THE EUROPEAN
PHYSICAL JOURNAL C

CrossMark

Obtaining the non-relativistic quantum mechanics from quantum field theory: issues, folklores and facts

What happens to the antiparticles when you take the non-relativistic limit of QFT?

T. Padmanabhan[a]

IUCAA, Post Bag 4, Ganeshkhind, Pune 411007, India

Abstract Given the classical dynamics of a non-relativistic particle in terms of a Hamiltonian or an action, it is relatively straightforward to obtain the non-relativistic quantum mechanics (NRQM) of the system. These standard procedures, based on either the Hamiltonian or the path integral, however, do not work in the case of a relativistic particle. As a result we do not have a single-particle description of relativistic quantum mechanics (RQM). Instead, the correct approach requires a transmutation of dynamical variables from the position and momentum of a single particle to a field and its canonical momentum. Particles, *along with antiparticles*, reappear in a very nontrivial manner as the excitations of the field. The fact that one needs to adopt completely different languages to describe a relativistic and non-relativistic free particle implies that obtaining the NRQM limit of QFT is conceptually nontrivial. I examine this limit in several approaches (like, for e.g., Hamiltonian dynamics, Lagrangian and Hamiltonian path integrals, field theoretic description etc.) and identify the precise issues which arise when one attempts to obtain the NRQM from QFT in each of these approaches. The dichotomy of NRQM and QFT does not originate just from the square root in the Hamiltonian or from the demand of Lorentz invariance, as is sometimes claimed. The real difficulty has its origin in the necessary existence of antiparticles to ensure a particular notion of relativistic causality. Because of these conceptual issues, it turns out that one *cannot*, in fact, obtain some of the popular descriptions of NRQM by *any* sensible limiting procedure applied to QFT. To obtain NRQM from QFT in a seamless manner, it is necessary to work with NRQM expressed in a language closer to that of QFT. This fact has several implications, especially for the operational notion of space coordinates in quantum theory. A close examination of these issues, which arise when quantum theory is combined with special relativity, could offer insights in the context of attempts to combine quantum theory with general relativity.
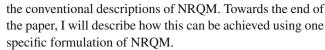
Contents

a e-mail: paddy@iucaa.in

1 Motivation and summary

Given the classical theory of a non-relativistic particle, there is a systematic way of obtaining its quantum version (NRQM), using either a Hamiltonian approach or one based on path integrals. For a system with, say, $H(\boldsymbol{x}, \boldsymbol{p}) = (\boldsymbol{p}^2/2m) + V(\boldsymbol{x})$, these approaches lead to the same quantum theory. This success, however, turns out to be more of an exception than a rule in the description of Nature. There is no guarantee that the standard (Hamiltonian or path integral) procedures of quantization will allow you to construct a quantum theory – *in terms of the same dynamical variables* – if you try to impose some extra constraints, like for e.g. Lorentz invariance,[1] general covariance, or the notion of relativistic causality, which exist in the classical theory.

An important example of a well-defined physical system, which has a simple classical description but does not have a corresponding quantum description in terms of the *same* dynamical variables is provided by a relativistic free particle. The usual procedures which work for NRQM do not work in this case. Bringing together the principles of special relativity and quantum mechanics leads to a *change in the dynamical variables*, the existence of antiparticles, and several other complications leading, eventually, to what is called Quantum Field Theory (QFT). The formalism and the language are completely different in QFT and in NRQM.

Though we have all learnt to live at peace with this development for decades, it is downright surprising when you think about it.

We do know that both QFT and NRQM work quite well in their respective domains. In the classical limit, the *equations of motion* describing a relativistic particle *does* go over to those describing a non-relativistic particle[2] when you take the limit $c \to \infty$. This suggests that, in the corresponding quantum avatars, one should be able to get NRQM from QFT by taking the limit $c \to \infty$. But if the language and even the dynamical variables used in QFT and NRQM are completely different, how can you get NRQM from QFT seamlessly? Several text books and articles deal with these issues rather too glibly (and inadequately). A large part of this paper will be devoted to pointing out that the transition from QFT to NRQM is *not* possible if your aim is to reproduce many of

the conventional descriptions of NRQM. Towards the end of the paper, I will describe how this can be achieved using one specific formulation of NRQM.

