Eur. Phys. J. C (2016) 76:227
DOI 10.1140/epjc/s10052-016-4079-8

THE EUROPEAN PHYSICAL JOURNAL C

CrossMark

Regular Article - Theoretical Physics

Quantum mechanics of 4-derivative theories

Alberto Salvio[1,a], **Alessandro Strumia**[2,3]

[1] Departamento de Física Teórica, Universidad Autónoma de Madrid and Instituto de Física Teórica IFT-UAM/CSIC, Madrid, Spain
[2] Dipartimento di Fisica dell'Università di Pisa and INFN, Pisa, Italy
[3] Theory Division, CERN, Geneva, Switzerland

Received: 1 March 2016 / Accepted: 8 April 2016 / Published online: 25 April 2016

Abstract A renormalizable theory of gravity is obtained if the dimension-less 4-derivative kinetic term of the graviton, which classically suffers from negative unbounded energy, admits a sensible quantization. We find that a 4-derivative degree of freedom involves a canonical coordinate with unusual time-inversion parity, and that a correspondingly unusual representation must be employed for the relative quantum operator. The resulting theory has positive energy eigenvalues, normalizable wavefunctions, unitary evolution in a negative-norm configuration space. We present a formalism for quantum mechanics with a generic norm.

Contents

1 Introduction

Newton invented classical mechanics putting two time derivatives in his equation $F = m\ddot{x}$, which corresponds to a kinetic energy with two time derivatives, $m\dot{x}^2/2$. Later Ostrogradski proved a no-go theorem: non-degenerate classical systems with more than two time derivatives contain arbitrarily negative energies and develop fatal run-away instabilities [1]. Classically, they do not make sense.

The discovery that nature is relativistic and quantum opened the quest for an extension of Newtonian gravity. A century ago Einstein and Hilbert found the classical theory of relativistic gravity. However, its quantum version is not renormalizable in 3+1 space-time dimensions. Sticking to the observed number of space-time dimensions, a renormalizable extension of general relativity is found by adding terms quadratic in the curvature tensor to the Einstein–Hilbert Lagrangian, such that the graviton acquires a 4-derivative kinetic term. Stelle proposed and dismissed this extension [2] (see also [3–13]).

Recently the Higgs mass hierarchy problem brought interest to dimension-less theories. In this context, gravitons must have dimension 0 (being a dimension-less metric) and thereby *must* have a 4-derivative kinetic term. If these theories could make sense at the quantum level, despite the negative classical energy, a great deal would be gained: relativistic quantum gravity, plus hierarchies among dynamically generated mass scales [12], plus inflation [14,15].

Quantization can eliminate arbitrarily negative classical energies. The following example is well known: the classical relativistic spin 1/2 field is described by a spinor $\Psi(x)$ with Dirac Lagrangian $\mathscr{L} = \bar{\Psi}(i\partial\!\!\!/ - m)\Psi$ containing one time derivative. Treating Ψ as a classical field (as initially proposed by Schrödinger), and inserting the plane-wave expansion

$$\Psi(x) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3\sqrt{2E_p}}[a_{p,s}u_{p,s}\mathrm{e}^{-ip\cdot x} + b^\dagger_{p,s}v_{p,s}\mathrm{e}^{ip\cdot x}] \tag{1}$$

in the Hamiltonian one finds negative energies in half of the configurations space:[1]

$$H = \int \frac{\mathrm{d}^3 p}{(2\pi)^3}\, E_p[a^\dagger_{p,s}a_{p,s} - b_{p,s}b^\dagger_{p,s}],$$
$$E_p = \sqrt{m^2 + \vec{p}^2} \tag{2}$$

[a] e-mail: alberto.salvio@cern.ch

[1] Unlike in the case of the hydrogen atom emphasized by Woodard [16,17], where the instability eliminated by quantum mechanics occurs only in one point of the configuration space.

This classical arbitrarily negative energy is avoided by quantization with anti-commutators, if the vacuum state is appropriately chosen. Indeed, the two-state solution to $\{b, b^\dagger\} = 1$ shows that one can switch annihilation with creation operators by choosing the vacuum to be the state with lower energy.

The spin 0 and spin 1 relativistic fields (described by dimension-1 fields with two derivatives) do not have this issue: the negative-frequency solutions to the Klein–Gordon equation correspond to Hamiltonians with positive energy.

The general lesson is that quantization depends on the number of time derivatives.

The goal of this study is describing if/how systems with four derivatives can be quantized obtaining a consistent theory, in particular of quantum gravity. We will find that a unique structure emerges, which again involves switching annihilation and creation operators.

This will bring us into the territory of negative-norm quanta, avoided like a plague by serious theorists that call them 'ghosts', and explored only by notorious crackpots such as Dirac [18], Pauli [19], Heisenberg [20], Pais, Uhlenbeck [21], Lee, Wick [22,23], Cutkosky [24], Coleman [25], Feynman [26], Gross [27], Hawking [28,29], 't Hooft [30,31] and others, also more recently [32–52]. These works sometimes contain bizarre and confusing statements and obsolete motivations, together with interesting ideas and ad hoc prescriptions.

This paper is structured as follows. In Sect. 2 we review the canonical Ostrogradski formalism. In Sect. 3 we present negative-norm quantum mechanics, the negative-norm harmonic oscillator (Sect. 3.3), and the associated negative-norm representation of a canonical coordinate (Sect. 3.4), with unusual parity under time-inversion T. In Sect. 4 we recall that a 4-derivative degree of freedom $q(t)$ is described by two canonical coordinates: $q_1 = q$ and $q_2 = \dot{q}$. While q_1 is T-even as usual, q_2 is T-odd: we argue that thereby it naturally follows the negative-norm representation. The resulting quantum theory is unitary: time evolution preserves the negative norm. The path integral formulation is discussed in Sect. 5. In Sect. 6 we discuss the interacting theory, outline the extension to quantum field theory, and discuss the issue of giving a sensible interpretation to negative norms, via a postulate that generalizes the Born rule. Conclusions are given in Sect. 7.

2 The Ostrogradski classical canonical formalism

Let us now introduce the main issues in the simplest relevant case. Our final goal will be 4-derivative gravity; however, the graviton components can be Fourier expanded into modes with given momentum and four time derivatives, and at leading order in the perturbative expansion one has decoupled harmonic oscillators. So, we start considering a single mode $q(t)$, described by the Lagrangian

$$\begin{aligned}\mathscr{L} &= -\frac{\ddot{q}^2}{2} + (\omega_1^2+\omega_2^2)\frac{\dot{q}^2}{2} - \omega_1^2\omega_2^2\frac{q^2}{2} - V(q)\\ &= -\frac{1}{2}q\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}+\omega_1^2\right)\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}+\omega_2^2\right)q - V(q)\\ &\quad + \text{total derivatives}.\end{aligned}$$

where $V(q)$ is some interaction. We assume real ω_1, ω_2, because we are interested in ghosts (negative kinetic and potential energy), not in tachyonic instabilities (potential unstable with respect to the kinetic term). The $-$ sign means that the ghost is the state with larger ω; we choose $\omega_1 > \omega_2$ and do not explicitly discuss here the degenerate case $\omega_1 = \omega_2$.

Ostrogradski introduced an *auxiliary* coordinate q_2 that allows one to describe the 4-derivative oscillator in canonical Hamiltonian form (see also ref. [21] for a review of this method):

$$\begin{aligned}&q_1 = q, \quad p_1 = \frac{\delta\mathscr{L}}{\delta\dot{q}_1} = (\omega_1^2+\omega_2^2)\dot{q} + \dddot{q},\\ &q_2 = \lambda\dot{q}, \quad p_2 = \frac{\delta\mathscr{L}}{\delta\dot{q}_2} = -\frac{\ddot{q}}{\lambda},\end{aligned} \tag{3}$$

where for a generic variable x we have introduced the variational derivative

$$\frac{\delta\mathscr{L}}{\delta x} = \frac{\partial\mathscr{L}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathscr{L}}{\partial\dot{x}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial\mathscr{L}}{\partial\ddot{x}} + \cdots. \tag{4}$$

While Ostrogradski assumed $\lambda = 1$, we introduced an arbitrary constant λ. The system in Eq. (3) can be solved for q and its time derivatives,

$$\begin{aligned}&q = q_1, \quad \dot{q} = \frac{q_2}{\lambda}, \quad \ddot{q} = -\lambda p_2,\\ &\dddot{q} = p_1 - (\omega_1^2+\omega_2^2)\frac{q_2}{\lambda},\end{aligned} \tag{5}$$

and the Hamiltonian turns out to be

$$\begin{aligned}H &= \sum_{i=1}^{2} p_i\dot{q}_i - \mathscr{L} = \frac{p_1q_2}{\lambda} - \frac{\lambda^2}{2}p_2^2 - \frac{\omega_1^2+\omega_2^2}{2\lambda^2}q_2^2\\ &\quad + \frac{\omega_1^2\omega_2^2}{2}q_1^2 + V(q_1).\end{aligned} \tag{6}$$

In view of its first term, the classical Hamiltonian H has no minimal energy configuration: this is the essence of the Ostrogradski no-go classical theorem. Using the Poisson parentheses $\{\ ,\ \}$ one computes the Hamiltonian equations of motion:

$$\begin{cases} \dot{q}_1 = \{q_1, H\} = \dfrac{\partial H}{\partial p_1} = \dfrac{q_2}{\lambda}, & \dot{p}_1 = \{p_1, H\} = -\frac{\partial H}{\partial q_1} = -\omega_1^2\omega_2^2 q_1 - V'(q_1), \\ \dot{q}_2 = \{q_2, H\} = \frac{\partial H}{\partial p_2} = -\lambda^2 p_2, & \dot{p}_2 = \{p_2, H\} = -\frac{\partial H}{\partial q_2} = -\frac{p_1}{\lambda} + \frac{\omega_1^2+\omega_2^2}{\lambda^2} q_2. \end{cases} \tag{7}$$

For any λ they imply the classical Lagrangian equation of motion. Setting $V = 0$, it is

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega_1^2\right)\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega_2^2\right) q = \frac{\mathrm{d}^4 q}{\mathrm{d}t^4} + (\omega_1^2 + \omega_2^2)\frac{\mathrm{d}^2 q}{\mathrm{d}t^2} + \omega_1^2\omega_2^2 q = 0. \tag{8}$$

The corresponding classical solution, for given initial conditions $q_0, \dot{q}_0, \ddot{q}_0, \dddot{q}_0$ at $t = 0$, is

$$q(t) = -\frac{\omega_2^2 q_0 + \ddot{q}_0}{\omega_1^2 - \omega_2^2}\cos(\omega_1 t) + \frac{\omega_1^2 q_0 + \ddot{q}_0}{\omega_1^2 - \omega_2^2}\cos(\omega_2 t) - \frac{\omega_2^2 \dot{q}_0 + \dddot{q}_0}{\omega_1(\omega_1^2 - \omega_2^2)}\sin(\omega_1 t) + \frac{\omega_1^2 \dot{q}_0 + \dddot{q}_0}{\omega_2(\omega_1^2 - \omega_2^2)}\sin(\omega_2 t). \tag{9}$$

This is a well-behaved oscillator without run-away issues because the positive-energy and negative-energy components are decoupled. Run-away solutions appear when they interact through a generic interaction, such as a $V \neq 0$.

