moving particles, the Feynman propagator is meant to represent virtual particles moving inside the interaction region of subatomic scattering processes. As Peskin and Schroeder (1995) put it in their well-known textbook on quantum field theory,

The Green’s function $D_F(x - y)$ is called the Feynman propagator for a Klein–Gordon particle, since it is, after all, a propagation amplitude. Indeed, the Feynman propagator will turn out to be part of the Feynman rules: $D_F(x - y)$ (or $\tilde{D}_F(p)$) is the expression that we will attach to internal lines of Feynman diagrams, representing the propagation of virtual particles. [p. 31]

Expressed as propagators, virtual particles contribute to the probability amplitudes for scattering events calculated with the S -matrix. More to the point, they add to the rate of collisions in the Feynman diagrams wherein they appear as internal lines. The overall scattering probability is then computed by summing up the contributions of all diagrams in the set corresponding to the expansion in Dyson series. However, the jury is still out concerning the ontological interpretation of virtual particles. The main reason is that they are never directly observed in high-energy experiments. In fact, in real target scattering processes the interaction region is treated as a black-box, and hence it is, so to speak, surrounded by a bubble of ignorance. As such, the presence of virtual particles is only inferred from the fact that their contributions to scattering

³Along the lines of Sect. 2, in the present formalism states $\langle \Omega |$ and $| \Omega \rangle$ are obtained by evolving $\langle 0 |$ and $| 0 \rangle$ for time t and $-t$ respectively, and then taking the infinite-time limit in an imaginary direction, i.e. $t \rightarrow (1 - \varepsilon i)\infty$, so that the correlation function (15) can finally be written as

$$\langle \Omega | T \{ \phi(x) \phi(y) \} | \Omega \rangle = \lim_{t \rightarrow (1 - \varepsilon i)\infty} \frac{\langle 0 | T \left\{ \phi_I(x) \phi_I(y) \exp \left[-i \int_{-t}^t V(\tau) d\tau \right] \right\} | 0 \rangle}{\langle 0 | T \left\{ \exp \left[-i \int_{-t}^t V(\tau) d\tau \right] \right\} | 0 \rangle},$$

where $\phi_I(t, x) = e^{iH_0 \cdot (t-t_0)} \phi(t_0, x) e^{-iH_0 \cdot (t-t_0)}$ is the time-evolved field operator in the interaction picture.

amplitudes are needed in order to recover correct experimental data. But then the question arises whether they are just posited for computational purposes, or they correspond to real physical effects.

A naive explanation for the lack of observability of virtual particles is that, due to Heisenberg uncertainty relation between time and energy, they would come into being for an extremely short lapse of time, thereby disappear well before we can possibly detect them. Again, this interpretation is well illustrated by a quotation from Peskin and Schroeder (1995):

[e]ven when there is not enough energy for pair creation, multiparticle states appear, for example, as intermediate states in second-order perturbation theory. We can think of such states as existing only for a very short time, according to the uncertainty principle $\Delta E \cdot \Delta t = \hbar$. As we go to higher orders in perturbation theory, arbitrarily many such “virtual” particles can be created. [p. 13]

In our view, though, this alleged explanation rests on an untenable reading of Heisenberg’s uncertainty principle. As Busch (2008) pointed out, the time-energy relation is supposed to hold between pairs of quantum operators, yet the time measured by clocks, namely the “external time” that measures a duration, is just a parameter of the experiments (cfr. also Griffiths (1987) for a similar objection). Surely, as Roberts and Butterfield (2020) showed, one can formulate an appropriate operator relation for the “internal time” (e.g. the Mandelstam-Tamm uncertainty principle), but even so it seems unreasonable to maintain that, on the basis of such a relation, energy experiments are responsible for the creation of virtual particles. What is more, it should be noted that an energy-time uncertainty principle can be written for non-relativistic quantum mechanics, where there is no room for virtual particles. So, it can hardly be the reason why the latter should exist for too little time to be detected. In addition to these negative remarks, there are further critiques to the view that virtual particles are actually real particles. Let us proceed to review the two main anti-realist objections, namely the alleged violation of conservation of energy and the argument for superposition, so as to explain what they show exactly about the ontology of virtual particles.

