The Correspondence Principle
In Quantum Field Theory
And Quantum Gravity

Damiano Anselmi

Dipartimento di Fisica "Enrico Fermi", Università di Pisa
and INFN, Sezione di Pisa,
Largo B. Pontecorvo 3, 56127 Pisa, Italy
damiano.anselmi@unipi.it

Abstract

We discuss the fate of the correspondence principle beyond quantum mechanics, specifically in quantum field theory and quantum gravity, in connection with the intrinsic limitations of the human ability to observe the external world. We conclude that the best correspondence principle is made of unitarity, locality, proper renormalizability (a refinement of strict renormalizability), combined with fundamental local symmetries and the requirement of having a finite number of fields. Quantum gravity is identified in an essentially unique way. The gauge interactions are uniquely identified in form. Instead, the matter sector remains basically unrestricted. The major prediction is the violation of causality at small distances.

1   Introduction

Bohr's correspondence principle is a guideline for the selection of theories in quantum mechanics. It is useful to guess the right Hamiltonians of several microscopic systems. Basically, it states that the laws of classical physics must be obtained from the laws of quantum physics in a suitable limit, which is typically the limit of large quantum numbers. More generally, the limit is called *classical limit*.

It may sound a redundant requirement. Any newly discovered laws of nature have to be consistent with the knowledge already available, since nature is one and cannot contradict itself. At the same time, if we view the principle as a guideline for the preselection of theories, it could be argued that anticipating the laws of physics might be dangerous. It could mislead into discarding possibilities that might turn out to be right. This may delay or prevent the discoveries of new laws. In view of these arguments, it might be better to put the idea of a correspondence principle aside and proceed as in every other domain of science, which means making as many experiments as possible and expressing the laws of physics in mathematical language. Quantum physics will exhibit an approximatively classical behavior whenever nature says so.

The whole point, however, is that this procedure is not sufficient at the quantum level, because our possibilities of observing the microscopic world are severely handicapped by the laws of physics themselves. Certainly, they are much smaller than our possibilities to observe the macroscopic world.

This fact is hidden in the meaning of the word *quantization*, which understands that a quantum theory is not built from scratch, but instead guessed from another theory, typically a classical one, which is later quantized. So, there must be some sort of correspondence between the two. *Per se*, this way of proceeding is kind of awkward, since the quantum theory is supposed to be the right one and the classical theory is supposed to be an approximation of it. How can we get the final theory correct, if we guess it from a limit of it? There must be many theories with the same limit. How can we decide which one is the right one?

Yet, for a variety of reasons that we are going to emphasize here, this guesswork is the only thing we can do. And it gets worse when we plan to explore smaller scales of magnitude, where quantum field theory plays a key role. There, the correspondence between quantum physics and classical physics becomes weaker. For example, quantum chromodynamics has a classical limit that has little to do with classical physics. The problem becomes even worse when we move to smaller distances, where quantum gravity

becomes important. We might expect that at some point all sort of correspondence with the classical world will eventually fade away and disappear, leaving us powerless.

We think that the lessons learned from the standard model and certain recent developments in quantum gravity make us ready to take stock of the situation and properly address the problem of the correspondence principle beyond quantum mechanics. Collecting the pieces of the puzzle, we see that there is indeed a need to supplement the scientific method with some sort of guidelines, although we know in advance that their power is doomed to become weaker and weaker when the distances get smaller and smaller.

In this paper we consider assumptions, principles and requirements of various types that have been proved to be important in quantum field theory and high-energy physics. We comment on their ranges of validity and use them to offer an upgraded version of the correspondence principle that better fits the knowledge gathered so far.

To begin with, we stress that we take quantum field theory for granted. We might as well add

0. Quantum field theory

as the zeroth assumption of the upgraded correspondence principle. We think that the success of quantum field theory in particle physics is a sufficient reason to justify the assumption without further comment. Moreover, for the reasons better explained in the next section, we have to be as conservative as possible. Our ability to observe the microscopic world is intrinsically limited, so it is not convenient to depart too much from the kind of correspondence that has worked so far, which is the only source of light we have in a relative darkness. In this respect, we think that approaches alternative to quantum field theory, or even colorful twists of quantum field theory, like holography, have very few chances of success.

2   Do we need a correspondence principle?

In this section we discuss the need of a correspondence principle in quantum field theory, what its meaning is supposed to be and what it should be useful for. We start by considering the intrinsic biological limitations of the human being, which necessarily affect the way we perceive and describe the world. With obvious modifications, the arguments apply to any living being.

First, we must take into account that we do have a size. Our body has a size, like our brains and cells, and the atoms of which we are made. Second, most of our direct

perceptions cover a limited range of energies and distances. The other perceptions of ours, for larger and smaller distances, are indirect, mediated by the instruments we build.

The range of our direct perceptions defines the size of the environment we are placed in. For example, we have eyes that perceive up to a certain resolution, which is $3 \cdot 10^{-4}$ radians [1]. Getting closer to an object, we can improve the accuracy of our visual perception. Combining our mobility with our eye resolution, we can still perceive only a very limited portion of the universe, which we call *direct perception range* (DPR). For the rest, we must rely on indirect measurements and observations. In the long run, indirect perceptions may introduce and propagate errors, particularly in connection with the notions of time and cause.

For concreteness and for various arguments that we plan to develop in the next section, we focus on causality. Everyday experience tells us that many events are connected by relations of cause and effect, that time has an ordering (past $\rightarrow$ present $\rightarrow$ future), that the past influences the present and the future, but the future cannot affect the past and the present. However, time is an extremely delicate perception. Strictly speaking, we have never seen time. We can see the three space dimensions, but time remains confined to our imagination. In some sense, time is the way we organize memories. This makes the entire concept kind of fuzzy.

Specifically, the human brain can process an image perceived for around 10^{-3}s [2] (being optimistic). We call it the *direct perception time resolution* (DPTR). Basically, we see the world at around 1000fps and establish direct relations of cause and effect between pairs of events that are separated by at least the DPTR. Yet, very unlikely the causality principle breaks down right below the DPTR: that would mean that the universe conspired to trick us. We have to assume that causality does hold for shorter time intervals.

If we help ourselves with instruments, we can resolve time intervals that are way shorter than the limits of our direct perception. There exist $5 \cdot 10^{12}$fps cameras, which can capture light in motion [3]. The shortest time interval ever measured is about 10^{-18}s [4]. There are elementary particles with mean lifetimes of about 10^{-25}s.