2.1 Quantizing the Ostrogradski Hamiltonian

The classical equation differs from the usual 2-derivative equation $\mathrm{d}(q + ip)/\mathrm{d}t = i\omega(q + ip)$, so that, trying to quantize the theory, we do not define the usual annihilation operator $a_i \propto q_i + ip_i$. Rather, it is convenient to define the operators a_i as the coefficients of a given frequency:

$$q(t) = a_1 \mathrm{e}^{-i\omega_1 t} + a_2 \mathrm{e}^{-i\omega_2 t} + \text{h.c.} \tag{10}$$

The a_1, a_2 can be expressed in terms of canonical Hamiltonian coordinates:

$$a_1 = \frac{\lambda\omega_1\omega_2^2 q_1 - i\omega_1^2 q_2 + ip_1\lambda - \omega_1 p_2\lambda^2}{2\lambda\omega_1(\omega_2^2 - \omega_1^2)}, \tag{11}$$

$$a_2 = \frac{\lambda\omega_1^2\omega_2 q_1 - i\omega_2^2 q_2 + ip_1\lambda - \omega_2 p_2\lambda^2}{2\lambda\omega_2(\omega_1^2 - \omega_2^2)}. \tag{12}$$

Using the canonical quantization $[q_i, p_j] = i\delta_{ij}$ one finds the commutation relations for the a_i:

$$[\tilde{a}_1, \tilde{a}_1^\dagger] = -1, \quad [\tilde{a}_2, \tilde{a}_2^\dagger] = 1, \quad \text{all other commutators vanish.} \tag{13}$$

having normalized $\tilde{a}_1 = \sqrt{2\omega_1(\omega_1^2 - \omega_2^2)}a_1$ and $\tilde{a}_2 = \sqrt{2\omega_2(\omega_1^2 - \omega_2^2)}a_2$. The Hamiltonian is

$$H = -\omega_1\frac{\tilde{a}_1\tilde{a}_1^\dagger + \tilde{a}_1^\dagger\tilde{a}_1}{2} + \omega_2\frac{\tilde{a}_2\tilde{a}_2^\dagger + \tilde{a}_2^\dagger\tilde{a}_2}{2}. \tag{14}$$

The state 1 with higher frequency $\omega_1 > \omega_2$ is a ghost.

As better discussed later in Sect. 3.3, this system can be quantized in two different ways:

1. *Positive norm, negative energy* One redefines $\tilde{a}_1' = \tilde{a}_1^\dagger$, such that it has the usual commutation $[\tilde{a}_1', \tilde{a}_1'^\dagger] = 1$. The vacuum state $|\tilde{0}\rangle$ is defined as usual by $\tilde{a}_1'|\tilde{0}\rangle = 0$ and $\tilde{a}_2|\tilde{0}\rangle = 0$. By solving this condition as a differential equation for $\psi_{\tilde{0}}(q_1, q_2) = \langle q_1, q_2|\tilde{0}\rangle$ with $p_i = -i\partial/\partial q_i$ one obtains the ground-state wavefunction:

$$\psi_{\tilde{0}}(q_1, q_2) = \exp\left(-\frac{q_1^2\omega_1\omega_2 + q_2^2/\lambda^2}{2}(\omega_1 - \omega_2) + iq_1\frac{q_2}{\lambda}\omega_1\omega_2\right). \tag{15}$$

2. *Negative norm, positive energy* The vacuum is now defined as $a_1|0\rangle = 0$ and $a_2|0\rangle = 0$. Using $p_i = -i\partial/\partial q_i$ one obtains the ground-state wavefunction

$$\psi_0(q_1, q_2) \propto \exp\left(\frac{-q_1^2\omega_1\omega_2 + q_2^2/\lambda^2}{2}(\omega_1 + \omega_2) - iq_1\frac{q_2}{\lambda}\omega_1\omega_2\right). \tag{16}$$

If $\lambda = 1$ the situation is bad, as emphasized by [16,17]: the positive-norm quantization gives a normalizable wavefunction $\psi_{\tilde{0}}$ but negative energies; the negative-norm quantization gives a ground-state wavefunction not normalizable in $q_2 = \dot{q}$. Excited states have the same problem.

However, as we will show in Sect. 4, consistency requires the negative-norm Dirac–Pauli representation of a canonical coordinate which roughly amounts to choosing an imaginary λ, e.g. $\lambda = -i$. One then obtains positive energy, negative norm, and a wavefunction $\psi_0(q_1, q_2)$ normalizable in q_1 and $q_2 = -i\dot{q}$. As we will now discuss, despite the strange i factor, $\dot{q} = iq_2$ as well as the Ostrogradski Hamiltonian $H = iq_2 p_1 + \cdots = \dot{q}\, p_1 + \cdots$ is self-adjoint, so that time evolution is unitary.

3 Quantum mechanics with negative norm

We here discuss quantum mechanics with negative norm from a general point of view. Negative-norm states require putting some minus sign here and there. It is convenient to be more general and consider a Hilbert-like space with generic, possibly negative, constant norm (called Krein space by mathematicians) and develop a basis-independent formalism. This will let us to clarify confusions, in particular about self-adjoint operators that are represented (in some basis) by non-hermitian matrices, allowing us to understand the unusual imaginary λ introduced in the previous section.

We follow the notations used in general relativity, rewriting the quantum-state metric as $\langle_n|_m\rangle = \eta_{nm}$ and defining the inverse metric $(\eta)^{nm} \equiv (\eta^{-1})_{nm}$, the contravariant ket $|^n\rangle = \eta^{nm}|_m\rangle$ such that $\langle^n|^m\rangle = \eta^{nm}$ and $\langle^n|_m\rangle = \delta^n_m = \langle_n|^m\rangle$. Summations over repeated indices are implicit. As usual, bras denote complex conjugate of kets.

A generic *state* $|\psi\rangle$ can be expanded in either the 'covariant' or the 'contravariant' basis:

$$\psi_n \equiv \langle_n|\psi\rangle, \quad \psi^n \equiv \langle^n|\psi\rangle. \tag{17}$$

Then

$$|\psi\rangle = \psi^n|_n\rangle = \psi_n|^n\rangle. \tag{18}$$

A generic *linear operator* A can be written as a matrix in 4 different ways:

$$A_{nm} \equiv \langle_n|A|_m\rangle, \quad A^{nm} \equiv \langle^n|A|^m\rangle, \quad A_n{}^m \equiv \langle_n|A|^m\rangle,$$
$$A^n{}_m \equiv \langle^n|A|_m\rangle. \tag{19}$$

Then

$$A = A^{nm}|_n\rangle\langle_m| = A_{nm}|^n\rangle\langle^m| = A_n{}^m|^n\rangle\langle_m| = A^n{}_m|_n\rangle\langle^m|. \tag{20}$$

The components of the matrices are related by $A_n{}^m = \eta_{nn'}A^{n'}{}_{m'}\eta^{m'm}$, which is an *iso-spectral* transformation: the eigenvalues do not change because the matrix A gets left-multiplied by η and right-multiplied by its inverse.

The *unity operator* is represented by $1_{nm} = \eta_{nm}$ and $1^{nm} = \eta^{nm}$ and expanded as

$$1 = \eta^{nm}|_n\rangle\langle_m| = \eta_{nm}|^n\rangle\langle^m| = |_n\rangle\langle^n| = |^n\rangle\langle_n|. \tag{21}$$

Operator multiplication becomes, in components, $(AB)_{nm} = A_{nn'}\eta^{n'm'}B_{m'm}$. Expectation values are given by $\langle\psi|A|\psi\rangle/\langle\psi|\psi\rangle$.

The *adjoint* $A^\dagger$ of an operator A is defined, as usual, as the operator such that $|\psi'\rangle = A|\psi\rangle$ implies $\langle\psi'| = \langle\psi|A^\dagger$. Thereby for generic matrix elements one has $\langle\psi_2|A^\dagger|\psi_1\rangle \equiv \langle\psi_1|A|\psi_2\rangle^*$, and the relation for the components

$$(A^\dagger)_{nm} = A^*_{mn} \quad \text{i.e.} \quad (A^\dagger)^{nm} = A^{mn*}$$
$$\text{i.e.} \quad (A^\dagger)_n{}^m = (A^m{}_n)^*. \tag{22}$$

The covariant components of a self-adjoint operator A satisfy the usual condition: a self-adjoint operator is described by a hermitian matrix, A_{nm}. The same result holds for the contravariant matrix A^{nm}. The mixed components satisfy a different condition, where complex conjugation is supplemented by an iso-spectral transformation: $A_n{}^m = (\eta A^{*T}\eta^{-1})_n{}^m$.[2]

A *self-adjoint* operator, $A^\dagger = A$ has real expectation values $\langle\psi|A|\psi\rangle/\langle\psi|\psi\rangle$, although the matrix $A^m{}_n$ that represents it can be anti-hermitian.

The mixed components directly enter into the *eigenvector equation* $A|\psi\rangle = A_\psi|\psi\rangle$:

$$A_n{}^m\psi_m = A_{nm'}\eta^{m'm}\psi_m = A_\psi\psi_n \quad \text{or}$$
$$A^n{}_m\psi^m = \eta^{nn'}A_{n'm}\psi^m = A_\psi\psi^n \tag{23}$$

where A_ψ is the eigenvalue. Let us now consider a self-adjoint operator H (later it will be the Hamiltonian), with eigenstates $|E_n\rangle$ and eigenvalues E_n. The identity

$$\langle E_n|H|E_m\rangle = \langle E_n|E_m\rangle E_m = E^*_n\langle E_n|E_m\rangle \tag{24}$$

tells that H can have three different kinds of eigenstates:

+) orthogonal eigenstates $\langle E_n|E_m\rangle = 0$ with real E_n and norm $\langle E_n|E_n\rangle = +1$;
−) orthogonal eigenstates $\langle E_n|E_m\rangle = 0$ with real E_n and norm $\langle E_n|E_n\rangle = -1$;
0) pairs of complex conjugated eigenvalues, $E_n = E^*_m$ with $\langle E_n|E_m\rangle \neq 0$ and zero norm, $\langle E_n|E_n\rangle = 0$.

In the classical analog, the latter possibility corresponds to a ghost which is also a tachyon, which is a different kind of instability, to be avoided even in absence of ghosts.

It is often convenient to choose a basis of eigenstates of H: $|_n\rangle = |E_n\rangle$. The associated contravariant states $|^n\rangle$ then satisfy $H|^n\rangle = E^*_n|^n\rangle$. In this basis the space splits into two sectors: positive norm and negative norm, plus the possible pairs of zero-norm states. The two sectors experience a joint dynamics only if the initial state has a quantum entanglement among them.

3.1 Unitary time evolution

The *evolution equation* $i\,\partial_t|\psi\rangle = H|\psi\rangle$ becomes

$$i\frac{\partial}{\partial t}\psi_n = H_{nm}\eta^{mm'}\psi_{m'} \quad \text{or} \quad i\frac{\partial}{\partial t}\psi^n = \eta^{nn'}H_{n'm}\psi^m. \tag{25}$$

[2] Here T denotes the matrix transpose; $*$ denotes complex conjugation; $\dagger$ denotes the adjoint operation, which generalizes the usual hermitian conjugation and reduces to it in the positive-norm case. We never use $\dagger$ to denote hermitian conjugation of a matrix.