Some authors have argued that virtual particles do not exist because they do not satisfy the relativistic energy-momentum relation. In fact, they are often said to be off mass-shell, contrary to real particles that are supposed to be on mass-shell. In particular, the particles undergoing scattering that are associated with the free initial and final states in the S -matrix are on mass-shell. An instance of this argument is discussed in another renowned textbook by Griffiths (1987) as follows:

although energy and momentum must be conserved at each vertex, a virtual particle does not carry the same mass as the corresponding free particle. In fact, a virtual particle can have any mass – whatever the conservation laws require. In the business, we say that virtual particles do not lie on their mass shell. External lines, by contrast, represent real particles, and these do carry the “correct” mass. [p. 60]

However, we submit that the claim that virtual particles violate conservation of energy is misleading on different grounds. For one, after the above quotation Griffiths himself goes on to observe that, if one takes into account the overall process, energy is always conserved anyway. Furthermore, it should be added that it is not necessarily true that mass does not take on the expected value along the internal lines of Feynman diagrams: as a matter of fact, virtual particles may as well be on mass-shell. Finally, Jaeger (2019) also objects that the alleged violation of the energy-momentum relation is based on a classical understanding of physical processes, whereby the requirement of conservation of energy comes to the detriment of Lorentz covariance. As such, it should not be enforced in a relativistic version of quantum mechanics, like perturbative quantum field theory, which is Lorentz covariant. In light of these critical facts, we do not find the anti-realist objection that virtual particles are off mass-shell really compelling.

Another claim leveled against the reality of virtual particles, which Fox (2008) deems as the most forceful anti-realist objection, stems from Weingard's (1982) argument for superposition, which we already mentioned in the previous section in connection with Redhead's (1988) criticism of the representational power of Feynman diagrams. It is based on the fact that there is an infinite superposition of Feynman diagrams in correspondence with the terms of the full Dyson series: then, since different virtual particles arise in each diagram depending of its specific order, it follows that "neither the number nor kinds of virtual particles are sharp" (Weingard 1988). Accordingly, virtual particles would lack properties that are typically ascribed to real particles, hence casting doubts onto their existence. Nevertheless, we resist this anti-realist conclusion of the argument for superposition. To begin with, let us stress that, in addition to the number and type of virtual particles, even the final state of the S -matrix remains indefinite until a scattering measurement is made. As Jaeger (2019) argues, in response to Weingard, this fact is similar to the more familiar scenario in ordinary non-relativistic quantum mechanics, wherein before a measurement is performed the state of the system can be in a superposition of the eigenstates of the measured observable, like in the famous double-slit experiment. In that case, although quantum superposition may have other puzzling consequences, it does not call into question the existence of the physical system that is in such a state. So, likewise, it should not undermine the reality of virtual particles either. As Jaeger puts it,

when a measurement for a state containing a particle of a given type is made, an ostensibly virtual particle may be found and, then, uncontroversially considered real and present to anyone accepting the existence of particles at all. [p. 12]

In this respect, virtual particles are to be regarded as real as the particles undergoing scattering, to which the initial and final states in the S -matrix are associated.

In our view, the problem with the objection raised by the argument for superposition is that it presupposes a particle interpretation. In fact, the requirement that the number and type of virtual particles be determinate would apply just if they are to be interpreted as being particles. Nevertheless, there is no consensus on the ontology of quantum field theory, that is on the nature of the subatomic systems involved in scattering phenomena. Other interpretations of the theory are available that can

better accommodate virtual particles. For instance, as a valuable alternative to a particle ontology, subatomic systems can be interpreted as fields, which have rather distinct properties from particles. In particular, fields are not required to be countable, contrary to particles that are instead supposed to be in a determinate number. As a consequence, Weingard's objection loses its force, for it does not show that virtual particles are not real, but only that they cannot be particles. Actually, some of the authors advocating a realist position towards virtual particles, such as Jaeger (2019) and Valente (2011), endorse a field ontology of quantum field theory, whereby not even the subatomic systems that are observed to undergo scattering are particles, but rather they are excitations of the underlying field. According to this interpretation, virtual particles are the carriers of the interaction bringing about the scattering, which is encoded in the contributions of the Feynman propagators to the S -matrix that are required to recover empirically correct predictions. In line with this interpretation, in the following section we aim to refine the indispensability argument in favor of the reality of virtual particles.

5 Virtual particles are as real as interactions

The most appealing positive argument for the existence of virtual particles, at the bottom of the realist position of authors like Jaeger (2019) and Valente (2011) among others, is based on a variant of the so-called no miracle argument in support of scientific realism (see Boyd (1983)). Crudely put, the argument goes that one ought to believe that even abstract claims about the entities postulated by our most successful physical theories are true (or at least approximately true), otherwise their success would be a miracle, with no explanation whatsoever. Although virtual particles are not directly observed, they are postulated by perturbative quantum field theory to make successful empirical predictions about scattering phenomena. As such, according to the no miracle argument, one ought to believe that they correspond to true physical effects. In line with scientific realism. Here, the inference to the existence of unobservable virtual particles is enforced by the essential role they play in the perturbation scheme: for, the transition probabilities computed with the S -matrix are in agreement with the experimental data only when we include the contributions of the intermediate states in the internal lines of Feynman diagrams. Virtual particles are indeed indispensable to the empirical success of the theory. From the perspective of scientific realism, that gives a compelling reason to believe that they are real.