If we think a moment, we build our instruments on the work hypothesis that the validity of the laws of nature, in particular causality, can be extended below the DPTR. When we use the instruments and cross check that everything works as expected, we get an *a-posteriori* validation of the assumption. This allows us to conclude that, indeed, causality holds well below the DPTR. Yet, having checked that it extends to, say, one billionth of a billionth of the DPTR is still not enough to prove that it holds for arbitrarily short time intervals or large energies. Eventually, it may break down.

Let us have a broader look at the issue. Classically, we can get to the correct physical laws relatively easily, because we can "turn on the light", i.e. throw a huge number of photons on the object of our observation without disturbing it. Our eyes then collect the photons reflected/emitted by the object, which are also a huge number. In practice, we make a very large number of experimental observations at once and pay no price. This is a very lucky situation. Instead, our exploration of the infinitesimally small distances is necessarily handicapped. We cannot turn on the light there, because even a single photon disturbs the object we want to detect. It is like trying and detect a rock by throwing a building at it and checking the results of the scattering.

Luckily, when the distances we explore are not too small, the laws of nature keep some similarity with the classical laws we are accustomed to. The correspondence principle, as we normally understand it, deals with this kind of similarity down to the atomic distances, which are the realm of quantum mechanics. As vague as the principle may sound, a recipe rather than a true principle, it is useful. It describes the relation that links the world we live in to the one we wish to explore, the world that shapes our thinking to the world that makes us what we physically are.

However, we expect that exploring smaller and smaller distances, the correspondence will become weaker and weaker and the devices we build will not help us indefinitely. Actually, it may happen, as quantum mechanics told us, that the possibilities of our experiments are limited by the very same laws of physics. For example, we have to live with the fact that we cannot measure the position and the velocity of a particle simultaneously with arbitrary precision.

The first descent to smaller distances after quantum mechanics is quantum field theory. What is called classical limit there is not necessarily related to classical phenomena, as quantum chromodynamics shows. Nevertheless, an upgraded version of the correspondence principle has worked successfully, so far, at least in the absence of gravity, where a similarity with a (sort of) classical world has more or less survived. The upgraded principle can be codified by means of the requirements of unitarity, locality and renormalizability.

The second descent is quantum gravity, which might require to reconsider or refine the correspondence principle in a nontrivial way.

We may view the matter like this. Our thought is shaped by our interactions with the environment that surrounds us. In some sense, it is a "classical" thought. The very same keywords of our logic (existence, origin, principle, consequence, cause, effect, time, etc.) are inherited from that environment, which means that they might just be useful

approximations and effective descriptions with limited ranges of applicability. When we apply them to the rest of the universe, we assume that our knowledge is "universal", which is far from justified. At some point, we might have to radically modify the laws of physics and even the basic principles of our thinking. Like it or not, the indeterminacy principle is still there, despite so many people whined, tirelessly objected and desperately searched for hidden variables. Actually, we should consider ourselves lucky that we did manage to get somewhere by renouncing determinism. But that means precisely that the principles suggested by our classical experiences were not principles. And that there is no principle that can be trusted to the very end.

In such a situation we might not have much more at our disposal than some sort of "correspondence", even if we know in advance that the correspondence will get weaker and weaker at smaller and smaller distances. What happens at a billionth of a billionth of the DPTR? How would the world appear to us if we could see it at a speed of 10^{26}fps? Is it accurate to talk about causes and effects down there? And what about past, present and future? The same can be said about every other concepts and notions we normally use, since they are inherited from experiences made in a radically different environment and derived from a very approximate and rough perception, if compared with the one that would be required. These handicaps limit our possibilities dramatically. A vague correspondence with what we understand better may be all we can hope for. And we cannot afford to be picky.

Summarizing, the reason why a correspondence principle may be useful in quantum field theory and quantum gravity is rooted in how the quantization works, since the right quantum theory must be identified by starting from a non quantum theory. The environment we wish to explore is so different from the environment we are placed in, that a correspondence between the two may be all we can get. It may help us organize the guesswork to make progress in the relative darkness we have to face.

3   The correspondence principle

In this section, we examine principles, assumptions, properties and requirements that are relevant to quantum field theory. We compare the version of the correspondence principle that fits the standard model in flat space to the versions that fit quantum gravity and the combination of the two. The most important properties are unitarity, locality and renormalizability, which we separate from the rest, since they encode a great part of the

correspondence principle. We first recall some important facts.

The standard model of particle physics is a quantum field theory of gauge fields and matter in flat space. So far, it has been very successful. A large number of predictions have been confirmed, in some cases with high precision. No serious contradiction has emerged and, if right-handed neutrinos are included, even the neutrino masses can be accounted for, without having to modify the fundamental principles. Quantum gravity has been elusive for decades, but recently a consistent theory was formulated [5] (see also [6, 7]) by means of a new quantization prescription that turns certain poles of the free propagators into fake particles, or "fakeons" [8], which can be consistently projected away from the physical spectrum. The theory can be coupled to the standard model with no effort.

The fakeons behave like physical particles in several situations, e.g. when they mediate interactions or decay into physical particles. However, they cannot be detected directly, since the cross section of every process that involves them as initial or final states vanishes. Thanks to this property, the fakeons can be consistently projected away from the physical spectrum. The projected theory is unitary.

The fakeons cannot be eliminated completely from the correlation functions, though, since they are mediators of interactions. This fact leads to an important physical prediction: the violation of causality at energies larger than the fakeon masses. The masses of the fakeons cannot vanish, otherwise causality would be broken at arbitrary distances, contrary to evidence. Actually, they must be sufficiently large to avoid contradiction with the experimental data.

The violation of microcausality survives the classical limit [9], which opens the way to study its effects in nonperturbative configurations, rather than insisting with elementary processes. If we add that the masses of the fakeons are not known and could be much smaller that the Planck mass, we infer that there is hope to detect the effects of the violation in the foreseeable future. They could be the first signs of quantum gravity.