The norm of any state $|\psi(t)\rangle$ is conserved by time evolution if H is self-adjoint:

$$i\frac{\partial}{\partial t}\langle\psi'(t)|\psi(t)\rangle = \langle\psi'|H - H^\dagger|\psi\rangle = 0. \tag{26}$$

A self-adjoint Hamiltonian H leads to unitary time evolution. The explicit solution can be written as $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ with $U(t) = T\mathrm{e}^{-i\int H(t)\,\mathrm{d}t}$, where T is the usual time-ordering. In components,

$$\begin{aligned}&\psi_n(t) = U_n{}^m\psi_m(0) = U_{nm'}\eta^{m'm}\psi_{m'}(0) \quad \text{or}\\ &\psi^n(t) = U^n{}_m\psi^m(0) = \eta^{nn'}U_{n'm}\psi^m(0).\end{aligned} \tag{27}$$

Having written generic-metric quantum mechanics in an abstract formalism that resembles as much as possible the usual positive-norm formalism, let us now emphasize the key differences. For simplicity, let us consider a time-independent H. One can then expand $U = \mathrm{e}^{-iHt} = \sum_{n=0}^{\infty}(-iHt)^n/n!$.

- Writing U in mixed components, $U_n{}^m$ is the naive exponentiation of the matrix $H_n{}^m$. However, the mixed components of a self-adjoint H do not form a hermitian matrix. Rather, the self-adjoint condition in Eq. (22) dictates that they are hermitian up to an iso-spectral transformation.
- The covariant components of a self-adjoint H satisfy the usual Hermiticity $H^*_{nm} = H_{mn}$. However, the covariant components U_{nm} are not given by the naive matrix exponentiation of H_{nm}. Rather, extra metric factors appear to covariantize the expansion:

$$\begin{aligned}U_{nm} = \eta_{nm} &+ \eta_{nn'}(-iHt)^{n'm'}\eta_{m'm}\\ &+\frac{1}{2}\eta_{nn'}(-iHt)^{n'r'}\eta_{r's'}(-iHt)^{s'm'}\eta_{m'm} + \cdots\end{aligned} \tag{28}$$

  Correspondingly, the unitarity condition $U^\dagger U = 1$ written in covariant components is $U^*_{n'n}\eta^{n'm'}U_{m'm} = \eta_{nm}$, while in mixed components one gets the usual $U^{k*}_{\ n}U_k{}^m = \delta^n_m$.

Practical computations often employ perturbation theory, which can now easily be generalized to generic norm. Decomposing $H = H_0 + V(t)$, the interaction picture is related to the Schrödinger picture as $A_I = \mathrm{e}^{iH_0t}A\mathrm{e}^{-iH_0t}$ where A is any operator (including V). Time evolution is given by

$$\begin{aligned}U_I(t_i,t_f) = \mathrm{T}\,\mathrm{e}^{-i\int_{t_i}^{t_f}\mathrm{d}t\,V_I(t)} &= 1 - i\int_{t_i}^{t_f}\mathrm{d}t'\,V_I(t')\\ &- \int_{t_i}^{t_f}\mathrm{d}t'\int_{t_i}^{t'}\mathrm{d}t''\,V_I(t')V_I(t'') + \cdots.\end{aligned} \tag{29}$$

The above explicit form of U_I shows that the energy conserved by quantum evolution (up to the usual quantum uncertainty $\Delta t\,\Delta E \geq \hbar$) are the eigenvalues of H. Let us consider for example a time-independent interaction V and an initial state and a final state which are energy eigenstates with eigenvalues E_i and E_f. Defining $V^f{}_i = \langle f|V|_i\rangle$, at first order one has

$$\begin{aligned}|\langle^f|U|_i\rangle|^2 &\simeq \left|\int_0^t \mathrm{d}t'\,\mathrm{e}^{i(E_f-E_i)t'}V^f{}_i\right|^2\\ &= \frac{4|V^f{}_i|^2}{|E_f-E_i|^2}\sin^2\frac{(E_f-E_i)t}{2}\\ &\overset{t\to\infty}{\simeq} 2\pi t|V^f{}_i|^2\delta(E_f-E_i).\end{aligned} \tag{30}$$

This means that energy conservation reads $E_f = E_i$, up to the usual quantum uncertainty $1/(t_f - t_i)$. Higher order corrections give the usual sum over intermediate quasi-on-shell states.

3.2 Example: the indefinite-norm two-state system

Let us consider a two-state system: $|_+\rangle$ with positive unit norm, and $|_-\rangle$ with negative unit norm. Without loss of generality, by redefining the relative phase of the two states and adding a constant overall energy, one can trivially write the most generic self-adjoint Hamiltonian as

$$H = \frac{1}{2}\begin{matrix}\langle+|\\ \langle-|\end{matrix}\begin{pmatrix}E_R & -iE_I\\ iE_I & E_R\end{pmatrix}\begin{matrix}|_+\rangle & |_-\rangle\end{matrix} = \frac{1}{2}\begin{matrix}\langle+|\\ \langle-|\end{matrix}\begin{pmatrix}E_R & iE_I\\ iE_I & -E_R\end{pmatrix}\begin{matrix}|^+\rangle & |^-\rangle\end{matrix} \tag{31}$$

having used $|^\pm\rangle = \pm|_\pm\rangle$. We see that the H_{nm} components are hermitian, unlike the $H_n{}^m$ components. The eigenvalues of H are $E_\pm = \pm E$ with $E = \sqrt{E_R^2 - E_I^2}/2$. The corresponding eigenstates are

$$\begin{aligned}|_{E_+}\rangle &= \sqrt{\frac{\gamma+1}{2}}|_+\rangle - i\sqrt{\frac{\gamma-1}{2}}|_-\rangle,\\ |_{E_-}\rangle &= i\sqrt{\frac{\gamma-1}{2}}|_+\rangle + \sqrt{\frac{\gamma+1}{2}}|_-\rangle,\end{aligned} \tag{32}$$

where $\gamma = 1/\sqrt{1 - E_I^2/E_R^2}$ is a 'boost factor' that substitutes the usual mixing angle.

- If $E_I < E_R$ the eigenvalues of H are real, the orthogonal eigenvectors satisfy $\langle_{E_\pm}|_{E_\pm}\rangle = \pm 1$, and tend to get closer to the 'light-cone' of zero-norm states as E_I increases. The components of $U = \mathrm{e}^{-iHt}$ oscillate in time:

$$U = \begin{matrix}\langle+|\\ \langle-|\end{matrix}\begin{pmatrix}\cos(Et) - i\gamma\sin(Et) & \sqrt{\gamma^2-1}\sin(Et)\\ \sqrt{\gamma^2-1}\sin(Et) & \cos(Et) + i\gamma\sin(Et)\end{pmatrix}\begin{matrix}|^+\rangle & |^-\rangle\end{matrix}. \tag{33}$$

The unusual feature is that $|\langle\pm|U|\pm\rangle|^2$ oscillates between 1 and $\gamma^2 \geq 1$.

- In the critical case, $E_R = E_I$, such that $\gamma = \infty$, the two eigenstates become degenerate with energy $E = 0$. The two eigenvectors also become degenerate, and parallel to the zero-norm state $\propto |+\rangle + i|-\rangle$. The evolution operator is

$$U = \begin{matrix}\\ \langle+| \\ \langle-|\end{matrix}\begin{matrix}\begin{matrix}|^+\rangle & |^-\rangle\end{matrix}\\ \begin{pmatrix}1 - iE_Rt/2 & E_Rt/2 \\ E_Rt/2 & 1 + iE_Rt/2\end{pmatrix}\end{matrix}. \tag{34}$$

This exemplifies a more general result: zero-norm eigenstates with complex eigenvalues appear when, increasing the interaction, a level crossing between a positive-norm and a negative-norm eigenstate takes place; the Hamiltonian becomes degenerate at the critical transition.

- If the interaction E_I is strong enough, $E_I > E_R$, one has a pair of complex conjugated eigenvalues, with zero-norm eigenvectors that satisfy $\langle E_+|E_-\rangle = 1$ and describe tachyonic ghosts: their time-evolution factor $\mathrm{e}^{-iE_\pm t}$ also contains a real exponential, in analogy to tachyonic states present in positive-norm theories. In the extreme limit $E_I \gg E_R$ the eigenvalues of H are $\pm iE_I/2$, and the time evolution operator is

$$U = \begin{matrix}\\ \langle+| \\ \langle-|\end{matrix}\begin{matrix}\begin{matrix}|^+\rangle & |^-\rangle\end{matrix}\\ \begin{pmatrix}\cosh(E_It/2) & \sinh(E_It/2) \\ \sinh(E_It/2) & \cosh(E_It/2)\end{pmatrix}\end{matrix}. \tag{35}$$

This runaway happens whenever H has a pair of complex eigenvalues $E_+ = E_-^*$, as clear writing time evolution in terms of energy eigenstates, $|\psi(t)\rangle = \psi^{E_+}\mathrm{e}^{-iE_+t}|_{E_+}\rangle + \psi^{E_-}\mathrm{e}^{-iE_-t}|_{E_-}\rangle$. Both the norm of $|\psi(t)\rangle$ and the real expectation value of H are preserved by time evolution:

$$\frac{\langle\psi(t)|H|\psi(t)\rangle}{\langle\psi(t)|\psi(t)\rangle} = \frac{E_+\psi^{E_+}\psi^{E_-*} + \text{c.c.}}{\psi^{E_+}\psi^{E_-*} + \text{c.c.}}. \tag{36}$$

3.3 The negative-norm harmonic oscillator

We here study the concrete system that lies at the basis of perturbative Quantum Field Theory: the harmonic oscillator. As discussed by Lee and Wick [22,23] it admits two inequivalent quantizations: positive norm, and indefinite norm.

Let us first recall the standard oscillator, described (up to irrelevant constants) by the Hamiltonian $H = \frac{1}{2}(p^2 + q^2)$ with $[q, p] = i$. Defining

$$a = \frac{q + ip}{\sqrt{2}}, \quad a^\dagger = \frac{q - ip}{\sqrt{2}} \tag{37}$$

one has $[a, a^\dagger] = 1$ and $H = (aa^\dagger + a^\dagger a)/2$.

Let us next consider a more general system described by the following Hamiltonian and commutation relations:

$$H = s_H\frac{a^\dagger a + aa^\dagger}{2}, \quad [a, a^\dagger] = s. \tag{38}$$

For $s = s_H = 1$ this reduces to the usual oscillator. We now show that $s = s_H = -1$ defines another consistent positive-energy theory. The symbol $a^\dagger$ here indicates the adjoint of a, which generalizes the Hermitian conjugate to negative norm.