We think this indispensability argument goes in the right direction. However, as it stands, it faces two shortcomings. First, the argument only establishes the reality of virtual particles, but it does not say exactly what kind of entities they are supposed to be. The more so because, owing to the argument for superposition we discussed in the previous section, they cannot be interpreted as particles in the traditional sense. Hence, there remains open the ontological question what physical effects in the real world virtual particles correspond to. Second, the perturbation scheme for subatomic scattering wherein virtual particles arise as internal lines of Feynman diagrams employs two distinct limiting procedures, namely the infinite-time limits for the initial and final states of the S -matrix and the infinite expansion in Dyson series

for the time-evolution operator $U(t, t_0)$. As we saw in Sect. 3, though, the Paradox of Infinite Limits arises when the fictitious limit system is indispensable to the explanation of some physical phenomenon. So, insofar the infinite idealizations constructed in the perturbation scheme to explain subatomic scattering are essential, one would face a conflict with scientific realism. In that case, somewhat surprisingly, indispensability, rather than licensing the purported inference to the reality of virtual particles, would instead block it. There thus seems to be a *prima facie* tension between the indispensability argument for the reality of virtual particles and the indispensability of the infinite limits employed in the perturbation scheme. To provide a consistent physical interpretation of virtual particles without committing one to any particle ontology, here we propose a refined version of the indispensability argument for their existence, which rests upon the notion of essential approximations, and therefore avoids the Paradox of Infinite Limits: accordingly, virtual particles correspond to real effects of the underlying interaction responsible for sub-atomic scattering, rather than to unobservable sub-atomic particles with extremely short life span.

The solution of the Paradox of Infinite Limits advocated by Palacios and Valente (2021) requires one to de-idealize the fictitious limit system, so as to give up statement (2). In other words, one needs to show that the limit $n \rightarrow \infty$ does not introduce an essential idealization, in that an explanation of the phenomenon of interest can be given in terms of approximations of the relevant properties of the target system, which prove to be good enough already for finite values of the variable n . One thereby dispenses the infinite limit from the purported explanation, in compliance with scientific realism. In the case of subatomic scattering, we argue that the infinite Dyson series as well as the infinite-time limits can, and indeed should, be dispensed. More to the point, although the interaction region is surrounded by a bubble of ignorance, and as such it is treated as a black-box, for the perturbation scheme to provide an explanation of subatomic scattering it must fulfill (at least) two conditions. First, it needs to recover correct empirical values, in agreement with the data gathered in high-energy experiments. Second, it ought to account for how the underlying interaction propagates throughout the process: else, there would lack reference to the physical effect responsible for scattering itself to take place. As we show next, the idealizations introduced by the two limiting procedures employed in the perturbation scheme, though, fail to fulfill these conditions, and as such they should be dispensed. That is just the point where virtual particles become indispensable to explain sub-atomic scattering within the perturbative scheme described by Feynman diagrams.

Let us first discuss the idealization arising in the limits for infinite negative and positive times. It guarantees that in the infinite distant past and future the particles undergoing scattering are free. Allegedly, these infinite-time limits $t \rightarrow \pm\infty$ are indispensable for the construction of the S -matrix, which requires the initial and final states of the S -matrix be eigenstates of the free Hamiltonian. Yet, if we take the systems constructed in the infinite-time limits $t \rightarrow \pm\infty$ as essential idealizations, statement (2) of the Paradox of Infinite Limits would be instantiated, in contrast with the fact that subatomic scattering processes take place during finite, extremely short intervals of time t , as required by statement (1). This would pose a challenge to scientific realism along the lines of the Paradox of Infinite Limits. Nevertheless, we contend that there are reasons why one should de-idealize the infinite-time limit systems. To begin with, for Williams (2023) the assumption of free particles underlying

the infinite-time limits has a straightforward spatial justification, in that in the distant past and future the particles are supposed to be so far apart that their interaction vanishes. To ground this justification, though, states must be given spatio-temporal localization. However, with the method for localizing states described at the end of Sect. 2.1, convergence (12) of the interacting states to the free asymptotic states can only be intended in a weak sense. For, as long as particles are described in terms of wave-packets having finite width at finite time, when one takes the infinite-time limit it is typically assumed that they spread over the entire space, so that the probability of being localized at each point vanishes. In fact, in the construction of the S -matrix, one does not consider a packet of plane waves with momentum centered at some p , but rather a single planar wave with a fixed momentum p . Accordingly, while it is still possible to take the limit for the momenta envelope to converge to a Dirac delta, such a limit is just distributional. Hence, the sought-after convergence is guaranteed in a very weak sense. In this respect, we agree with Duncan (2012) that the limit $t \rightarrow \pm\infty$ should not be taken literally, but rather as a useful mathematical fiction of infinite past and future (p. 98). Indeed, it seems too strong a demand to elevate infinite systems defined in a distributional limit to the status of idealizations.