We stress again that in quantum mechanics and quantum field theory the quantization is in some sense a dynamical logical process that builds the right theory starting from a limit of it. When gravity is switched off, the starting classical action is also the action that describes the classical limit $\hbar \to 0$ (which might have nothing to do with classical physics, as in the case of quantum chromodynamics). In quantum gravity, instead, the starting action, which is (3.4), is just an *interim action*, because it misses the projection that throws away the fake degrees of freedom from the physical spectrum. The projection is determined by the quantization process itself. The correct classical limit $\hbar \to 0$ can be obtained by "classicizing" the quantum theory and is encoded in the *finalized classical*

action [9], which is in general nonlocal (when fakeons are present). The interim classical action and the finalized classical action coincide in flat space, where the projection is trivial.

Different quantization prescriptions can lead to inequivalent physical predictions from the same interim classical action. Some quantization prescriptions may even lead to unacceptable consequences. For example, if the interim classical action (3.4) is quantized in the standard way, it has ghosts, instead of fakeons. In that case, unitarity is violated [10]. Thus, a requirement that we may include in a temporary version of the correspondence principle is that an acceptable quantum field theory should only contain physical particles and possibly fakeons, but not ghosts.

If the fakeons are excluded as well, it is possible to have causality at all energies. Then, however, it is not possible to explain quantum gravity by means of a quantum field theory that is local, unitary and renormalizable. For the reasons outlined in the previous section, enforcing microcausality is a stretch, given the limitations of our capabilities to perceive the external world. So, the first principle that has to be sacrificed when we descend to the realm of quantum gravity is microcausality (together with the presumption that we can impose restrictions on nature based on our personal tastes). Moreover, the violation of microcausality might actually be a bonus, instead of a price to pay, since, as stressed above, there are hopes to detect it.

Principles

Now we examine the various candidate ingredients of the correspondence principle, organized in different tiers. The first tier collects the most important principles, which are almost sufficient to encode the upgraded correspondence principle.

1 Unitarity

It is the statement that the scattering matrix S is unitary, i.e. $SS^\dagger = 1$. In particular, there exists a physical Fock space V such that, if $|n\rangle$ denotes an orthonormal basis of V, the identity

$$\sum_{|n\rangle \in V} \langle \alpha|S|n\rangle \langle n|S^\dagger|\beta\rangle = \langle \alpha|\beta\rangle \qquad (3.1)$$

holds for every states $|\alpha\rangle, |\beta\rangle \in V$. In some cases, the physical subspace is identified by means of a suitable projection $W \to V$ starting from a larger, often unphysical Fock space W. Example of unphysical fields that can be consistently projected away are the Faddeev-Popov ghosts and the temporal and longitudinal components of the gauge

fields. Examples of unphysical fields that cannot be consistently projected away are the ghosts of higher-derivative propagators (which typically appear when the poles due to the higher derivatives are quantized using the Feynman prescription). Examples of fields which are neither physical nor unphysical and can be consistently projected away are the fakeons.

The requirement encoded by unitarity is the conservation of probabilities. It ensures that the theory is in some sense complete, i.e. a state cannot appear from nowhere or disappear into nowhere. Basically, there cannot be a preexisting, external source or reservoir. The cosmological constant Λ_C is an example of preexisting conditions, so we cannot demand unitarity in a strict sense in the presence of a nonvanishing cosmological constant, where a meaningful S matrix might not even exist [11]. In the far infrared limit (once the massive fields have been integrated out) Λ_C flows to a constant value Λ_C^* . If $\Lambda_C^* = 0$ unitarity is exact, otherwise it is "anomalous". Note that Λ_C^* is not the measured value of the cosmological constant (indeed, we cannot reach the far infrared limit), so $\Lambda_C^* = 0$ is not in contradiction with the observations. The measured value ($\Lambda_C = 4.33 \cdot 10^{-66} \text{eV}^2$) could be due to the first radiative corrections (indeed, $m_\nu^4/M_{\text{Pl}}^2 \sim 7 \cdot 10^{-65} \text{eV}^2$ for neutrino masses m_ν of the order of 10^{-2}eV).

1a Perturbative unitarity

Perturbative unitarity is the unitarity equation $SS^\dagger = 1$ expressed order by order in the perturbative expansion. It is encoded in a set of diagrammatic identities, called cutting equations [12]. It is often useful to refer to perturbative unitarity rather than unitarity, since the resummations of the perturbative series typically lead to nontrivial widths, which make the particles decay (like the muon in the standard model). Perturbative unitarity allows us to consider processes where the decaying (physical) particles are detected directly while they are "still alive".

2 Locality

Locality is probably the assumption that is more intrinsically related to the quantization process and the correspondence principle. The final quantum theory is encoded in its S matrix, or the generating functional Γ of the one-particle irreducible correlation functions. Both are nonlocal. The finalized classical action is also nonlocal, unless fakeons are absent. Thus, the requirement of locality can only apply to the interim classical action. If we relax this assumption, the quantization process loses most of its meaning. We would have to guess the generating functional Γ directly and run into the problems

described in the previous section. We recall that, in the exploration of the quantum world, we have no possibility of making infinitely many observations in a finite amount of time and/or without disturbing the system. This means that we have no way to determine a theory with an infinite degree of arbitrariness.

The limitations pointed out in the previous section suggest that nature is not arranged to be fully understood or explained by us humans. Therefore, as far as we can tell, the ultimate theory of the universe might well be infinitely arbitrary. Some signs that this is the case are already available (check the discussion below about uniqueness versus the arbitrariness of the matter sector). Nevertheless, the success of quantum field theory and the recent results about quantum gravity give us reasons to believe that we might still have something interesting to say, as long as we do not renounce locality. For these reasons, we regard the locality of the interim classical action as a cornerstone of the correspondence principle and the quantization process.

2a Perturbative locality

Perturbative locality is the version of locality that applies to the nonrenormalizable theories, where the classical action contains infinitely many terms and an arbitrary number of higher derivatives. Resumming those terms leads in general to a nonlocal classical action. The usual perturbative expansion is defined in combination with the expansion of the classical action in powers of the fields and their derivatives. Every truncation of the latter is obviously local. Perturbative locality is the assumption that it makes sense to work with such truncations as approximations of the complete action.

A nonrenormalizable theory (like Einstein gravity equipped with the counterterms turned on by renormalization [13]) is predictive at low energies. For the reasons already stressed, the ultimate theory of the universe might well be nonrenormalizable. However, since we have a better option for the moment, which is the interim classical action (3.8) quantized as explained in ref. [5], i.e. a local, unitary and renormalizable theory that explains both quantum gravity and the standard model, we think that we can postpone this possibility and focus on locality and renormalizability.