We again define the vacuum as $a|0\rangle = 0$ and the excited states as $|_n\rangle = a^\dagger|_{n-1}\rangle/\sqrt{n} = (a^\dagger)^n|0\rangle/\sqrt{n!}$. Its inverse is $a|_n\rangle = s\sqrt{n}|_{n-1}\rangle$. The state metric is $\eta_{nm} \equiv \langle_m|_n\rangle = s^n\delta_{nm}$. The norm is determined by the dynamics, and odd states have negative norm for $s = -1$. The inverse metric is $\eta^{nm} = s^{-n}\delta_{nm}$ and the contravariant states are $|^n\rangle = s^{-n}|_n\rangle$. In components one has

$$a = \begin{matrix}\\ \langle0| \\ \langle1| \\ \langle2| \\ \langle3| \\ \vdots\end{matrix}\begin{matrix}\begin{matrix}|0\rangle & |1\rangle & |2\rangle & |3\rangle & \cdots\end{matrix}\\ \begin{pmatrix}0 & s & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2}s^2 & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{3}s^3 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots\end{pmatrix}\end{matrix}$$
$$= \begin{matrix}\\ \langle0| \\ \langle1| \\ \langle2| \\ \langle3| \\ \vdots\end{matrix}\begin{matrix}\begin{matrix}|0\rangle & |^1\rangle & |^2\rangle & |^3\rangle & \cdots\end{matrix}\\ \begin{pmatrix}0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{3} & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots\end{pmatrix}\end{matrix} \tag{39}$$

$$a^\dagger = \begin{matrix}\\ \langle0| \\ \langle1| \\ \langle2| \\ \langle3| \\ \vdots\end{matrix}\begin{matrix}\begin{matrix}|0\rangle & |1\rangle & |2\rangle & |3\rangle & \cdots\end{matrix}\\ \begin{pmatrix}0 & 0 & 0 & 0 & \cdots \\ s & 0 & 0 & 0 & \cdots \\ 0 & \sqrt{2}s^2 & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{3}s^3 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots\end{pmatrix}\end{matrix}$$
$$= s\begin{matrix}\\ \langle0| \\ \langle1| \\ \langle2| \\ \langle3| \\ \vdots\end{matrix}\begin{matrix}\begin{matrix}|0\rangle & |^1\rangle & |^2\rangle & |^3\rangle & \cdots\end{matrix}\\ \begin{pmatrix}0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots\end{pmatrix}\end{matrix} \tag{40}$$

In components the commutation relations read

$$[a, a^\dagger]_{nm} = (a\cdot\eta\cdot a^\dagger - a^\dagger\cdot\eta\cdot a)_{nm} = s^{n+1}\delta_{nm}$$
$$= s\eta_{nm} \quad \text{i.e.} \quad [a, a^\dagger] = s\sum_n |_n\rangle\langle^n| = s1 \tag{41}$$

and the Hamiltonian is

$$H_{nm} = \left(n + \frac{1}{2}\right) s_H s^{n+1} \delta_{nm} = E_n \eta_{nm}$$

$$\text{i.e.} \quad H = \sum_{n=0}^{\infty} E_n |_n\rangle\langle^n| \tag{42}$$

where $E_n = (n + \frac{1}{2})ss_H$ are the Hamiltonian eigenvalues, $H|_n\rangle = E_n|_n\rangle$. We see that positive-energy eigenvalues are obtained for $s = s_H = 1$ (the usual case with positive H and positive norm), but also for $s = s_H = -1$ (negative H and negative norm).

Concerning the negative-norm case, $s = -1$, notice that the harmonic oscillator does not predict tachyonic ghosts with zero norm. Furthermore the matrix elements $a_n{}^m$ are not the hermitian conjugates of $(a^\dagger)_n{}^m$, such that the operators $q = (a + a^\dagger)/\sqrt{2}$ and $p = i(a^\dagger - a)/\sqrt{2}$ are represented by matrices $q_n{}^m$ and $p_n{}^m$ which are not Hermitian. This is why various authors who look at these matrices improperly speak of 'anti-Hermitian' operators. Nevertheless, q and p are self-adjoint operators. We will now find their coordinate representation.

3.4 The negative-norm coordinate representation

Starting from the harmonic oscillator, we now describe a more general representation of a pair of canonical coordinate variables q, p. Parity flips $q \to -q$ and $p \to -p$. In the harmonic oscillator case, this means $a \to -a$ and $a^\dagger \to -a^\dagger$: so eigenstates $|_n\rangle$ with even (odd) n are even (odd) under parity. In the negative-norm quantization, states with odd n also have negative norm. Going to the coordinate wavefunction representation (we use the notation x for the coordinate, which later will become field space), this means that the norm is

$$\begin{aligned}\langle\psi'|\psi\rangle &= \int \mathrm{d}x\, [\psi'^*_{\text{even}}(x)\psi_{\text{even}}(x) - \psi'^*_{\text{odd}}(x)\psi_{\text{odd}}(x)] \\ &= \int \mathrm{d}x\, \psi'^*(x)\psi(-x).\end{aligned} \tag{43}$$

The corresponding unit operator is $1 = \int \mathrm{d}x|-x\rangle\langle x|$. Switching to the formalism appropriate for generic norm, one has the norm $\langle_{x'}|_x\rangle = \delta(x + x')$. Thereby the contravariant state is $|^x\rangle = |_{-x}\rangle$ and it satisfies the usual $\langle^{x'}|_x\rangle = \delta(x - x')$. As already discussed in the text surrounding Eq. (17), a state can be expanded as $|\psi\rangle = \int \mathrm{d}x\, \psi(_x)|^x\rangle = \int \mathrm{d}x\, \psi(^x)|_x\rangle$ with $\psi(_x) = \langle_x|\psi\rangle$ and $\psi(^x) = \langle^x|\psi\rangle = \psi(_{-x})$.

What is emerging from the harmonic oscillator computation is a more general structure: a coordinate space representation of a pair q, p of conjugated canonical variables that differs from the usual positive-norm representation

$$q|x\rangle = x|x\rangle, \quad p|x\rangle = +i\frac{\mathrm{d}}{\mathrm{d}x}|x\rangle, \tag{44}$$

which implies $\langle x|p|\psi\rangle = (-i\,\mathrm{d}/\mathrm{d}x)\psi(x)$ so that it satisfies $\langle x|[q, p]|\psi\rangle = i\langle x|\psi\rangle$.

The negative-norm coordinate representation, originally discussed by Dirac [18] and Pauli [19], is

$$\boxed{q|_x\rangle = ix|_x\rangle, \quad p|_x\rangle = +\frac{\mathrm{d}}{\mathrm{d}x}|_x\rangle.} \tag{45}$$

Although q looks anti-hermitian, taking into account the extra i as well as the negative norm, these unusual features combine to form a self-adjoint q:

$$\begin{aligned}\langle_{x'}|q^\dagger|_x\rangle &= \langle_x|q|_{x'}\rangle^* = [ix'\delta(x + x')]^* \\ &= ix\delta(x + x') = \langle_{x'}|q|_x\rangle.\end{aligned} \tag{46}$$

This means that $\langle\psi|q|\psi\rangle = \int \mathrm{d}q\ \psi^*(-q)iq\,\psi(q)$ is real. A similar result holds for p. When acting on wavefunctions one has $\langle_x|q|\psi\rangle = -ix\psi(_x)$ and $\langle_x|p|\psi\rangle = (+\mathrm{d}/\mathrm{d}x)\psi(_x)$, giving the desired $[q, p] = i$ commutator. Defining momentum eigenstates as $p|_p\rangle = ip|_p\rangle$ one finds $\langle_q|_p\rangle = \mathrm{e}^{ipq}/\sqrt{2\pi}$, $\langle_{p'}|_p\rangle = \delta(p + p')$. The operator q acts as $\langle_q|q|_p\rangle = (-\mathrm{d}/\mathrm{d}p)\langle_q|_p\rangle$. One can again define $|^p\rangle = |_{-p}\rangle$ such that $1 = \int \mathrm{d}p\,|^p\rangle\langle_p|$.

The i factor that differentiates the usual representation from the Dirac–Pauli representation has an impact on the time-inversion parity. As usual, a positive-energy spectrum requires that the time inversion symmetry is anti-unitary. Then, in the Dirac–Pauli quantization q is naturally T-odd and p is naturally T-even (while the opposite holds in the usual quantization, unless T is defined adding ad hoc extra signs). This will play a key role in Sect. 4.

We are now ready to come back to the harmonic oscillator. Inserting into the condition $\langle x|a|0\rangle = 0$ the standard positive-norm representation such that $a = (q + ip)/\sqrt{2} = (x + s\,\mathrm{d}/\mathrm{d}x)/\sqrt{2}$ gives a differential equation which implies the ground-state wavefunction $\psi_0(x) \propto \mathrm{e}^{-sx^2/2}$. This is normalizable for $s = 1$ (positive norm) and non-normalizable for $s = -1$, where s was defined in Eq. (38). This problem was emphasized e.g. by Woodard [16,17] who thereby dismissed the negative-norm quantization as purely formal. However, the problem arises because the positive-norm representation of q, p was used together with the negative-norm oscillator: the problem is just a manifestation of the inconsistency of the assumptions. Consistency requires that the negative-norm harmonic oscillator must be accompanied by the negative-norm Dirac–Pauli coordinate space representation of the self-adjoint q, p operators, Eq. (45). Then the condition $\langle x|a|0\rangle = 0$ leads to a normalizable wavefunction for the ground state $\psi_0 \propto \mathrm{e}^{-x^2/2}$, as well as for the excited states. The Dirac–Pauli choice thereby provides a self-consistent description of the negative-norm oscillator. Furthermore, as discussed in the next section, in the 4-derivative case the Dirac–Pauli representation is required by simple considerations.

Table 1 Coordinate representations of a pair of canonical variables $[q,\, p] = i$, and the associated ground-state wavefunctions for the positive-energy harmonic oscillator

Norm	$\langle x\|q\|\psi\rangle$	T-parity	$\langle x\|p\|\psi\rangle$	T-parity	Harmonic oscillator with $E > 0$
Positive	$x\psi(x)$	Even	$-i\, d\psi/dx$	Odd	$\psi_0(q) \propto \mathrm{e}^{-q^2/2}$ and $H = +\frac{1}{2}(q^2 + p^2)$
Indefinite	$-ix\psi(x)$	Odd	$d\psi/dx$	Even	$\psi_0(q) \propto \mathrm{e}^{-q^2/2}$ and $H = -\frac{1}{2}(q^2 + p^2)$

4 For four derivatives Dirac–Pauli is desirable

As discussed in the previous section, and as summarized in Table 1, quantum mechanics has two faces: a canonical coordinate can be represented

(i) in the usual way with positive norm;
(ii) in the Dirac–Pauli way, with negative norm, Eq. (45).

As we now show, for theories with four derivatives the latter quantization choice (which, in the gravitational case, corresponds to a renormalizable theory with positive energy) is desirable.

A single 4-derivative real coordinate $q(t)$ contains two degrees of freedom. In the Ostrogradski procedure (Sect. 2) one rewrites the theory as a Hamiltonian system of two canonical coordinates, $q_1 = q$ and $q_2 = \lambda\dot{q}$. The key new feature that arises in 4-derivative theories is that $\dot{q}$ becomes an extra canonical coordinate. In the classical theory q_2 is just an auxiliary variable, and λ is an irrelevant constant: Ostrogradski used $\lambda = 1$.

In the quantum theory, q_1 and q_2 are operators that allow one to define the basis $|q_1, q_2\rangle$. We now show that the usual quantization must be used for q_1 and that the Dirac–Pauli quantization must be chosen for q_2, which is equivalent to (and more transparent than) fixing an imaginary λ and using the canonical representation.

As usual, the operator $q_1 = q$ is invariant under time-reversal $t \to -t$, and thereby it can follow the usual T-even representation. On the other hand, the operator $\dot{q}$ is T-odd, because of the time derivative: the time-inversion operator T transforms it as $T\dot{q}T^{-1} = -\dot{q}$. This is the novel key feature.

Taking into account that T is anti-unitary, one can equivalently define a usual T-even coordinate $q_2 = \lambda\dot{q}_1$ by choosing an imaginary λ.[3] However, it is simpler to forget the λ factors and just declare that the self-adjoint operator $\dot{q}$ is T-odd and thereby it follows the T-odd Dirac–Pauli representation. Then the Ostrogradski Hamiltonian of Eq. (6) is T-even. The states satisfy $T|q, \dot{q}\rangle = |q, \dot{q}\rangle$ since $\dot{q}$ has imaginary eigenvalues and since T is anti-unitary.