In addition, the infinite-time limits only tell us that the particles are free when they are separated by very large spatial distances. Yet, that does not provide any information about how particles behave as they approach each other, the more so because their scattering takes place within a very short interval of time. In particular, there completely lacks reference to their mutual interaction, which is responsible for the real process to take place. What is more, the Møller wave operators Ω_{\pm} we used in (8) to construct the limit time-operator $U(-\infty, +\infty)$, are provably not unitary: such a fact, if interpreted in light of Haag's theorem,⁴ provides a no-go result for the possibility of merging the free theory with the interaction picture. The consequences of Haag's theorem are actually what led Fraser (2020) to conclude that the entire framework of perturbative quantum field theory cannot even support an idealization of real scattering processes, but it ought to be demoted just to a scheme to produce approximations in Norton's sense. All in all, the idealization introduced by the infinite-time limits does not correspond to the real target system, that is a system of two (or more) quantum fields undergoing subatomic scattering due to their mutual interaction. In fact, the interaction is left out of the picture, and hence the infinite idealization fails to provide a physical explanation of the phenomenon.

Regarding the infinite expansion in Dyson series, that is the other limiting procedure employed in the perturbation scheme, recall from Sect. 2.1 that there is an ongoing debate about the representational power of Feynman diagrams, which is related to the issue whether or not the infinite expansion is indispensable. We claim that it is not. Indeed, taken in the limit $n \rightarrow \infty$, the sum (5) becomes an asymptotically divergent series. As such, it cannot recover empirical correct results. Instead, truncating the Dyson series to lower orders, i.e. to some finite n , yields values that

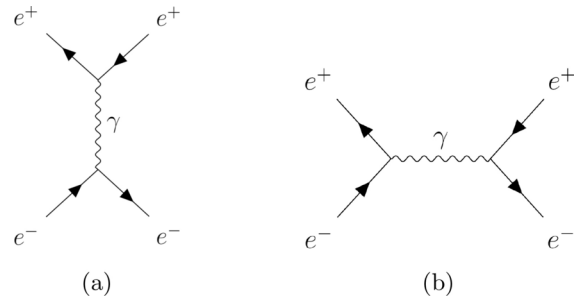
⁴We cannot go into the details of Haag's theorem here, the more so because there are several versions of it (cfr. Earman (2006) and Fraser (2009) for a philosophical discussion). However, we can summarize its importance with the following two statements. (i) If two pure ground states are not equal, then they generate unitarily inequivalent irreducible representations. (ii) If two local quantum fields are unitarily equivalent at any given time, then both fields are free if one of them is free.

are very close to the actual experimental data. On this matter, we concur with Mario Bacelar Valente's pragmatic account of Feynman diagrams Valente (2011), according to which

[w]e do not have an infinite expansion..., what we have are applications of the theory resting in an approximate scheme of description of the interaction between two distinct fields that cannot be taken beyond a few order calculations" [p. 45].

As it turns out, for computational purposes one must apply renormalization methods. Yet, one cannot renormalize the whole infinite series: rather, the renormalization procedure can only be performed after truncating the series at a chosen order. Reducing (5) to lower terms is not just a pragmatic way of coping with our mathematical inability to sum infinitely many terms, or to provide an analytic expression of the full sum; the truncation is even necessary to extract the value of the physical quantity under study, in the face of the divergence of the series. For the sake of recovering empirically correct results, in practice physicists consider just a few terms of the expansion, and sometimes even the first one may actually be sufficient. In the example of Bhabha scattering introduced in the previous section, the real sub-atomic process is approximately represented by the combination of two second-order interactions depicted Fig. 2, comprising the Feynman diagram for a scattering contribution (a) and another Feynman diagram for an annihilation contribution (b). So, to make empirical predictions, one does not take the whole infinite collection of Feynman diagrams, but just the contributions of these two diagrams, together with the virtual particles arising as internal lines in each of them.⁵ The full Dyson series must thus be dispensed with,

Fig. 2 Second-order contribution of the Bhabha scattering. The two interactions represented here are a scattering process (a) and an annihilation process (b). The timeline is intended from left to right



⁵ To put it technically, the application of the Feynman rules leads one to the following expression for the scattering amplitude:

$$\mathcal{A} = -e^2 (\bar{v}_{p1} \gamma^\mu v_{p3}) \frac{1}{(p_1 - p_3)^2} (\bar{u}_{p4} \gamma_\mu u_{p2}) + e^2 (\bar{v}_{p1} \gamma^\nu u_{p2}) \frac{1}{(p_1 + p_2)^2} (\bar{u}_{p4} \gamma_\nu v_{p3}) \quad (17)$$

where the first addendum comes from the scattering and the second one from the annihilation, while we set p_1 and p_3 to represent the four-momentum of the positron and p_2 and p_4 the four-momentum of the electron. In particular here we are using the standard notation, that is γ^μ are the Gamma matrices, u and \bar{u} are the four-component spinors for fermions, and v and \bar{v} are the four-component spinors for anti-fermions.