3 Renormalizability

Renormalizability has to do with the fact that an interim classical action is not guaranteed to be "stable" with respect to the radiative corrections generated by the perturbative expansion, in particular the removal of their divergent parts. If this kind of stablility

fails, new terms, multiplied by independent parameters, must be added to the interim classical action in order to stabilize it. The theory is nonrenormalizable if the only chance to stabilize it is to include infinitely many terms (which typically means: all the local terms one can build, up to field redefinitions and nonanomalous symmetry requirements), multiplied by independent parameters. It is renormalizable if a finite number of terms or independent parameters is sufficient. Nonrenormalizable theories are predictive at low energies, where perturbative locality ensures that only a finite number of terms are important (those participating in the truncation). Renormalizable theories can in principle be predictive at all energies.

3a Strict renormalizability

Strict renormalizability means that the physical parameters have nonnegative dimensions in units of mass, with respect to the power counting that governs the high-energy behavior of the theory. The counterterms are of finitely many types.

3b Super-renormalizability

Super renormalizability means that all the physical parameters have strictly positive dimensions in units of mass. Then the divergences are finitely many. Super renormalizability does not seem to be favored to describe high-energy physics.

3c Proper renormalizability

We introduce this refinement of the notion of strict renormalizability, because it is particularly useful, in combination with other key requirements, to single out a unique theory of quantum gravity (see below). Proper renormalizability means that the gauge couplings (including the Newton constant) must be dimensionless (with respect to the power counting governing the ultraviolet behaviors of the correlation functions) and the other physical parameters must have nonnegative dimensions in units of mass.

4 Fundamental symmetry requirements

The symmetries can be global or local; exact, explicitly broken, spontaneously broken or anomalous. The local symmetries are also called (generalized) gauge symmetries and mediate interactions. They must be exact or spontaneously broken, otherwise unitarity is violated. They include both the gauge (i.e. Yang-Mills) symmetries as such and the local symmetries of gravity (i.e. invariance under general changes of coordinates and local Lorentz invariance).

4a   Global Lorentz invariance

Lorentz invariance is a global symmetry in flat space. There, it can be explicitly broken without violating unitarity, in which case many vertices that are normally nonrenormalizable become renormalizable [14], even if the fakeons are forbidden. However, in quantum gravity Lorentz invariance is a local symmetry, which cannot be dynamically or explicitly broken without violating unitarity. Thus, where gravity exists, Lorentz invariance must be exact or spontaneously broken. Since gravity exists at low energies, Lorentz invariance can be violated only at high energies. If this occurred, gravity would be a low-energy effective phenomenon, emerging from a radically different high-energy picture, of which, however, there is at present no idea. For these reasons, in most arguments of this paper we assume that Lorentz symmetry is exact or spontaneously broken.

4b   General covariance

4c   Local Lorentz invariance

4d   Gauge invariance

The second tier of properties collects consequences of the first tier and supplementary requirements.

5   Uniqueness

By uniqueness, or essential uniqueness, we mean that the theory is determined up to a finite number of independent physical parameters (which need to be measured experimentally) and a finite number of options for the quantization prescription. We could specify that the total number of possibilities must be "small" (with respect to our biological and physical limits), but there is no objective definition of smallness that we can use here, so we prefer to leave this point unresolved. Yet, it is important to emphasize that the requirement of uniqueness excludes the nonrenormalizable theories and most theories with infinitely many fields, unless their physical parameters are somehow related to one another and ultimately just depend on a finite number of independent ones.

Investigations about consistent reductions on the number of independent physical parameters in renormalizable theories dates back to Zimmermann and Oheme [15]. The reductions in nonrenormalizable theories in flat space have been studied in refs. [16]. The extension of the reduction to the nonrenormalizable theory of quantum gravity (which is the Hilbert-Einstein theory equipped with all the corrections turned on by

renormalization and cleared of all its higher-derivative quadratic terms by means of field redefinitions [13]) poses serious problems, since there is no way to have a perturbative control on it. However, the reduction has a chance in higher-derivative gravity, due to the presence of the higher-derivative quadratic terms, which play a crucial role [17].

A correspondence principle that points to a unique theory would overcome our handicapped perception of the microscopic world. However, various signals, like the arbitrariness of the matter sector of the standard model, tell us that we must probably cope with the fact that this goal is utopian. Nevertheless, we can have uniqueness in gravity and, to some extent, within the gauge interactions, as expressed by the following points 5a and 5b.

5a   Uniqueness in form of the gauge interactions

The combination of unitarity, 1, locality, 2, proper renormalizability, 3c, and Lorentz invariance, 4a, is a very powerful correspondence principle in flat space. Indeed, if we also forbid the presence of fakeons, 10d, the set of these requirements implies 4d, 8 and 10, i.e. it determines the gauge transformations [18], the form of the action and even that the spacetime dimension D must be equal to four. The action is the Yang-Mills one,

$$S_{\mathrm{YM}} = -\frac{1}{4} \int \mathrm{d}^4 x \sqrt{-g} F_{\mu\nu}^a F^{a\mu\nu}, \qquad (3.2)$$

where $F_{\mu\nu}^a$ denotes the field strength.

If we replace 3c with 3a or 3, the allowed dimensions are 4, 3, 2, and 1. However, the universe predicted by quantum field theory is too simple below four dimensions, so we regard four as the minimum value. In this sense, 3c, 3a and 3 can be considered equivalent at this level.

If we relax 10d by allowing massive fakeons, which is requirement 10c, then we just have macrocausality, 10a, instead of causality, and there are solutions in every even spacetime dimensions $D \geqslant 6$. Their interim classical actions are

$$S_{\mathrm{YM}}^D = -\frac{1}{4} \int \mathrm{d}^D x \sqrt{-g} \left[F_{\mu\nu}^a P_{(D-4)/2}(\mathcal{D}^2) F^{a\mu\nu} + \mathcal{O}(F^3) \right], \qquad (3.3)$$

where $P_n(x)$ is a real polynomial of degree n in x and $\mathcal{D}$ is the covariant derivative, while $\mathcal{O}(F^3)$ are the Lagrangian terms that have dimensions smaller than or equal to D and are built with at least three field strengths and/or their covariant derivatives. Observe that by 3c the gauge coupling is dimensionless. We have used Bianchi

13

identities and partial integrations to simplify the quadratic terms. The coefficients of the polynomial $P_{(D-6)/2}$ must satisfy a few restrictions, so that, after projecting away the gauge modes (by working, for example, in the Coulomb gauge), the poles of the propagators have squared masses with nonnegative real parts and the massless poles have positive residues.