The strange extra factor of i has been justified from first principles. With hindsight, it was not so strange. After all, it is well known that the self-adjoint spatial gradient is $i\vec{\nabla}$ rather than $\vec{\nabla}$. In a relativistic theory, one could have guessed that similarly the self-adjoint time derivative is $i\partial/\partial t$ rather than $\partial/\partial t$. Loosely speaking, while from a classical perspec-

[3] Alternative routes lead to the same conclusion. For example, one can use the T-even $\ddot{q}$ (instead of $\dot{q}$) as second canonical coordinate. In the Ostrogradski formalism $\ddot{q} = -p_2$ is a momentum. So again one gets a canonical coordinate with unusual T-parity (normally a momentum is T-odd). In general, the invariance of the commutation relation $[q_2,\, p_2] = i$ under the anti-unitary time-inversion implies that q_2 and p_2 have opposite T-parities. One can switch $q_2 \leftrightarrow p_2$ in order to restore their usual T-parities: but their commutator changes sign, implying again negative-norm quantization. This is indeed what happens in the auxiliary variable formalism, used in various forms in the literature

Footnote 3 continued
as an alternative to the canonical Ostrogradski formalism (see e.g. [28,29,32,49]). This formalism is convenient when dealing with quantum field theory instead of quantum mechanics with a finite number of degrees of freedom. In order to facilitate the contact, we summarize the auxiliary variable formalism below. Restarting from the Lagrangian in Eq. (3), we add zero as a perfect square containing an auxiliary variable a:

$$\mathscr{L} = \mathscr{L} + \frac{1}{2}\left[\ddot{q} + (\omega_1^2 + \omega_2^2)\frac{q}{2} - \frac{a}{2}\right]^2. \tag{47}$$

Expanding the square cancels both the second-order and the fourth-order kinetic terms for q leaving

$$\mathscr{L} = -\frac{a\ddot{q}}{2} + (\omega_1^2 - \omega_2^2)^2\frac{q^2}{8} - (\omega_1^2 + \omega_2^2)\frac{aq}{4} + \frac{a^2}{8} - V(q). \tag{48}$$

Going to the free theory $V = 0$, we can diagonalize the kinetic and mass term through the field redefinition

$$\begin{cases} a = \sqrt{\omega_1^2 - \omega_2^2}(\tilde{q}_2 - \tilde{q}_1) \\ q = (\tilde{q}_2 + \tilde{q}_1)/\sqrt{\omega_1^2 - \omega_2^2} \end{cases} \quad \text{i.e.}$$
$$\tilde{q}_{1,2} = \frac{q}{2}\sqrt{\omega_1^2 - \omega_2^2} \mp \frac{a}{2\sqrt{\omega_1^2 - \omega_2^2}} \tag{49}$$

obtaining two decoupled oscillators,

$$\mathscr{L} = \frac{\dot{\tilde{q}}_2^2 - \omega_2^2\tilde{q}_2^2}{2} - \frac{\dot{\tilde{q}}_1^2 - \omega_1^2\tilde{q}_1^2}{2}. \tag{50}$$

From its classical solution, $a = 2\ddot{q} + (\omega_1^2 + \omega_2^2)q$, we see that a roughly corresponds to the Ostrogradski p_2. Furthermore, inserting such classical solution in Eq. (49) one recovers the formalism used in [28,29]

$$\tilde{q}_2 = \frac{\ddot{q} + \omega_1^2 q}{\sqrt{\omega_1^2 - \omega_2^2}}, \quad \tilde{q}_1 = -\frac{\ddot{q} + \omega_2^2 q}{\sqrt{\omega_1^2 - \omega_2^2}}. \tag{51}$$

tive $\dot{q}$ was the natural auxiliary variable, from a quantum perspective the natural extra coordinate operator is $i\dot{q}$.[4]

Using the Heisenberg representation, one has $q(t) = U^\dagger(t)q(0)U(t)$ and $\dot{q} = -i[q, H] = U^\dagger(t)\dot{q}(0)U(t)$ with unitary U, so $q(t)$ keeps real eigenvalues and $\dot{q}(t)$ keeps imaginary eigenvalues at any t (these statements are not contradictory, given that $q(t)$ also depends on $p_1(0)$ and $p_2(0)$).

4.1 The frequency eigenstates

We conclude this section by computing what the Dirac–Pauli representation adopted for $q_2 = \dot{q}$ implies for the frequency eigenstates. We restart from the Hamiltonian Eq. (6) and bring it in diagonal form,

$$H = -\frac{1}{2}\left(\tilde{p}_1^2\tilde{\lambda}^2 + \omega_1^2\frac{\tilde{q}_1^2}{\tilde{\lambda}^2}\right) + \frac{1}{2}(\tilde{p}_2^2 + \omega_2^2\tilde{q}_2^2) \tag{52}$$

through the canonical transformation

$$q_1 = \frac{\tilde{q}_2 - \tilde{\lambda}\tilde{p}_1/\omega_1}{\sqrt{\omega_1^2 - \omega_2^2}}, \quad \frac{q_2}{\lambda} = \frac{\tilde{p}_2 - \omega_1\tilde{q}_1/\tilde{\lambda}}{\sqrt{\omega_1^2 - \omega_2^2}},$$
$$p_1 = \omega_1\frac{\omega_1\tilde{p}_2 - \omega_2^2\tilde{q}_1/\tilde{\lambda}}{\sqrt{\omega_1^2 - \omega_2^2}}, \quad p_2\lambda = \frac{\omega_2^2\tilde{q}_2 - \omega_1\tilde{\lambda}\tilde{p}_1}{\sqrt{\omega_1^2 - \omega_2^2}}, \tag{53}$$

which satisfies $q_1p_1 - q_2p_2 = \tilde{p}_2\tilde{q}_2 - \tilde{p}_1\tilde{q}_1$. For the sake of generality, we here allow for generic factors λ and $\tilde{\lambda}$. The non-vanishing commutators, $[\tilde{q}_1, \tilde{p}_1] = i$ and $[\tilde{q}_2, \tilde{p}_2] = i$, can be rewritten in terms of $\tilde{a}_2 = \sqrt{\omega_2/2}(\tilde{q}_2 + i\,\tilde{p}_2/\omega_2)$ and of $\tilde{a}_1 = \sqrt{\omega_1/2}(\tilde{q}_1/\tilde{\lambda} - i\tilde{\lambda}\tilde{p}_1/\omega_1)$ reproducing the Hamiltonian of Eq. (14) and the commutators of Eq. (13). The ground-state wave function is easily computed imposing $\langle\tilde{q}_1, \tilde{q}_2|\tilde{a}_{1,2}|0\rangle = 0$ finding

$$\psi_0(\tilde{q}_1, \tilde{q}_2) \propto \exp\left[-\omega_2\frac{\tilde{q}_2^2}{2} + \omega_1\frac{\tilde{q}_1^2}{2\tilde{\lambda}^2}\right]. \tag{54}$$

For $\tilde{\lambda} = 1$ it is not normalizable [43,44]. It is normalizable if instead $|\mathrm{Im}\,\tilde{\lambda}| > |\mathrm{Re}\,\tilde{\lambda}|$.

The Dirac–Pauli representation for q_2, p_2 corresponds to imaginary λ. Imposing that q_2, p_1 are T-odd and that q_1, p_2 are T-even (i.e. that q_2 and p_2 have the unusual T parity) implies that $\tilde{q}_1$, $\tilde{p}_2$ are T-odd and that $\tilde{q}_2$, $\tilde{p}_1$ are T-even (i.e. that the canonical coordinates of the negative-norm mode $\tilde{q}_1$ and $\tilde{p}_1$ have the unusual T parity). This is obtained for imaginary $\tilde{\lambda}$.

As a check, let us connect the q_1, q_2 basis with the $\tilde{q}_1, \tilde{q}_2$ basis for generic λ and $\tilde{\lambda}$. It is convenient to start from the T-odd basis $\tilde{q}_1$, $\tilde{p}_2$, in which the ground-state wavefunction is

$$\psi_0(\tilde{q}_1, \tilde{p}_2) \propto \exp\left[-\frac{\tilde{p}_2^2}{2\omega_2} + \omega_1\frac{\tilde{q}_1^2}{2\tilde{\lambda}^2}\right]. \tag{55}$$

Next, the transition to the T-odd variables p_1, q_2 is simply

$$\langle p_1, q_2|\tilde{q}_1, \tilde{p}_2\rangle \propto \delta\left(\frac{\tilde{q}_1}{\tilde{\lambda}} - \frac{p_1 - q_2\omega_1^2/\lambda}{\omega_1\sqrt{\omega_1^2 - \omega_2^2}}\right) \times \delta\left(\tilde{p}_2 - \frac{p_1 - q_2\omega_2^2/\lambda}{\sqrt{\omega_1^2 - \omega_2^2}}\right). \tag{56}$$

Inserting the change of variables dictated by the δ functions into $\psi_0(\tilde{q}_1, \tilde{p}_2)$ one obtains

$$\psi_0(p_1, q_2) \propto \exp\left[-\frac{p_1^2 + 2\omega_1\omega_2 p_1 q_2/\lambda - \omega_1\omega_2(\omega_1^2 + \omega_1\omega_2 + \omega_2^2)(q_2/\lambda)^2}{2\omega_1\omega_2(\omega_1 + \omega_2)}\right] \tag{57}$$

where q_2 and p_1 are both complex and linked by $\mathrm{Re}\,p_1 = \omega_1^2\mathrm{Re}\,(q_2/\lambda)$ and $\mathrm{Im}\,p_1 = -\omega_2^2\mathrm{Im}\,(q_2/\lambda)$. ψ_0 can be trivially analytically continued to real p_1, q_2. For $\lambda = \pm i$ it remains a bounded Gaussian. Finally, one performs the Fourier transform from p_1 to q_1, obtaining from $\psi_0(p_1, q_2)$ the ground-state wavefunction $\psi_0(q_1, q_2)$, which agrees with Eq. (16). The same equality holds for excited states, which can be computed acting with creation operators on the ground state.

In the limit $\omega_1 = \omega_2$ one gets the critical situation described in Sect. 3.2.

5 Path-integral quantization

We now present the path-integral quantization of the same 4-derivative theory.

5.1 Path integral for generic norm

Our generic-norm formalism makes easy to write down the equivalent path integral formalism, an issue already considered in [27]. Inserting $1 = \int \mathrm{d}q\,|^q\rangle\langle_q|$ at intermediate times $t_m = t_i + m\,\mathrm{d}t$ one has

$$\langle^{q_f,t_f}|_{q_i,t_i}\rangle = \prod_m \int \mathrm{d}q_m\,\langle^{q_{m+1},t_{m+1}}|_{q_m,t_m}\rangle. \tag{58}$$

[4] The possibility of converting a non-normalizable wavefunction $\psi_0 \propto \mathrm{e}^{z^2/2}$ into a normalizable one by restricting $z = x + iy$ to the imaginary axis, rather than along the real axis, was presented as an ad hoc recipe to get something sensible in earlier works by Bender and Mannheim [33–42]. At the technical level, their approach differs from ours because they added an i factor to the variable q, rather than to $\dot{q}$. Our approach follows from general considerations, and it has the advantage that q, p and thereby the Hamiltonian are self-adjoint (although their matrix representations look anti-hermitian), such that the generalization to an interacting theory will be immediate (Sect. 6.1).