Notation: Latin indices range over $0, 1, 2, \ldots, n = D - 1$ where, usually, $D = 4$. The Greek indices range over spatial coordinates, $1, 2, \ldots, n = D - 1$. I will set $\hbar = 1, c = 1$ when it will not lead to any confusion. The signature is mostly negative. I denote by $p.x$ the on-shell dot product in which p_0 is a given function of $\boldsymbol{p}$, like e.g., $p_0 = (\boldsymbol{p}^2 + m^2)^{1/2}$, while $p_a x^a$ will denote the off-shell dot product. I will omit the superscripts in x^i, p^i etc. when it is clear from the context, like e.g., use the notation $\psi(x)$ for $\psi(x^i)$. The symbol $\equiv$ in an equation tells you that the equation is used to define some quantity.

1.1 Does the emperor have clothes?

Let me briefly describe a series of issues which arise when you try to think of NRQM as the $c \to \infty$ limit of QFT.[3] These should alert you that the situation is not as straightforward as the folklore might suggest.

(1) In NRQM, a description based on the Schroedinger wave function $\psi(x)$ (which is a *c-number complex function* in the coordinate representation) has a distinct technical advantage over the one based on the Heisenberg picture. In QFT, however, the Heisenberg picture is better suited for the description and one uses, say, a, *real*, scalar field *operator* $\hat{\phi}(x)$ which satisfies the Klein–Gordon equation. Of course, operators remain operators and real functions remain real when you take $c \to \infty$ limit; so to get $\psi(x)$ from $\hat{\phi}(x)$ one has to do something more than just taking the $c \to \infty$ limit. A favorite procedure adopted in the textbooks is the identification of $e^{-imt}\langle \boldsymbol{k}|\hat{\phi}(x)|0\rangle$ (where $|\boldsymbol{k}\rangle$ is a one-particle state with momentum $\boldsymbol{k}$) with the Schroedinger wave function. While it is trivial to show that this function, in the appropriate limit, satisfies the Schroedinger equation, this construction is rather ad hoc. More importantly, it leads to another serious issue:

What happens to antiparticles when you take the $c \to \infty$ of QFT? After all, a massive antiparticle has every right to remain at rest (or in a low-energy state) such that it could be described by NRQM! So when you take the appropriate limit of QFT you should be able to get the NRQM of *both* particles and antiparticles in a seamless manner. Many of the conventional procedures (including the one mentioned above) will not do this. At best you will get the Schroedinger equation for the particle and will have to forget about the antiparticle which, of course, is unsatisfactory.[4]

---

[1] Unless otherwise specified, I use the expression 'Lorentz invariance' to mean intrinsic Lorentz invariance and not manifest Lorentz invariance.

[2] Aside: but the *Lagrangian* for the relativistic particle does *not* go over to that of a non-relativistic particle in this limit, contrary to what some text books would like us to believe; instead, the Lagrangian blows up. If you subtract the constant mc^2 – which blows up – and redefine the Lagrangian, you lose Lorentz invariance of the action. In fact, it is not possible to construct any Lorentz invariant action for a relativistic free particle which will give the $L_{\text{NR}} = (1/2)mv^2$ in the $c \to \infty$ limit; see Chapter 15 of Ref. [1].

[3] Throughout the paper, I will only deal with a non-interacting, massive, scalar field because it is enough to illustrate the issues I am interested in. Spin and interactions add extra complications which I want to avoid so that I can highlight these issues in the simplest possible context.

[4] For example, there are (wrong) claims in the literature that the *real* scalar field in QFT has no NRQM limit because, for a real scalar field,

(2) Another issue of interpretation has to do with the very different roles played by the spatial coordinate $\boldsymbol{x}$ in QFT and NRQM. In QFT we will deal with $\hat{\phi}(t, \boldsymbol{x})$, which is an operator with both t and $\boldsymbol{x}$ acting as labels. This is necessary since Lorentz transformations will mix space and time; so if t is a label so should $\boldsymbol{x}$ be. But in NRQM the spatial coordinate itself will acquire an operator status $\hat{\boldsymbol{x}}(t)$ labeled by t. Stated in another way, the dynamical variables in NRQM are $\hat{x}^{\alpha}(t)$ and $\hat{p}_{\beta}(t)$ obeying the equal time commutation rule (ETCR), $[\hat{x}^{\alpha}(t), \hat{p}_{\beta}(t)] = i\delta_{\beta}^{\alpha}$. On the other hand, in QFT the dynamical variables are $\hat{\phi}(x)$ and $\hat{\pi}(x)$, which obey the ETCR given by $[\hat{\phi}(t, \boldsymbol{x}), \hat{\pi}(t, \boldsymbol{y})] = i\delta(\boldsymbol{x} - \boldsymbol{y})$. But there is no way of obtaining the position operator of NRQM from the basic field operators of QFT. Text books do pay homage to this fact by mumbling something about the inability to localize a particle in QFT but that does not answer the technical question of how the appropriate limit has to be taken so that you get the dynamical variables and the ETCR of NRQM from the dynamical variables and ETCR of the QFT. This, in fact, turns out to be impossible; *you cannot get there from here.* As we shall see, to make a seamless transition you need to describe NRQM in a language which is closer to that of QFT; not the other way around.