The poles with negative or complex residues, as well as those with positive residues but complex masses, must be quantized as fakeons. The poles with positive residues and nonvanishing real masses can be quantized either as fakeons or physical particles. Those with vanishing masses must be quantized as physical particles.

In all the situations just described, the gauge group remains essentially free, as long as it is unitary and (toghether with the matter content) satisfies the anomaly cancellation conditions (which are other consequences of unitarity). The knowledge we have today does not explain why the gauge group of the standard model is precisely the product of the three simplest groups, $U(1)$, $SU(2)$ and $SU(3)$, and why, say, $SU(37)$, $SU(41)$, etc., are absent. We cannot anticipate if other gauge groups will be discovered (possibly completely broken or confined). There could even be a sort of "periodic table" of the gauge groups and we might just have grabbed the three simplest representatives. This issue is tied in some way to the non uniqueness of the matter sector (see 5c below).

5b **Uniqueness of the gravitational interactions**

Gravity does not have this problem, because its local symmetry (invariance under diffeomorphisms times local Lorentz invariance) is unique. The requirements 1, 2, 3c, 4b, 4c, 8 and 10c lead to the unique interim classical action

$$S_{\text{QG}} = -\frac{1}{2\kappa^2} \int \sqrt{-g} \left[2\Lambda_C + \zeta R + \alpha \left(R_{\mu\nu} R^{\mu\nu} - \frac{1}{3} R^2 \right) - \frac{\xi}{6} R^2 \right] \qquad (3.4)$$

in four dimensions, and also select the quantization prescriptions that are physically acceptable. In formula (3.4) α, ξ, ζ and κ are real positive constants, and Λ_C can be positive or negative. The action must be quantized as explained in ref. [5]. It propagates the graviton, a scalar ϕ of squared mass $m_\phi^2 = \zeta/\xi$ (which can be quantized as a physical particle or a fakeon) and a spin-2 fakeon $\chi_{\mu\nu}$ of squared mass $m_\chi^2 = \zeta/\alpha$ [1].

We recall that if we use the Feynman quantization prescription for all the fields, we obtain the Stelle theory [10], where $\chi_{\mu\nu}$ is a ghost. In that case, unitarity is violated.

---

[1]In this formula, we are neglecting a small correction due to the cosmological constant [7].

The assumption 8 that the spacetime dimensions are four has been included explicitly. If we eliminate it, the requirements 1, 2, 3c, 4b, 4c and 10c admit solutions in every even dimensions greater than or equal to four. Their interim classical actions read

$$S_{\text{QG}}^D = -\frac{1}{2\kappa^2} \int \sqrt{-g} \left[2\Lambda_C + \zeta R + R_{\mu\nu} \mathcal{P}_{(D-4)/2}(\mathcal{D}^2) R^{\mu\nu} \right.$$
$$\left. + R\mathcal{P}'_{(D-4)/2}(\mathcal{D}^2) R + \mathcal{O}(R^3) \right], \qquad (3.5)$$

where $\mathcal{P}_n$ and $\mathcal{P}'_n$ denote other real polynomials of degree n and $\mathcal{O}(R^3)$ are the Lagrangian terms that have dimensions smaller than or equal to D and are built with at least three curvature tensors and/or their covariant derivatives. Again, the squared masses must have nonnegative real parts and the poles of the propagators must be quantized as explained before.

5c Non uniqueness of the matter sector

As far as we know today, quantum field theory cannot predict the matter content of the theory that describes nature. For example, we can enlarge the standard model coupled to quantum gravity by including new massive particles and/or massive fakeons, as long as they satisfy the anomaly cancellation conditions and are heavy enough, so that their presence does not affect the experimental results available today. The ultimate theory could even contain infinitely many matter fields. This is a point where the correspondence principle has been almost completely powerless. Probably, it is a sign of the fading correspondence and it might be impossible to remedy in the future. Let us remark that every attempt to relate the matter content to the interactions beyond the anomaly cancellation conditions (grand unification, supersymmetry, string theory and so on) has failed.

6 Analyticity

Analyticity ensures that it is sufficient to calculate an amplitude, or a loop diagram, in any open subset of the space $\mathcal{P}$ of the complexified external momenta to derive it everywhere in $\mathcal{P}$ by means of the analytic continuation. It holds if the theory contains only physical particles.

6a Regionwise analyticity

Regionwise analyticity is the generalization of analyticity that holds when the theory contains fakeons in addition to physical particles. The space $\mathcal{P}$ is divided into disjoint

regions of analyticity. An amplitude, or a loop diagram, is analytic in each region, but the relation between the results found in different regions is not analytic. The main region is the Euclidean one, which contains neighborhoods of the imaginary energies. It is sufficient to calculate an amplitude, or a loop diagram, in any open set of the Euclidean region to derive it everywhere in $\mathcal{P}$ by means of a nonanalytic operation, called average continuation. The average continuation is the arithmetic average of the two analytic continuations that circumvent a branch point [19, 8].

7 Existence of interactions

7a Existence of the gauge interactions

7b Existence of gravity

8 Four spacetime dimensions

Quantum field theory predicts that if the spacetime dimensions of the universe were smaller than four, then the universe would be too simple. So, they must be at least four. It is an experimental fact that they are four at large distances, but in principle they could be more at small distances. If that were that case, they would have to be compactified at low energies. This makes the higher-dimensional theories equivalent to four-dimensional theories with infinite sets of matter fields. Moreover, the compactification involves manifolds of even dimensions $\geqslant 2$, which cause the appearance of infinitely many independent parameters (the moduli of the compactification). Since we do not have robust arguments to restrict the matter content, as already pointed out, these possibilities remain open.

9 Finite numbers of fields and independent physical parameters

A restriction on the matter content is to demand that the set of fields be at least finite. Only in $D = 4$ the theory can have finitely many fields and finitely many parameters, because the need of a compactification to $D = 4$ introduces an infinite arbitrariness in the matter sector of the theories with $D > 4$.