Each step $\langle^{q_{m+1},t_{m+1}}|_{q_m,t_m}\rangle$ can be evaluated as

$$\langle^{q_{m+1}}|\mathrm{e}^{-i\,H\,\mathrm{d}t}|_{q_m}\rangle = \int \mathrm{d}p_m\ \langle^{q_{m+1}}|^{p_m}\rangle\langle_{p_m}|\mathrm{e}^{-i\,H\,\mathrm{d}t}|_{q_m}\rangle = \int \frac{\mathrm{d}p_m}{2\pi}\mathrm{e}^{i[p_m(q_{m+1}-q_m)-H_{\mathrm{cl}}\mathrm{d}t]}, \tag{59}$$

having defined

$$H_{\mathrm{cl}} \equiv \frac{\langle p|H|q\rangle}{\langle p|q\rangle} \tag{60}$$

and using $\langle^q|^p\rangle = \mathrm{e}^{ipq}/\sqrt{2\pi}$ and $\langle_p|_q\rangle = \mathrm{e}^{-ipq}/\sqrt{2\pi}$. The final result is the path integral

$$\langle^{q_f,t_f}|_{q_i,t_i}\rangle = \int Dq\ Dp\ \mathrm{e}^{i\int \mathrm{d}t[p\dot{q}-H_{\mathrm{cl}}]}$$
$$\text{where } Dq\ Dp = \lim_{\mathrm{d}t\to 0}\prod_m \frac{\mathrm{d}q_m\mathrm{d}p_m}{2\pi} \tag{61}$$

and with boundary conditions $q(t_i) = q_i, q(t_f) = q_f$.

5.2 Path integral for 4-derivative quantum theories

Applying the generic path integral of Eq. (61) to the 4-derivative oscillator in the canonical Ostrogradski formalism, one gets the transition amplitude

$$\langle^{q_{1f},q_{2f},t_f}|_{q_{1i},q_{2i},t_i}\rangle \propto \int Dq_1 Dp_1 Dq_2 Dp_2 \times \exp\left[i\int \mathrm{d}t\,[p_1\dot{q}_1 + p_2\dot{q}_2 - H_{\mathrm{cl}} + J_1q_1 + J_2q_2]\right] \tag{62}$$

where for generality we added currents $J_{1,2}$ (such that acting with functional derivatives with respect to them one can form more general matrix elements of time-ordered operators; J_1 is T-even and J_2 is T-odd). The Dirac–Pauli representation for $\dot{q}$ manifests in two ways:

1. A propagator with an unusual $-$ in its external state.

Rewriting the transition amplitude in the usual positive-norm formalism, it acquires an usual $-$ sign, becoming $\langle q_f, -\dot{q}_f, t_f|q_i,\dot{q}_i,t_i\rangle$. In the limit $t_f \to t_i$ one has $\langle^{q_f,\dot{q}_f}|_{q_i,\dot{q}_i}\rangle = \delta(q_f-q_i)\delta(\dot{q}_f-\dot{q}_i)$, so that the unusual $-$ sign is equivalent to the Dirac–Pauli negative norm.[5] Furthermore, the T-odd nature of $\dot{q}$ is hardwired in the path integral, as a geometrical feature. For each trajectory $q(t)$ with boundary conditions $q(t_{i,f}) = q_{i,f}$ and $\dot{q}(t_{i,f}) = \dot{q}_{i,f}$ the time-inverted trajectory has the same action and the following boundary conditions:

$$q_i \to q_f,\quad q_f \to q_i,\quad \dot{q}_i \to -\dot{q}_f,\quad \dot{q}_f \to -\dot{q}_i. \tag{64}$$

Thereby the propagator given by the path integral satisfies the identity

$$\langle q_f, -\dot{q}_f, t_f|q_i,\dot{q}_i,t_i\rangle = \langle q_i,\dot{q}_i,t_f|q_f,-\dot{q}_f,t_i\rangle \tag{65}$$

which is equivalent to the operator identity $\langle\psi_f|\psi_i\rangle = \langle T\psi_i|T\psi_f\rangle$ given that $T|q,\dot{q},t\rangle = |q,\dot{q},-t\rangle$.

2. An unusual classical Hamiltonian.

Inserting the Ostrogradski Hamiltonian of Eq. (6) in the generic path integral of Eq. (61) one gets the following classical Hamiltonian:

$$H_{\mathrm{cl}} = \frac{\langle p_1,p_2|H|q_1,q_2\rangle}{\langle p_1,p_2|q_1,q_2\rangle} = ip_1q_2 + \frac{p_2^2}{2} + \frac{\omega_1^2+\omega_2^2}{2}q_2^2 + \frac{\omega_1^2\omega_2^2}{2}q_1^2 + V(q_1). \tag{66}$$

This is the same as Eq. (6) with $\lambda = -i$. H_{cl} can be complex because q_2, p_2, in the Dirac–Pauli representation, have complex eigenvalues.[6] Thanks to the unusual i, it is invariant under time-reversal.

Let us now try to evaluate the path integral. As usual, one can perform the Gaussian Dp_1Dp_2 path integrals. The Dp_1 path integral formally gives the Dirac delta function $\delta(q_2-\lambda\dot{q}_1)$, allowing to eliminate the Dq_2 path integral, leaving

$$\langle^{q_f,\dot{q}_f,t_f}|_{q_i,\dot{q}_i,t_i}\rangle \propto \int Dq\ \exp\left[i\int \mathrm{d}t\,[\mathscr{L}(q) + J_1q + J_2\lambda\dot{q}]\right], \tag{67}$$

where $\mathscr{L}$ coincides with the original 4-derivative Lagrangian. By partial integration, the source term for $\dot{q}$ can be transformed into a source for q or for $\ddot{q}$ (like in the auxiliary-field method). This computation however has three problems:

1. The Dp_1 path integral is, in general, divergent. Thereby the subsequent result is only formal.
2. The $\delta(q_2-\lambda\dot{q}_1)$ always vanishes if q_1 and q_2 are real. Thereby the Dq_2 path integral is only formal.
3. Once interactions are turned on, the Lagrangian admits classical runaway solutions, reflected in the path integral.

Given that the theory is well defined in the operator formalism, somehow this path integral must have a sense.

[5] In the limit $\mathrm{d}t = t_f - t_i \to 0$ the classical action becomes

$$S_{\mathrm{cl}} \simeq \frac{6}{\mathrm{d}t^3}\left(q_f - q_i - \mathrm{d}t\,\frac{\dot{q}_i+\dot{q}_f}{2}\right)^2 + \frac{(\dot{q}_f-\dot{q}_i)^2}{2\,\mathrm{d}t} + \cdots \tag{63}$$

which is minimal for a motion with constant speed $(\dot{q}_i + \dot{q}_f)/2$. The classical action satisfies $S(q_f,\dot{q}_f,t_f;q_i,\dot{q}_i,t_i) = -S(q_i,\dot{q}_i,t_i;q_f,\dot{q}_f,t_f)$ as well as $S_{\mathrm{cl}}(q_f,\dot{q}_f,t_f;q_i,\dot{q}_i,t_i) = -S_{\mathrm{cl}}(q_f,-\dot{q}_f,t_i;q_f,-\dot{q}_i,t_i)$.

[6] The classical Hamiltonian is real if one instead uses the equivalent oscillator basis $\tilde{q}_1,\tilde{q}_2,\tilde{p}_1,\tilde{p}_2$ of Eq. (53).

5.3 Euclidean path integral for 4-derivative quantum theories

A sensible path integral is found by starting from Eq. (62) and continuing it to Euclidean time, $it = t_E$, such that $\mathrm{d}q/\mathrm{d}t = i\,\mathrm{d}q/\mathrm{d}t_E$ i.e. $\dot{q} = iq'$. One gets the Euclidean path integral

$$\langle q_{1f}, q_{2f}, t_{Ef} | q_{1i}, q_{2i}, t_{Ei}\rangle \propto \int Dq_1 Dq_2 Dp_1 Dp_2 \times \exp\left[\int \mathrm{d}t_E (ip_1 q_1' + ip_2 q_2' - H_{\rm cl} + J_1 q_1 + J_2 q_2)\right]. \tag{68}$$

Now the Dp_1 integral is convergent and gives $\delta(q_2 - q_1')$, such that the Dq_2 path integral just fixes $q_2 = q_1'$. Next, the remaining terms in $H_{\rm cl}$ are a sum of positive squares so all other integrals are convergent. Performing them one finds the Lagrangian Euclidean path integral:

$$\boxed{\langle q_f, q_f', t_{Ef} | q_i, q_i', t_{Ei}\rangle \propto \int Dq\, \exp\left[-\int \mathrm{d}t_E \left[\mathscr{L}_E(q) + J_1 q + J_2 q'\right]\right]} \tag{69}$$

where the classical Euclidean Lagrangian corresponding to Eq. (3) is

$$\mathscr{L}_E = \frac{1}{2}\left(\frac{\mathrm{d}^2 q}{\mathrm{d}t_E^2}\right)^2 + \frac{\omega_1^2 + \omega_2^2}{2}\left(\frac{\mathrm{d}q}{\mathrm{d}t_E}\right)^2 + \frac{\omega_1^2\omega_2^2}{2}q^2 + V(q). \tag{70}$$

Let us now check the result. The classical free solution is

$$q(t_E) = a_1 \mathrm{e}^{-\omega_1 t_E} + a_2 \mathrm{e}^{-\omega_2 t_E} + b_1 \mathrm{e}^{\omega_1 t_E} + b_2 \mathrm{e}^{\omega_2 t_E}. \tag{71}$$

It already contains runaway exponentials, characteristic of any Euclidean theory. Interactions compatible with the positivity of the action lead to an equally good path integral. By imposing the boundary conditions $q = q' = 0$ at $t_{Ei} = -\infty$ and evaluating the classical action, one finds the normalizable ground-state wave function

$$\langle q, q', t_E = 0 | 0, 0, t_E = -\infty\rangle \propto \exp\left[-\frac{q^2\omega_1\omega_2 + q'^2}{2}(\omega_1 + \omega_2) + qq'\omega_1\omega_2\right]. \tag{72}$$

This agrees with the ground-state wavefunction $\psi_0(q_1, q_2)$ in Eq. (16), which was computed in the Dirac–Pauli formalism in Minkowski space, after identifying $q = q_1$ and $q' = q_2$. In other words, $q' = \mathrm{d}q/\mathrm{d}t_E = -i\mathrm{d}q/\mathrm{d}t$ coincides with q_2, as computed for $\lambda = -i$. The novel feature introduced by four derivatives is that q' must *not* be continued into an imaginary $-i\dot{q}$ (which would give divergent wavefunctions), because it already describes the T-odd variable q_2, which contains the Dirac–Pauli i factor of Eq. (45).[7] The final result is that the Minkowskian theory is an unusual analytic continuation of the Euclidean theory.

6 Interactions, quantum field theory, probability

Summarizing, we so far considered a 4-derivative harmonic oscillator. One might think that we achieved nothing [43,44]. After all, a classical 4-derivative harmonic oscillator has no runaway problems, see Eq. (9), given that it splits into two decoupled oscillators, one with negative energy and one with positive energy. The classical trouble starts when they interact. In this section we will explain that we have achieved instead something useful in an interacting quantum field theory.

6.1 Adding interactions

The quantum formalism was so far developed for the harmonic oscillator (which corresponds to the modes of a free 4-derivative quantum field theory), finding that the quantum theory has a positive-energy spectrum and no runaway behaviors. Adding interactions, the quantum interacting inherits all these good properties, as long as interactions are perturbative and as long as the interacting Hamiltonian H remains self-adjoint.