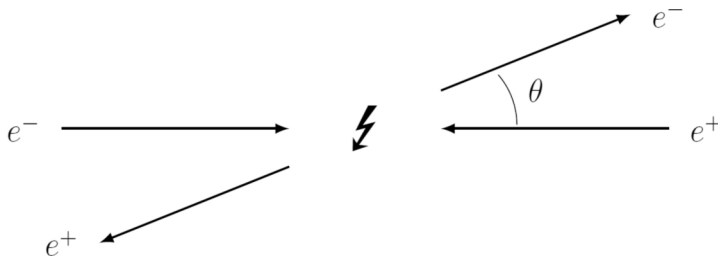


Fig. 3 The electron-positron scattering interaction in the center of mass frame of reference

in compliance with scientific realism. Indeed, for computational purposes, it is the approximation for finite n , instead of the infinite limit $n \rightarrow \infty$ for the series, that proves indispensable.

An additional point is worth making here concerning how to truncate the series. Since we only need to consider finitely many terms, one may wonder how we could actually choose the order of truncation. In Valente’s pragmatic account of Feynman diagrams, it may seem that it is entirely left to an arbitrary decision of the physicists. To the contrary, we submit that the order of truncation has an empirical underpinning. For, it is sensible to fix the finite order in such a way that the accuracy of the computation is compatible with the measurement error. To clarify what we mean by this we recall that, since a Dyson series is divergent, and not even Borel summable, it cannot return values as accurate as desired. Even worse, it is not possible to consider a Dyson series as a representation of a specific function. On the other hand, it is also true that each Dyson series has an optimal truncation order, and the accuracy of a truncated series is taken to correspond to the magnitude of the last term (even if, strictly speaking, this is not a rigorous mathematical statement). As discussed in Miller (2023), it is possible to adapt the state-space semantics to this situation, by modifying the so-called satisfaction function so that, if r is the value obtained by the series with an estimated theoretical error ϵ_t and \bar{r} is the experimental value with a measurement error ϵ_m , the satisfaction function is true when $(r - \epsilon_t, r + \epsilon_t) \subset (\bar{r} - \epsilon_m, \bar{r} + \epsilon_m)$. Accordingly, we can say that the accuracy of the computation is compatible to the measurement error. Again, we can illustrate the point with the Bhabha scattering. By applying the truncation to the second order, we can make precise calculations about, e.g., the cross section σ : specifically, if we fix the usual reference system depicted in Fig. 3, after a few mathematical steps (involving, among other things, the usage of the Mandelstam variables and the relativistic energy-momentum equation) we obtain the differential cross-section

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{8E^2} \left(\frac{1 + \cos^2 \theta}{2} + \frac{1 + \cos^4 \frac{\theta}{2}}{\sin^4 \frac{\theta}{2}} - \frac{2 \cos^4 \frac{\theta}{2}}{\sin^2 \frac{\theta}{2}} \right). \tag{18}$$

where the differential angular range of the scattered particle at angle θ is given by the solid angle $d\Omega = \sin \theta d\theta d\phi$. Already in the 1970 s, it was shown that such a theoretical prediction was in line with the empirical data measured in the Stanford Linear Accelerator Center positron-electron storage ring (SPEAR I), as one can see,

e.g., in the work published by Augustin et al. in 1975 (see Fig. 3 in Augustin (1975)). This indicates that, over and above being a pragmatic matter, the choice of the order of truncation is enforced by experimental constraints. That is the sense in which the approximation for finite n is essential to explain the phenomenon of subatomic scattering.

We claim that dispensing the idealizations introduced by the two limiting procedure employed in the perturbative scheme provides the ground for the indispensability argument for the reality of virtual particles. For one, the problem with the assumption of free particles introduced in the infinite-time limits is that, as it stands, it fails to fulfill a necessary condition for the explanation of subatomic scattering, since the pre- and post-collision states in the S -matrix are taken to be non-interacting. Virtual particles, represented by Feynman propagators (16), must therefore be added to represent the physical interaction propagating through the quantum fields. Besides such kind of physical indispensability, virtual particles satisfy also the demand of empirical indispensability we stated after the definition of essential approximations in Sect. 3. Indeed, as we showed above, it is necessary to truncate the infinite Dyson series in order to recover empirically correct values of the relevant properties of the target. As such, we submit that the perturbative scheme for subatomic scattering must be understood in terms of essential approximations, in compliance with scientific realism. Note that the proposed realist interpretation of virtual particles as propagators of the underlying interaction is not attached to any particle ontology of quantum field theory. For, if they were to be interpreted as particles in the traditional sense, then one would run against the argument for superposition, whereby one cannot fix the number or kind of particles supposedly hidden within the scattering region. Instead, under the present interpretation, the fact that one takes into account a superposition of different Feynman diagrams up to a finite order simply means that one must sum up all the contributions of the internal lines appearing in each diagram in the truncated set for the sake of approximating the real effects produced by the physical interaction between quantum fields that gives origin to subatomic scattering. The number of contributions one chooses to include in the interaction, that is the number of diagrams in the superposition, depends on how well one wishes to approximate the experimental data, without making reference to any unobservable particles, whether off-shell or not. That is the sense in which virtual particles, intended as carriers of the interaction, arise as approximations that matter.