In the last tier, we include other more or less important consequences and properties.

9 Consistency requirements

9a Well-defined Hilbert (Fock) space (positive definite norms)

9b Hermitian Hamiltonian bounded from below

10 Causality

Everyday experience tells us that the future is determined by the past and effects are determined by causes. On a closer look, however, these statements are rather vague, to say the least. In a deterministic framework, for example, it is true that the initial conditions uniquely determine the future, but we could turn the argument around and claim that the future uniquely determines the past, for the same reason. Is it appropriate to speak about causes, when the future is already determined?

To avoid paradoxes like this, we should define the notions of cause and effect with precision. However, it is not that easy. At most, we can replace the intuitive ideas with more formal or mathematical definitions on which we can agree. The drawback of this approach is that we might have to declassify them to secondary properties rather than fundamental principles. This is a risk that all "principles" face, actually, because at the end they are formal requirements (see section 4 for more comments on this), so which is principle and which is side property is a matter of how effective each of them is.

At the classical level, for example, causality can be formulated as the requirement that the field equations not involve the sources of interactions located in the future light cone and at spacelike separations.

If we adopt this definition, quantum field theory predicts that the laws of nature are causal in the classical limit, in flat space. On the other hand, quantum gravity predicts corrections to the (classical) field equations of general relativity that do require the knowledge of interactions at future times, albeit restricted to finite ranges of spacetime separations [9]. We can say that the corrected Einstein equations violate microcausality, but satisfy macrocausality.

We can define causality in quantum field theory by adding the requirement that the commutators of spacelike separated observables vanish.

10a Macrocausality

At the classical level, we define macrocausality as the requirement that there exist finite positive thresholds τ and σ such that the field equations in x not involve the sources of interactions located in $y \in \mathcal{U}_{x,\tau} \bigcup \mathcal{V}_{x,\sigma}$, where $\mathcal{U}_{x,\tau} = \{y : (x-y)^2 < -1/\tau\}$ and $\mathcal{V}_{x,\sigma}\{y : x^0 - y^0 < 0, (x-y)^2 > 1/\sigma\}$. Alternatively, they involve those sources by negligible amounts that tend to zero when $|(x-y)^2| \to \infty$.

At the quantum level, we define macrocausality by adding the requirement that the commutators of any two local observables $\mathcal{O}(x)$ and $\mathcal{O}(y)$ are negligible for every x and every $y \in \mathcal{U}_{x,\tau}$ and tend to zero for $|(x-y)^2| \to \infty$.

Macrocausality is supported by experimental evidence. In quantum field theory, it follows as a bonus if we do not compromise the other major requirements (in particular, we do not renounce the locality of the interim classical action), as long as a much weaker form of macrocausality holds, which we define below.

10b Microcausality

It is the mirror of macrocausality, where $\mathcal{U}_{x,\tau}$ is replaced by $\bar{\mathcal{U}}_{x,\tau} = \{y : -1/\tau \leqslant (x-y)^2 < 0\}$ and $\mathcal{V}_{x,\sigma}$ is replaced by $\bar{\mathcal{V}}_{x,\sigma} = \{y : x^0 - y^0 < 0, 0 \leqslant (x-y)^2 \leqslant 1/\sigma\}$. We are not aware of ways of violating it without also violating macrocausality, unless we compromise the basic architecture of quantum field theory, in particular locality.

10c Weak macrocausality

It is useful to define a weak notion of macrocausality, which is the requirement that the theory not contain massless fakeons. Indeed, in most cases macrocausality as defined above follows for from free, when weak macrocausality is assumed. Note that weak macrocausality allows the theory to contain massive fakeons.

10d Weak microcausality

It can be defined as the requirement that the theory have no fakeons.

11 Ultraviolet behavior

Some theories have particularly nice ultraviolet behaviors. For example, quantum chromodynamics is asymptotically free. Long ago, Weinberg [20] suggested a generalization of asymptotic freedom, which is asymptotic safety, where the ultraviolet limit is an interacting conformal field theory with a finite dimensional critical surface. Evidence of asymptotic safety in quantum gravity has been found in refs. [21]. These properties are interesting from the theoretical point of view. However, the standard model does not seem to have a nice ultraviolet behavior, so requirements like asymptotic freedom and asymptotic safety sound very restrictive.

11a Asymptotic freedom

11b Asymptotic safety

12 Positive definiteness of the (bosonic sector of the) Euclidean theory

This property concerns the theories in the absence of gravity, where the interim and finalized classical actions coincide. When fakeons are present, it is not meaningful to demand that the Euclidean version of the interim classical action, which is unprojected, be positive definite, even when it is purely bosonic.

Combinations of principles

Now we comment on the properties implied by various combinations of requirements. We start from the standard model in flat space, which suggests a correspondence principle made of (I) 1, 2, 3, 4a, 4d, 7a and 10d. These requirements imply 5a, 6, 8, 10 and 12, but they are not compatible with the existence of gravity, 7b. We obtain another correct correspondence if (II) we replace the assumption 10d with 6. If (III) we drop 6 and replace 10d with 10c, we may include massive fakeons and have Yang-Mills theories (3.3) in arbitrary even dimensions.

Coming to quantum gravity, the theory (3.4) and its higher dimensional versions (3.5) suggest a correspondence principle made of (IV) 1, 2, 3c, 4a, 4b, 4c, 9 and 10c, which implies 5b, 6a, 8 and 10a, but not 6, 10b and 10d. If we drop the assumption 9 we have to renounce the implication 8.

Quantum gravity coupled to the (covariantized) standard model suggests the correspondence principle made of (V) 1, 2, 3c, 4, 9 and 10c, which implies 5a, 5b, 6a, 8 and 10a, but not 6, 10b and 10d. Again, (VI) if we drop the assumption 9 we renounce the implication 8.