The second issue was the main obstacle to past attempts of adding ad hoc unusual i factors in order to make the quantum free theory consistent [33–42] (normalizable wavefunctions and unitary evolution with negative norm and positive energy eigenvalues): adding extra complex factors can render interactions complex, ruining the theory [43,44].

In our approach the only extra i factor arose from a principal reason: $\dot{q}$ is a T-odd coordinate that follows the negative-norm Dirac–Pauli representation. This satisfies all the properties of quantum mechanics, as generalized to negative norms: $\dot{q}$ itself is self-adjoint, like q and $\ddot{q}$. Thereby any interaction which is a real function of them is self-adjoint. Our procedure immediately generalizes to the interacting case (in agravity [12] all interactions are dictated by general covariance).

The perturbativity assumption means that, as long as the energy spectrum of the free oscillator gets slightly distorted by interactions, the energy eigenvalues will remain real and bounded from below (strongly interacting theories could also be good; however, they seem not needed for the physical application to agravity [12]).

[7] Hawking and Hertog [28,29] found a non-normalizable Minkowskian wavefunction because they expressed Eq. (72) in terms of $\dot{q}$, which is not the canonical coordinate q_2 appropriate for 4-derivative theories. Their proposal of integrating out $\dot{q}$ in the Euclidean before performing the analytic continuation to the Minkowskian is not necessary: the Minkowskian wavefunctions are normalizable if the appropriate analytic continuation is performed.

One might worry that, even if all energy eigenvalues are positive, the theory possesses negative-norm states with $\langle\psi|H|\psi\rangle < 0$. Equation (29) shows how transition amplitudes can be computed trough perturbation theory: we see that the energy eigenvalues are the quantity that enters into the conservation of energy. Thereby a theory where all eigenvalues of H (of H_0 in the perturbative expansion) are positive is consistent. As usual, perturbative computations can be systematized in terms of the propagator. By expressing $q = q_1$ in terms of the annihilation and creation operators $a_i, a_i^\dagger$ through Eqs. (11) and (12) and using the commutation relations $[\tilde{a}_i, \tilde{a}_i^\dagger] = s_i$ we find the propagator

$$\langle 0|Tq(t)q(t')|0\rangle$$

$$= \langle 0|\theta(t-t')q(t)q(t') + \theta(t'-t)q(t')q(t)|0\rangle \tag{73a}$$

$$= \frac{1}{\omega_1^2-\omega_2^2}\sum_i \frac{s_i}{2\omega_i}[\mathrm{e}^{i\omega_i(t-t')}\theta(t'-t)$$

$$+ \mathrm{e}^{i\omega_i(t'-t)}\theta(t-t')] \tag{73b}$$

$$= i\frac{1}{\omega_1^2-\omega_2^2}\int\frac{\mathrm{d}E}{2\pi}\sum_i \frac{s_i\,\mathrm{e}^{-iE(t-t')}}{E^2-\omega_i^2+i\epsilon} \tag{73c}$$

$$= \int\frac{\mathrm{d}E}{2\pi}\frac{-i\,\mathrm{e}^{-iE(t-t')}}{(E^2-\omega_1^2+i\epsilon)(E^2-\omega_2^2+i\epsilon)}, \tag{73d}$$

where ϵ is a small positive quantity and we used $s_1 = -1$ and $s_2 = 1$ in the last step.

One might worry that, using the Heisenberg picture, operators satisfy the time evolution equation $\dot{A} = -i[A, H]$, which looks dangerously similar to the classical equation of motion, as given by Poisson parentheses, which has runaway solutions. However, the quantum solutions are equal to the classical solutions only in a free theory. In general operators are not numbers, and the difference (in particular, the Dirac–Pauli representation) manifests when non-linear interactions are present. As is well known, the Heisenberg equations are in general solved by $A(t) = U^\dagger(t)A(0)U(t)$. So, all good properties of negative-norm states found in the Schrödinger picture remain valid in the Heisenberg picture, given that they are equivalent.

6.2 Extension to quantum field theory

As is well known, a single harmonic-oscillator degree of freedom $q(t)$ is the building block for a field such as $\phi(t, x, y, z)$ or $g_{\mu\nu}(t, x, y, z)$. The expansion of a field in Fourier modes with given momentum works in the 4-derivative case similarly to the 2-derivative case. As long as, at the end, we are only interested in S-matrix elements, all the detailed structure of the quantum mechanical theory, such as the wavefunctions, gets hidden behind the commutation relations of Eq. (13), which hold separately for each mode. The usual $i\epsilon$ prescription for the field propagator dictates that amplitudes can be analytically continued from the Euclidean case. Details will be presented elsewhere.

One would like to claim that quantum field theory inherits all good properties of quantum mechanics also when negative norms are present. However, while in quantum mechanics interactions can easily satisfy the condition that avoids 'tachyonic ghosts' (namely, the interaction strength between two opposite-norm states must be smaller than their energy difference as discussed in Sect. 3.2), any interesting quantum field theory leads to situations that might violate this condition. The simplest situation where this occurs is the decay of a ghost (for example a massive spin 2 graviton at rest), which can be degenerate with a multi-particle state (for example two photons going in opposite directions with energy equal to half of the ghost mass), such that the interaction, no matter how small, can be smaller than the energy difference. Actually, the ghost is degenerate with an infinite number of similar states, such that an appropriate limit procedure is needed: in the positive-norm case, entropic considerations allow to interpret this situation as particle decay. A 4-derivative kinetic term $\Pi(p) = -(p^2 - m_1^2)(p^2 - m_2^2)$ acquires a positive imaginary part. We will explore if 'ghost decay' can be interpreted like in [32].

6.3 Ghost does not play dice?

So far we carefully avoided talking about probabilities.

The theory is unitary in a negative-norm space. Thereby the only remaining difficulty is assigning an interpretation to states that entangle positive-norm components with negative-norm components. The Copenhagen interpretation added an extra ingredient external to the deterministic formalism of quantum mechanics: the Born postulate, according to which:

> "when an observable corresponding to a self-adjoint operator A is measured in a state $|\psi\rangle$, the result is an eigenvalue A_n of A with probability
>
> $$P_n = \frac{\langle\psi|\Pi_n|\psi\rangle}{\langle\psi|\psi\rangle} \quad \text{where } \Pi_n = \frac{|n\rangle\langle n|}{\langle n|n\rangle} \tag{74}$$
>
> is the projector over the eigenstate $|n\rangle$ of A".

For positive norms, these P_n satisfy the probability rules $0 \leq P_n \leq 1$ and $\sum_n P_n = 1$; the average value of A satisfies $\langle\psi|A|\psi\rangle/\langle\psi|\psi\rangle = \sum_n A_n P_n$.

At the moment we do not have a satisfactory generalization to indefinite norm. Even worse, the Born postulate is unsatisfactory by itself, given that it describes a non-local collapse of the wavefunction [53–55]. In order to make progress,

one needs to operate close to the heart of quantum mechanics. As is well known this presents fatal risks: physicists tend to become philosophers. We conclude by listing some interpretations of quantum mechanics, equivalent to the Copenhagen interpretation, which could lead to a satisfactory indefinite-norm quantum mechanics.

1. Feynman clarified the ontological basis of the Born postulate: it agrees with experiments, so 'shut up and compute'. All experiments have so far been performed with positive-norm states. The negative-norm graviton predicted by agravity is beyond the reach of present experiments. On the one hand, this is good because it means that Einstein's general relativity is recovered at large distances; on the other hand, however, we do not have experimental guidance. Lee and Wick proposed that the interpretation issue is bypassed, given that in quantum field theory we can only observe asymptotic states, which are made of positive-norm quanta [22,23]. The Lee–Wick idea may be applied to the gravitational theory proposed by Stelle [2], as discussed in [6,56].
2. Any self-adjoint Hamiltonian H gives unitary evolution with respect to many different norms, since each energy eigenstate evolves picking just a phase. Defining ghost parity $\mathcal{G}$ to be the metrics in the special basis of energy eigenstates and $|^{\psi}\rangle = \mathcal{G}^{-1}|_{\psi}\rangle$, a possible generalization of the Born postulate to generic norm is (see also [33–42])

$$P_n = \langle^{\psi}|\Pi_n|_{\psi}\rangle \quad \text{where } \Pi_n = |^n\rangle\langle_n|. \tag{75}$$

   The example of Sect. 3.2 gets converted into normal oscillations with mixing angle $\sin^2 2\theta = E_I^2/E_R^2$. However, $\langle_{\psi}|A|_{\psi}\rangle$ is real but does not have a probabilistic interpretation, while $\langle^{\psi}|A|_{\psi}\rangle$ has a probabilistic interpretation but can be complex.
3. Various authors claim that the Born postulate is just an emergent phenomenon (somehow like friction) that follows from the fundamental deterministic equations when applied to complex systems in view of spontaneous decoherence [57–59].
4. Cramer [60,61] proposed a "transactional interpretation", claiming that EPR non-locality results from a cancellation of advanced and retarded waves, in a time-symmetric set-up (see also [62]) inspired by the analogous formulation of classical electro-dynamics proposed by Dirac and Feynman–Wheeler. The $\langle\psi'|\psi\rangle$ amplitude in the Dirac–Pauli coordinate representation supports the interpretation as being the overlap of a wave ψ moving forward in time with a wave ψ' moving backwards in time.

We plan to further investigate such issues.

7 Conclusions (so far)

We presented the quantization of 4-derivative theories, finding that a unique structure emerges. We can summarize it as follows.

Quantum mechanics has its usual visible face, where a coordinate operator q is represented as $q|x\rangle = x|x\rangle$. But quantum mechanics also has a hidden face, where $q|x\rangle = ix|x\rangle$, as first pointed out by Dirac and Pauli. Both q and p of a canonical pair $[q, p] = i$ are self-adjoint in both representations. The main difference is that the usual representation implies positive norms and q is naturally even under time reflection T, while the DP representation leads to states with indefinite norm and to a naturally T-odd q (in view of the i factor and of the fact that T is anti-unitary).

The Ostrogradski formulation of a 4-derivative degree of freedom $q(t)$ (summarized in Sect. 2) employs two canonical coordinates: $q_1 = q$ and $q_2 = \dot{q}$. For the first time we have observed that q_1, which is T-even, naturally follows the usual representation, while q_2, which is T-odd, naturally follows the Dirac–Pauli negative-norm quantization. This leads to a sensible quantum theory with positive energies and normalizable wavefunctions, as discussed in Sect. 4.

In Sect. 3 we presented a new formalism appropriate for generic-norm quantum mechanics, introducing 'covariant' $|_n\rangle$ and 'contravariant' $|^n\rangle$ basis states. This clarifies why a self-adjoint linear operator can be represented by a matrix that, in some basis, is not hermitian. A self-adjoint Hamiltonian leads to unitary time evolution, in the sense that the negative norm is preserved. Given that $q, \dot{q}, \ddot{q}, \ldots$ are self-adjoint, a Hamiltonian which is a generic real function of them is self-adjoint, leading to sensible interacting quantum theory provided that one avoids tachyons, an observation that was previously overlooked. The usual condition that the theory should be free of tachyons is generalized to negative-norm quantum mechanics.

In Sect. 5 we presented the path integral formulation of negative-norm quantum mechanics. The classical Hamiltonian becomes complex. Another new result of this paper is the proof that the normalizable wavefunctions found in the operator formalism are recovered from the path integral after performing naive manipulations over ill-defined objects and/or analytic continuations. In particular, the version of the path integral in Euclidean time $t_E = it$ is well defined, and reproduces the usual wavefunctions taking into account that $\mathrm{d}q/\mathrm{d}t_E$ already coincides with the Dirac–Pauli $\dot{q}$.