Before concluding, in order to provide further strength to our claim that virtual particles are indispensable and thereby correspond to real effects in the physical world, it is insightful to submit the perturbative scheme for subatomic scattering to a de-idealizing procedure that in the philosophical literature has become known as Earman's Sound Principle (cfr. Jones (2006)). The principle is stated by Earman (2004) in the following terms:

While idealizations are useful and, perhaps, even essential to progress in physics, a sound principle of interpretation would seem to be that no effect can be counted as a genuine physical effect if it disappears when the idealizations are removed. [p. 191]

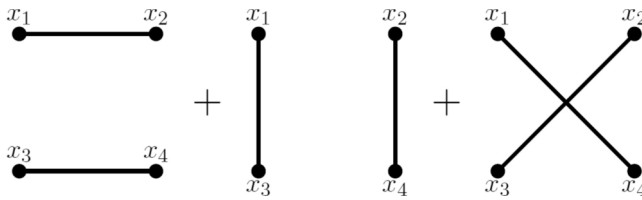


Fig. 4 Feynman diagram representation of the 4-point correlation function $\langle 0|T\{\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\}|0\rangle$

The underlying idea is that, if a certain phenomenon is explained on the basis of, say, an idealized infinite-limit system, a physically real effect must still remain even if one de-idealizes such a fictitious system. In the case of subatomic scattering processes, we argue that one can remove the external lines of Feynman diagrams, namely the asymptotic states of the idealized free particles, in such a way that one is left just with the internal lines, corresponding to the virtual particles, as residuals. To show this, let us stress that under some technical hypotheses it is possible to compute the *S*-matrix terms by means of the Lehmann–Symanzik–Zimmermann (LSZ) reduction formula. As we noted above, one of the troublesome aspects of taking the limit for infinite negative and positive times comes from Haag’s theorem, in that the Møller wave operators Ω_{\pm} employed to construct the operator $U(-\infty, +\infty)$ in Eq. (8) fail to be unitary. That poses deep difficulties on how to represent interactions. The LSZ formalism allows one to circumvent the trouble within the axiomatic version of quantum field theory: in fact, it operates in an integral form over the whole spacetime manifold, in such a way to avoid the limit for infinitely large times. Specifically, the LSZ formula prescribes a recursive procedure that under the action of a particular annihilation operator progressively removes particles in the initial and final states of scattering transitions, until no particle is left at all. At the end of the procedure, what remains is just the vacuum expectation value of the chronological product of field operators, or in other terms the internal lines of the Feynman diagrams.

For the sake of illustrating how the LSZ reduction works, let us discuss formula (13) for a 4-points function. In this case, by extending the treatment of the Feynman propagator (16) for a pair of fields, the propagator (14) becomes:

$$\begin{aligned} \langle 0|T\{\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\}|0\rangle &= D_F(x_1 - x_2)D_F(x_3 - x_4) \\ &+ D_F(x_1 - x_3)D_F(x_2 - x_4) \\ &+ D_F(x_1 - x_4)D_F(x_2 - x_3). \end{aligned} \tag{19}$$

On the basis of the intuitive visualization proposed by Feynman himself, we can imagine to depict each of the points x_i ($i = 1, 2, 3, 4$) by a dot, and each factor $D_F(x_i - x_j)$ by a line joining x_i to x_j , with $i \neq j$, thereby obtaining the sum of three Feynman diagrams as represented in Fig. 4.

The LSZ reduction formula applied to the 4-point function reads:

$$\left(\frac{i}{(2\pi)^{\frac{2}{3}}Z^{\frac{1}{2}}}\right)^4 \int dx_1 dx_2 dx_3 dx_4 e^{-i(p_1 x_1 + p_2 x_2)} e^{i(p_3 x_3 + p_4 x_4)} \quad (20)$$

$$(\square_1 + m^2)(\square_2 + m^2) \langle 0 | T \{ \phi(x_1) \phi(x_2) \phi(x_3) \phi(x_4) \} | 0 \rangle (\square_3 + m^2)(\square_4 + m^2),$$

where Z is a computable renormalization constant. In particular, the Fourier transform of the above Eq. (20) shows explicitly how the elements of the S -matrix reduce to the residues of the poles of the correlation functions when the four-momenta are put on-shell: that is,

$$\left(\frac{-i}{(2\pi)^{\frac{2}{3}}Z^{\frac{1}{2}}}\right)^4 (p_1 - m^2)(p_2 - m^2)(p_3 - m^2)(p_4 - m^2) \hat{G}(p_1, p_2, p_3, p_4), \quad (21)$$

with

$$\hat{G}(p_1, p_2, p_3, p_4) = \int dx_1 dx_2 dx_3 dx_4 e^{-i(p_1 x_1 + p_2 x_2)} e^{i(p_3 x_3 + p_4 x_4)} \langle 0 | T \{ \phi(x_1) \phi(x_2) \phi(x_3) \phi(x_4) \} | 0 \rangle. \quad (22)$$