The results found so far are summarized in the table

	Assumptions	Implications	Missing implications
(I)	1, 2, 3, 4a, 4d, 7a, 10d	5a, 6, 8, 10, 12	7b
(II)	1, 2, 3, 4a, 4d, 7a, 6	5a, 8, 10, 12	7b
(III)	1, 2, 3, 4a, 4d, 7a, 10c	5a, 6a, 10a	6, 7b, 8, 10b, 10d
(IV)	1, 2, 3c, 4a, 4b, 4c, 7b, 9, 10c	5b, 6a, 8, 10a	6, 10b, 10d
(V)	1, 2, 3c, 4, 7a, 7b, 9, 10c	5a, 5b, 6a, 8, 10a	6, 10b, 10d
(VI)	1, 2, 3c, 4, 7a, 7b, 10c	5a, 5b, 6a, 10a	6, 8, 10b, 10d

$$(3.6)$$

If we relax the requirement 3c of proper renormalizability by converting it into simple renormalizability, 3, then we lose the uniqueness properties 5a and 5b, because in every spacetime dimensions D infinitely many super-renormalizable theories of quantum gravity and gauge fields with fakeons are admitted. Their interim actions are obtained from (3.3) and (3.5) by raising the degrees n of the polynomials $\mathcal{P}_n$ and $\mathcal{P}'_n$.

Summarizing, a successful correspondence principle is made of 1, 2, 3c, 4 and 9, i.e.

$$
\begin{array}{c}
\text{unitarity} \\
\text{locality} \\
\text{proper renormalizability} \\
\text{fundamental symmetries}
\end{array}
\tag{3.7}
$$

and the requirements of having (i) finitely many fields and parameters and (ii) no massless fakeons. This combination implies quantum gravity coupled to the standard model in four dimensions, with interim classical action

$$
S_{\mathrm{QG}} + S_m,
\tag{3.8}
$$

where S_m is the covariantized action of the standard model (or one of its extensions), equipped with the nonminimal terms compatible with renormalizability.

If we drop the assumption 9, solutions with analogous properties exist in every even spacetime dimension D, made of (3.3), (3.5) and the matter sector. In that case, D plays the role of an additional physical parameter that must be measured experimentally.

With or without 9, microcausality is violated, analyticity is replaced by regionwise analyticity, the gravitational interactions are essentially unique, the Yang-Mills interactions are unique in form and the matter sector remains basically unrestricted.

With respect to the version of the correspondence principle that is successful in flat space, the only upgrade required by quantum gravity amounts to renouncing analyticity, 6, in favor of regionwise analyticity, 6a, and settle for macrocausality, 10c, instead of full causality, 10. As we wanted, the final solution is as conservative as possible. Moreover, the violation of microcausality is turned into a physical prediction, which might be confirmed experimentally if a suitable amplification mechanism is found.

4   Conclusions

The fates of determinism and possibly causality are there to remind us that the correspondence between the environment we live in, which shapes our thinking, and the microscopic world is doomed to become weaker and weaker as we explore smaller and smaller distances. At some point, we face the intrinsic limitations of our ability to understand the universe. Maybe the impossibility to effectively restrict the matter sector and the gauge group of the standard model is a further sign of the fading correspondence.

The surviving correspondence principle (3.7) is not made of physical requirements, but mostly formal ones. Unitarity, as said, is a requirement of completeness. Locality is intrinsically tied to the dynamics of the logical process that builds the quantum theory from an interim classical theory, the final theory being nonlocal anyway. As far as (proper) renormalizability is concerned, it it hard to view it as more than formal.

This is an interesting turn of events. Perhaps ironic, but consistent with what we have been remarking all over this paper. Insisting on physical requirements would be like requesting that nature adapt to us, rather than coping with the fact that we have to adapt to nature. What are the odds that a physical intuition shaped by a classical environment gets it right, when it comes to the phenomena of the infinitesimally small? Not many. We have been knowing that for a fact since the birth of quantum mechanics. And what if we cannot understand why the correspondence principle is made of (3.7) instead of something else? Even better, we might argue. Indeed, if we understood that, the principle would probably be inadequate, because it would be fit to describe our world, but unfit to describe the infinitesimal world.

Nevertheless, it is important to stress that the cornerstones of (3.7) are not chosen because they are appealing, but simply because they work. The opposite attitude — choose something appealing and pretend it works — is very fashionable nowadays, but spectacularly unsuccessful [22]. The failures accumulated during the past decades in the pursuit of beauty, mathematical elegance, symmetry, supersymmetry, unification, grandunification, theories of everything and theories with no parameters, and then holography and who knows what will be next, remind us of the infinite monkey theorem. Are humans that stupid? We cannot exclude it. The possibility that humans are ill equipped to embrace the challenges demanded by the exploration of the smaller and smaller distances is high. We are already witnessing a slow, relentless involution into a realm where the main judgment criteria are the authority principle or its modern distorsions, such as the counting of likes, called citations. Sadly, this turn of events is not even new: it is called Middle Ages. But the puzzling question is: how long will it last this time?

References

[1] M. Yanoff and J.S.Duker, *Ophthalmology*, MOSBY Elsevier, 3rd Edition (2009), p. 54.

[2] M.C. Potter, B. Wyble, C.E. Hagmann, et al., Detecting meaning in RSVP at 13 ms per picture, Atten Percept. Psychophys. 76 (2014) 270.

[3] See for example, P. Monckton, World's fastest camera captures over five trillion shots per second, Forbes, Apr 30, 2017.

[4] M. Ossiander, F. Siegrist, V. Shirvanyan, R. Pazourek, A. Sommer, T. Latka, A. Guggenmos, S. Nagele, J. Feist, J. Burgdörfer, R. Kienberger and M. Schultze, Attosecond correlation dynamics, Nature Physics 13 (2017) 280.

[5] D. Anselmi, On the quantum field theory of the gravitational interactions, J. High Energy Phys. 06 (2017) 086, 17A3 Renormalization.com and arXiv: 1704.07728 [hep-th].

[6] D. Anselmi and M. Piva, The ultraviolet behavior of quantum gravity, J. High Energ. Phys. 05 (2018) 27, 18A2 Renormalization.com and arXiv:1803.07777 [hep-th] and

[7] D. Anselmi and M. Piva, Quantum gravity, fakeons and microcausality, 18A3 Renormalization.com and arXiv:1806.03605 [hep-th].

[8] D. Anselmi, Fakeons and Lee-Wick models, J. High Energy Phys. 02 (2018) 141, 18A1 Renormalization.com and arXiv:1801.00915 [hep-th].

[9] D. Anselmi, Fakeons, microcausality and the classical limit of quantum gravity, 18A4 Renormalization.com and arXiv:1809.05037 [hep-th].