The fact that (1) our approach leads to normalizable wavefunctions and (2) these wavefunctions can also be deduced from a well-defined Euclidean path integral clearly show that the right quantization for $\dot{q}$ is the Dirac–Pauli one.

Two issues must be addressed before these results can be used to obtain a predictive renormalizable quantum theory of gravity: generalization to quantum field theory, and general-

ization of the Born probabilistic interpretation to negative norms.

Acknowledgments This work was supported by the ERC grant NEO-NAT, by the Spanish Ministry of Economy and Competitiveness under Grant FPA2012-32828, Consolider-CPAN (CSD2007-00042), the Grant SEV-2012-0249 of the "Centro de Excelencia Severo Ochoa" Programme and the Grant HEPHACOS-S2009/ESP1473 from the C.A. de Madrid. We thank all the colleagues who told us that working with ghosts is equivalent to abandoning physics, and Enrique Álvarez, José R. Espinosa, Antonio González-Arroyo, Martti Raidal, Hardi Veermae, Sergey Sibiryakov, and Guido Altarelli for useful comments and encouragement. This was the content of my last conversation with Guido—the present paper was finalized wondering what Guido would have said.

Funded by SCOAP[3].

References

1. M. Ostrogradski, Mem. Ac. St. Petersbourg **VI**, 385 (1850)
2. K.S. Stelle, Renormalization of higher derivative quantum gravity. Phys. Rev. D **16**, 953 (1977). Recently, the proof of renormalisability was repeated using modern techniques, see G. Rabuffo, thesis at Pisa University
3. E.S. Fradkin, A.A. Tseytlin, Renormalizable asymptotically free quantum theory of gravity. Nucl. Phys. B **201**, 469–491 (1982). doi:10.1016/0550-3213(82)90444-8
4. I.G. Avramidi, A.O. Barvinsky, Asymptotic freedom in higher derivative quantum gravity. Phys. Lett. B **159**, 269 (1985)
5. I.G. Avramidi, Covariant methods for the calculation of the effective action in quantum field theory and investigation of higher derivative quantum gravity. (1986). arXiv:hep-th/9510140
6. I. Antoniadis, E.T. Tomboulis, Gauge invariance and unitarity in higher derivative quantum gravity. Phys. Rev. D **33**, 2756 (1986)
7. E. Elizalde, S.D. Odintsov, A. Romeo, Renormalization group properties of higher derivative quantum gravity with matter in (4-epsilon)-dimensions. Nucl. Phys. B **462**, 315 (1995). arXiv:hep-th/9502131
8. G. Narain, R. Anishetty, Running couplings in quantum theory of gravity coupled with gauge fields. JHEP **1310**, 203 (2013). arXiv:1309.0473
9. I.L. Shapiro, A.G. Zheksenaev, Gauge dependence in higher derivative quantum gravity and the conformal anomaly problem. Phys. Lett. B **324**, 286 (1994)
10. E.V. Gorbar, I.L. Shapiro, Renormalization group and decoupling in curved space. 2. The standard model and beyond. JHEP **0306**, 004 (2003). arXiv:hep-ph/0303124
11. G. de Berredo-Peixoto, I.L. Shapiro, Higher derivative quantum gravity with Gauss–Bonnet term. Phys. Rev. D **71**, 064005 (2004). arXiv:hep-th/0412249
12. A. Salvio, A. Strumia, Agravity. JHEP **1406**, 080 (2014). arXiv:1403.4226
13. M.B. Einhorn, D.R.T. Jones, Naturalness and dimensional transmutation in classically scale-invariant gravity. JHEP **1503**, 047 (2015). arXiv:1410.8513
14. K. Kannike, G. Hütsi, L. Pizza, A. Racioppi, M. Raidal, A. Salvio, A. Strumia, Dynamically induced planck scale and inflation. JHEP **1505**, 065 (2015). arXiv:1502.01334
15. A. Farzinnia, S. Kouwn, Classically scale invariant inflation, supermassive WIMPs, and adimensional gravity. Phys. Rev. D **93**(6), 063528 (2016). doi:10.1103/PhysRevD.93.063528. arXiv:1512.05890 [hep-ph]
16. R.P. Woodard, Avoiding dark energy with $1/r$ modifications of gravity. Lect. Notes Phys. **720**, 403 (2006). arXiv:astro-ph/0601672
17. R.P. Woodard, Ostrogradsky's theorem on Hamiltonian instability. Scholarpedia **10**(8), 32243 (2015). doi:10.4249/scholarpedia.32243. arXiv:1506.02210 [hep-th]
18. P.A.M. Dirac, The physical interpretation of quantum mechanics. Proc. R. Soc. Lond. A **180**, 1 (1942)
19. W. Pauli, On Dirac's new method of field quantization. Rev. Mod. Phys. **15**, 175 (1943)
20. W. Heisenberg, Quantum theory of fields and elementary particles. Rev. Mod. Phys. **29**(3), 269 (1957)
21. A. Pais, G.E. Uhlenbeck, On field theories with nonlocalized action. Phys. Rev. **79**, 145 (1950)
22. T.D. Lee, G.C. Wick, Negative metric and the unitarity of the S matrix. Nucl. Phys. B **9**, 209 (1969)
23. T.D. Lee, G.C. Wick, Finite theory of quantum electrodynamics. Phys. Rev. D **2**, 1033 (1970)
24. R.E. Cutkosky, P.V. Landshoff, D.I. Olive, J.C. Polkinghorne, A non-analytic S matrix. Nucl. Phys. B **12**, 281 (1969)
25. S. Coleman, Acausality, in *Erice 1969, Ettore Majorana School On Subnuclear Phenomena*. (New York, 1970), pp. 282–327
26. R.P. Feynman, Negative probability, in *Quantum Implications: Essays in Honor of David Bohm*, ed. by B.J. Hiley, F.D. Peat (Routledge and Kegan Paul, London, 1987)
27. D.G. Boulware, D.J. Gross, Lee–Wick indefinite metric quantization: a functional integral approach. Nucl. Phys. B **233**, 1 (1984)
28. S.W. Hawking, T. Hertog, Living with ghosts. Phys. Rev. D **65**, 103515 (2001). arXiv:hep-th/0107088
29. See also H. Kleinert, J. Math. Phys. **27**, 12 (1986)
30. G. 't Hooft, Unitarity in the Brout–Englert–Higgs mechanism for gravity. (2007). arXiv:0708.3184
31. G. 't Hooft, A class of elementary particle models without any adjustable real parameters. Found. Phys. **41**, 1829 (2011). arXiv:1104.4543
32. B. Grinstein, D. O'Connell, M.B. Wise, Causality as an emergent macroscopic phenomenon: the Lee–Wick $O(N)$ model. Phys. Rev. D **79**, 105019 (2008). arXiv:0805.2156
33. P.D. Mannheim, Solution to the ghost problem in fourth order derivative theories. Found. Phys. **37**, 532 (2006)
34. P.D. Mannheim, A. Davidson, Dirac quantization of the Pais–Uhlenbeck fourth order oscillator. Phys. Rev. A **71**, 042110 (2004). arXiv:hep-th/0408104
35. C.M. Bender, Making sense of non-Hermitian Hamiltonians. Rep. Prog. Phys. **70**, 947 (2007). arXiv:hep-th/0703096
36. C.M. Bender, P.D. Mannheim, No-ghost theorem for the fourth-order derivative Pais–Uhlenbeck oscillator model. Phys. Rev. Lett. **100**, 110402 (2007). arXiv:0706.0207
37. C.M. Bender, P.D. Mannheim, Exactly solvable PT-symmetric Hamiltonian having no Hermitian counterpart. Phys. Rev. D **78**, 025022 (2008). arXiv:0804.4190
38. C.M. Bender, P.D. Mannheim, Giving up the ghost. J. Phys. A **41**, 304018 (2008). arXiv:0807.2607
39. P.D. Mannheim, Cosmological perturbations in conformal gravity. Phys. Rev. D **85**, 124008 (2011). arXiv:1109.4119
40. P.D. Mannheim, Comprehensive solution to the cosmological constant, zero-point energy, and quantum gravity problems. Gen. Relativ. Gravit. **43**, 703 (2009). arXiv:0909.0212

41. P.D. Mannheim, Making the case for conformal gravity. Found. Phys. **42**, 388 (2011). arXiv:1101.2186
42. P.D. Mannheim, Advancing the case for PT symmetry—the Hamiltonian is always PT symmetric. (2015). arXiv:1506.08432
43. A.V. Smilga, Ghost-free higher-derivative theory. Phys. Lett. B **632**, 433 (2005). arXiv:hep-th/0503213
44. A.V. Smilga, Comments on the dynamics of the Pais–Uhlenbeck oscillator. SIGMA **5**, 017 (2008). arXiv:0808.0139
45. K. Andrzejewski, J. Gonera, P. Maslanka, Euclidean path integral and higher-derivative theories. Prog. Theor. Phys. **125**, 247 (2009). arXiv:0904.3055
46. R. Di Criscienzo, S. Zerbini, Functional determinants in higher derivative Lagrangian theories. J. Math. Phys. **50**, 103517 (2009). arXiv:0907.4265 [Erratum-ibid. **51**, 059901 (2010)]
47. S.V. Ketov, G. Michiaki, T. Yumibayashi, Quantizing with a higher time derivative, in *Advances in Quantum Field Theory*. (InTech Publishers, 2012), pp. 49–73
48. T. Biswas, E. Gerwick, T. Koivisto, A. Mazumdar, Towards singularity and ghost free theories of gravity. Phys. Rev. Lett. **108**, 031101 (2012). arXiv:1110.5249
49. T. Chen, E.A. Lim, Stabilization of linear higher derivative gravity with constraints. JCAP **1405**, 010 (2014). arXiv:1311.3189
50. K. Andrzejewski, J. Gonera, P. Machalski, K. Bolonek-Lason, On the triviality of higher derivative theories. Phys. Lett. B **706**, 427 (2012). arXiv:1110.0672
51. J. Maldacena, Einstein gravity from conformal gravity. (2011). arXiv:1105.5632
52. H. Lu, Y. Pang, C.N. Pope, Conformal gravity and extensions of critical gravity. Phys. Rev. D **84**, 064001 (2011). arXiv:1106.4657
53. A. Einstein, B. Podolsky, N. Rosen, Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777 (1935)
54. J. Bell, On the Einstein–Podolsky–Rosen paradox. Physics **1**, 195 (1964)
55. A. Aspect, Bell's inequality test: more ideal than ever. Nature **398**, 189 (1999)
56. B. Hasslacher, E. Mottola, Phys. Lett. B **99**, 221 (1981)
57. R. Griffiths, Consistent histories and the interpretation of quantum mechanics. J. Stat. Phys **36**, 219 (1984)
58. R. Omnes, Consistent interpretation of quantum mechanics. Rev. Mod. Phys. **64**, 339 (1992)
59. M. Gell-Mann, J.B. Hartle, Classical equations for quantum systems. Phys. Rev. D **47**, 3445 (1993)
60. J.G. Cramer, An overview of the transactional interpretation. J. Theor. Phys. **27**, 227 (1988)
61. J.G. Cramer, The transactional interpretation of quantum mechanics and quantum nonlocality. (2015). arXiv:1503.00039
62. Y. Aharonov, P.G. Bergmann, J.L. Lebowitz, Time symmetry in the quantum process of measurement. Phys. Rev. **134**, B1410 (1964)