Note that the vacuum expectation value that appears in the last equation is just the Green's function for four points, given by the sum of all Feynman diagrams with four particles created or destroyed in x_1, x_2, x_3 and x_4 , as in Fig. 4. What is most relevant for the reduction procedure we wish to discuss in connection with Earman's principle, though, is that the wave operators \square_i (with $i = 1, 2, 3, 4$) in Eq. (22) have the effect to remove the propagators corresponding to the external legs of Feynman diagrams, as one can easily see in the impulse space where these factors are transformed into $(p^2 - m^2)$. So, the LSZ formula just tells us that the S -matrix elements are nothing but the Green function with the external legs amputated and with the four-momenta forced to be on-shell.

Crucially, after applying the reduction formula, the only surviving terms that contribute to the probability amplitudes in the S -matrix are precisely the propagators of the internal lines, while the asymptotic states of the incoming and outgoing free particles derived in the infinite-time limits are erased, thereby disassembling the set of Feynman diagrams for the scattering process. Since what remains after removing such idealizations are the lines corresponding to the intermediate states of the subatomic world hidden inside the scattering region, owing to Earman's principle the effects of the virtual particles arising therein as propagators of the interaction can be regarded as physically real.

6 Conclusion

Throughout this paper we have developed an indispensability argument in support of the existence of virtual particles in scattering processes. In order to avoid the Paradox of Infinite Limits, one needs to de-idealize the infinite systems introduced by the two limiting procedures employed in the perturbation scheme, namely the expansion in Dyson series and the limits for negative and positive infinite times associated with the assumption of free particles. We have shown that these limits do not introduce

essential idealizations, in compliance with scientific realism. Instead, according to our argument, unobservable virtual particles arise as essential approximations and they should be interpreted as propagators of the interaction responsible for subatomic scattering. In fact, their existence constitutes a case of approximations that matter.

Acknowledgements N. Cangioti was supported of the MIUR - PRIN 2022 project “Nonlinear dispersive equations in presence of singularities” (Prot. N. 20225ATSTP). G. Arioli was supported by the MIUR - PRIN 2022 project “Partial differential equations and related geometric-functional inequalities” (financially supported by the EU, in the framework of the “Next Generation EU initiative”). G. Valente was supported of the MIUR - PRIN 2022 project “Analogical Reasoning in Contemporary Physical Theories” (Prot. N. 2022F4Z8YH). The authors are grateful to Patricia Palacios for helpful comments on their discussion of the notion of essential approximations. They also thank three anonymous referees and the guest editors of this Topical Collection for extensive remarks on earlier versions of the manuscript.

Funding Open access funding provided by Politecnico di Milano within the CRUI-CARE Agreement.

Declarations

Conflict of interest The authors declare that they have no Conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Augustin, J.-E., et al. (1975). Measurement of $e^+e^- \rightarrow e^+e^-$ and $e^+e^- \rightarrow \mu^+\mu^-$. *Physical Review Letters*, 34, 233–236.
- Bain, J. (2000). Against particle/field duality: Asymptotic particle states and interpolating fields in interacting QFT (Or: Who’s Afraid of Haag’s Theorem? *Erkenntnis*, 53(3), 375–406.
- Baker, A. (2009). Mathematical explanation in science. *British Journal for the Philosophy of Science*, 60(3), 611–633.
- Bhabha, H. J. (1935). The creation of electron pairs by fast charged particles. *Proceedings of the Royal Society of London A: Mathematical and Physical and Engineering Sciences*, 152(877), 559–586.
- Batterman, R. (2004). Critical phenomena and breaking drops: Infinite idealizations in physics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 36(2), 225–244.
- Boyd, R. (1983). On the current status of the issue of scientific realism. *Erkenntnis*, 19(1–3), 45–90.
- Brenig, W., & Haag, R. (1959). Allgemeine Quantentheorie der Stoßprozesse. *Fortschritte der Physik*, 7, 183–242.
- Brown, J. R. (1996). Illustration and Inference. In B. Baigrie (Ed.), *Picturing Knowledge: Historical and Philosophical Problems Concerning the Use of Art in Science*. Toronto Canada: University of Toronto Press.
- Busch, P. (2008). The time-energy uncertainty relation. Lecture Notes in Physics In J. Muga, R. S. Mayato, & Í. Egusquiza (Eds.), *Time in quantum mechanics*. (Vol. 734). Springer.
- Butterfield, J. (2013). Reduction, emergence and renormalization. *The Journal of Philosophy*, 111(1), 5–49.