[10] K.S. Stelle, Renormalization of higher derivative quantum gravity, Phys. Rev. D 16 (1977) 953;

for the one-loop beta functions of the theory, see [6] and I. G. Avramidi and A. O. Barvinsky, Asymptotic freedom in higher derivative quantum gravity, Phys. Lett. B 159 (1985) 269;

N. Ohta, R. Percacci and A.D. Pereira, Gauges and functional measures in quantum gravity II: Higher-derivative gravity, Eur. Phys. J. C 77 (2017) 611 and arXiv:1610.07991 [hep-th];

for recent applications to phenomenological problems and the beta functions in the presence of matter, see A. Salvio and A. Strumia, Agravity, J. High Energy Phys. 06 (2014) 80 and arXiv:1403.4226 [hep-ph];

A. Salvio and A. Strumia, Agravity up to infinite energy, Eur. Phys. C 78 (2018) 124 and arXiv:1705.03896 [hep-th].

[11] See for example, S.B. Giddings, The boundary S-matrix and the AdS to CFT dictionary, Phys. Rev. Lett. 83 (1999) 2707 and arXiv:hep-th/9903048;

V. Balasubramanian, S.B. Giddings, A. Lawrence, What do CFTs tell us about anti-de Sitter spacetimes?, J. High Energy Phys. 9903 (1999) 001 and arXiv:hep-th/9902052;

J. Bros, H. Epstein, M. Gaudin, U. Moschella and V. Pasquier, Triangular invariants, three-point functions and particle stability on the de Sitter universe, Commun. Math. Phys. 295 (2010) 261 and arXiv:0901.4223 [hep-th];

E.T. Akhmedov and P. V. Buividovich, Interacting field theories in de Sitter space are non-unitary, Phys. Rev. D 78 (2008) 104005 and arXiv:0808.4106 [hep-th];

E.T. Akhmedov, Lecture notes on interacting quantum fields in de Sitter space, Int. J. Mod. Phys. D 23 (2014) 1430001 and arXiv:1309.2557 [hep-th]

D. Marolf, I. A. Morrison, and M. Srednicki, Perturbative S-matrix for massive scalar fields in global de Sitter space, Class. Quant. Grav. 30 (2013) 155023 and arXiv:1209.6039 [hep-th].

[12] R.E. Cutkosky, Singularities and discontinuities of Feynman amplitudes, J. Math. Phys. 1 (1960) 429;

M. Veltman, Unitarity and causality in a renormalizable field theory with unstable particles, Physica 29 (1963) 186.

[13] For a convenient way to organize the action, see for example, D. Anselmi, Properties of the classical action of quantum gravity, J. High Energ. Phys. 05 (2013) 028, 13A2 Renormalization.com and arXiv:1302.7100 [hep-th].

[14] See, for example, D. Anselmi and M. Halat, Renormalization of Lorentz violating theories, Phys. Rev. D76 (2007) 125011 and arXiv:0707.2480 [hep-th];

D. Anselmi, Standard model without elementary scalars and high energy Lorentz violation, Eur. Phys. J. C65 (2010) 523 and arXiv:0904.1849 [hep-ph].

[15] W. Zimmermann, Reduction in the number of coupling parameters, Comm. Math. Phys. 97 (1985) 211;

R. Oheme and W. Zimmermann, Relation between effective couplings for asymptotically free theories, Comm. Math. Phys. 97 (1985) 569;

R. Oehme, Reduction and reparametrization of quantum field theories, Prog. Theor. Phys. Suppl., 86 (1986) 215;

for recent phenomenological applications, see S. Heinemeyer, M. Mondragon, N. Tracas and G. Zoupanos, Reduction of the Parameters in MSSM, J. High Energ. Phys. 08 (2018) 150 and arXiv:1712.02729 [hep-ph];

S. Heinemeyer, J. Kubo, M. Mondragon, O. Piguet, K. Sibold, W. Zimmermann and G. Zoupanos, Reduction of couplings and its application in particle physics, Finite theories, Higgs and top mass predictions, PoS(Higgs & top)001 and arXiv:1411.7155 [hep-ph].

[16] D. Anselmi, Consistent irrelevant deformations of interacting conformal field theories, J. High Energy Phys. 10 (2003) 045 and arXiv:hep-th/0309251;

D. Anselmi, Infinite reduction of couplings in non-renormalizable quantum field theory, J. High Energy Phys. 08 (2005) 029 and arXiv:hep-th/0503131.

[17] M. Atance and J. L. Cortés, Effective field theory of pure gravity and the renormalization group, Phys. Lett. B387 (1996) 697 and arXiv:hep-th/9604076;

M. Atance, J.L. Cortés, Effective field theory of gravity, reduction of couplings and the renormalization group, Phys. Rev. D54 (1996) 4973 and arXiv:hep-ph/9605455.

[18] J.M. Cornwall, D.N. Levin and G. Tiktopoulos, Uniqueness of spontaneusly broken gauge theories, Phys. Rev. Lett. 30 (1973) 1268.

[19] D. Anselmi and M. Piva, A new formulation of Lee-Wick quantum field theory, J. High Energy Phys. 06 (2017) 066, 17A1 Renormalization.com and arXiv:1703.04584 [hep-th].

[20] S. Weinberg, Ultraviolet divergences in quantum theories of gravitation, in *An Einstein centenary survey*, Edited by S. Hawking and W. Israel, Cambridge University Press, Cambridge 1979, p. 790.

[21] O. Lauscher and M. Reuter, Ultraviolet fixed point and generalized flow equation of quantum gravity, Phys. Rev. D65 (2002) 025013 and arXiv:hep-th/0108040;

O. Lauscher and M. Reuter, Flow equation of quantum Einstein gravity in a higher-derivative truncation, Phys. Rev. D66 (2002) 025026 and arXiv:hep-th/0205062;

D.F. Litim, Fixed points of quantum gravity, Phys. Rev. Lett. 92 (2004) 201301 and arXiv:hepth/0312114;

for a recent update, see K.G. Falls, C.R. King, D.F. Litim, K. Nikolakopoulos, C. Rahmede, Asymptotic safety of quantum gravity beyond Ricci scalars, Phys. Rev. D 97 (2018) 086006 and arXiv:1801.00162 [hep-th].

[22] D. Falk, Why some scientists say physics has gone off the rails. Has the love of "elegant" equations overtaken the desire to describe the real world?, NBC News (2018).