- Dorato, M., & Rossanese, E. (2018). The nature of representation in Feynman diagrams. *Perspectives on Science*, 26(4), 443–458.
- Duncan, A. (2012). *The conceptual framework of quantum field theory*. Oxford University Press.
- Earman, J. (2004). Curie's principle and spontaneous symmetry breaking. *International Studies in the Philosophy of Science*, 18(2–3), 173–198.
- Earman, J., & Fraser, D. (2006). Haag's theorem and its implications for the foundations of quantum field theory. *Erkenntnis*, 64, 305–344.
- Fabbroni, S., et al. (2012). The Young-Feynman two-slits experiment with single electrons: Build-up of the interference pattern and arrival-time distribution using a fast-readout pixel detector. *Ultramicroscopy*, 116, 73–76.
- Feynman, R. P. (1949). Space-time approach to quantum electrodynamics. *Physical Review*, 76(6), 769–789.
- Fox, T. (2008). Haunted by the spectre of virtual particles: A philosophical reconsideration. *Journal for General Philosophy of Science*, 39(1), 35–51.
- Fraser, D. (2009). Quantum field theory: Underdetermination, inconsistency, and idealization. *Philosophy of Science*, 76(4), 536–567.
- Fraser, J. D. (2020). The real problem with perturbative quantum field theory. *The British Journal for the Philosophy of Science*, 71(2), 391–413.
- Griffiths, D. (1987). *Introduction to elementary particles*. Harper & Row.
- Jaeger, G. (2019). Are virtual particles less real? *Entropy*, 21(2), 141.
- Jones, N. J. (2006). *Ineliminable Idealizations, Phase Transitions, and Irreversibility*, PhD Thesis, Ohio State University, Columbus USA.
- Kaiser, D. (2005). *Drawing theories apart: The dispersion of Feynman diagrams in postwar physics*. University of Chicago Press.
- Meynell, L. (2008). Why Feynman diagrams represent. *International Studies in the Philosophy of Science*, 22(1), 39–59.
- Miller, M. E. (2023). Mathematical structure and empirical content. *British Journal for the Philosophy of Science*, 74(2), 511–532.
- Morgan, M. S. (1999). *Models as mediators: Perspectives on natural and social science*. Cambridge University Press.
- Norton, J. D. (2012). Approximation and idealization: Why the difference matters. *Philosophy of Science*, 79(2), 207–232.
- Palacios, P., & Valente, G. (2021). The paradox of infinite limits: A realist response. In T. D. Lyons & P. Vickers (Eds.), *Contemporary scientific realism: The challenge from the history of science*. Oxford University Press.
- Passon, O. (2019). On the interpretation of Feynman diagrams, or, did the LHC experiments observe $H \rightarrow \gamma\gamma$? *European Journal for Philosophy of Science*, 9(2), 20.
- Peskin, M. E., & Schroeder, D. V. (1995). *An introduction to quantum field theory*. Westview Press.
- Redhead, M. (1980). Models in physics. *British Journal for the Philosophy of Science*, 31(2), 145–163.
- Redhead, M. (1988). A philosopher looks at quantum field theory. In H. Brown & R. Harré (Eds.), *Philosophical foundations of quantum field theory*. Oxford University Press.
- Roberts, B. W., & Butterfield, J. (2020). Time-energy uncertainty does not create particles. *Journal of Physics: Conference Series*, 1638, 012005.
- Ruiz de Olano, P., et al. (2022). Taking approximations seriously: The cases of the Chew and Nambu–Jona-Lasinio models. *Studies in History and Philosophy of Science*, 93, 82–95.
- Stöltzner, M. (2017). Feynman diagrams as models. *The Mathematical Intelligencer*, 39, 46–54.
- Stöltzner, M. (2018). Feynman diagrams: Modeling between physics and mathematics. *Perspectives on Science*, 26(4), 482–500.
- Valente, M. B. (2011). Are virtual quanta nothing but formal tools? *International Studies in the Philosophy of Science*, 25, 39–53.
- Walton, K. L. (1990). *Mimesis as make-believe: On the foundations of the representational arts*. Harvard University Press.
- Weingard, R. (1982). Do virtual particles exist? *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 235–242, 1982.
- Weingard, R. (1988). Virtual particles and the interpretation of quantum field theory. In H. Brown & R. Harré (Eds.), *Philosophical foundations of quantum field theory*. Oxford University Press.
- Williams, P. (2023). *Philosophy of particle physics*. Cambridge University Press.

Wüthrich, A. (2018). The exigencies of war and the stink of a theoretical problem: Understanding the genesis of Feynman's quantum electrodynamics as mechanistic modelling at different levels. *Perspectives on Science*, 26(4), 501–520.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.