(3) Similar – and sometimes worse – difficulties arise when you approach the problem in the language of path integrals.[5] Whenever we have a well-defined classical action, we could try to quantize the system in terms of the path integral by performing the sum over all paths, connecting two events x_1 and x_2, in the expression

$$G(x_2, x_1) = \sum_{\boldsymbol{x}(t)} \exp i A[\boldsymbol{x}(t)]. \tag{1}$$

This works like a charm in NRQM. What is more, the resulting expression $G_{\mathrm{NR}}(x_2, x_1)$ has an equivalent interpretation as the matrix element of the time evolution operator:

$$\begin{aligned} G_{\mathrm{NR}}(x_2, x_1) &= \langle \boldsymbol{x}_2 | \exp[-i(t_2 - t_1)H] | \boldsymbol{x}_1 \rangle \\ &= \langle t_2, \boldsymbol{x}_2 | t_1, \boldsymbol{x}_1 \rangle. \end{aligned} \tag{2}$$

The interpretation relies on the fact that $|\boldsymbol{x}_1\rangle$ and $|\boldsymbol{x}_2\rangle$ are the eigenkets of a position operator $\hat{\boldsymbol{x}}(0)$ with eigenvalues $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$; therefore we can use $G_{\mathrm{NR}}(x_2, x_1)$ to propagate the wave function $\langle t_1, \boldsymbol{x}_1 | \psi \rangle$ to give $\langle t_2, \boldsymbol{x}_2 | \psi \rangle$. We run into several issues when we try to do any of these in an attempt to obtain a RQM.

To begin with, there are some technical issues in performing the sum in Eq. (1); most of the procedures which work well in NRQM do not work in this case. (This is because these procedures in NRQM work only if the Hamiltonian is quadratic in momentum.) There is one procedure, based on Euclidean lattice regularization, which *does* give the sensible result leading to what is usually called the Feynman propagator $G_R(x_2, x_1)$ in QFT. But the interpretation of this propagator is nontrivial because, roughly speaking, it contains information as regards both the particle and the antiparticle. Hence, it cannot be expressed in the form $G_R(x_2, x_1) = \langle t_2, \boldsymbol{x}_2 | t_1, \boldsymbol{x}_1 \rangle$; in fact, we do not have an analog of position operator $\hat{\boldsymbol{x}}(t)$ or its eigenstates, $|t, \boldsymbol{x}\rangle$ in a Lorentz invariant QFT; so one does not have an analog of Eq. (2) with the same interpretation in RQM.[6]

Thus there are serious issues in obtaining the NRQM based on position eigenstates $|t, \boldsymbol{x}\rangle$ and a wave function $\langle t, \boldsymbol{x} | \psi \rangle$ as a sensible limiting case of QFT. This conclusion remains valid irrespective of the procedure – Hamiltonian or path integral – adopted to construct the quantum theory of a relativistic particle.

1.2 Preview and summary

Let me next summarize the structure of the rest of the paper and the key results. In Sect. 2, I begin by constructing the quantum theory of a "free particle"[7] described by the Hamiltonian $H = H(|\boldsymbol{p}|)$. Since this form covers both non-relativistic and relativistic free particles, it is possible to compare the two situations at one go by studying such a system and probe why we cannot extend the standard ideas of NRQM to construct a RQM. Since a well-defined momentum operator and its eigenstates $|\boldsymbol{p}\rangle$ exist, it is possible to develop the quantum theory in momentum representation in a straightforward manner. Neither the square root structure of the Hamiltonian for a relativistic particle nor the requirement of Lorentz invariance introduces any serious difficulties in the momentum representation. But Lorentz invariance requires using a relativistically invariant normalization for momentum eigenkets (viz., $\langle \boldsymbol{p}' | \boldsymbol{p} \rangle = 2\omega_{\boldsymbol{p}} (2\pi)^n \delta(\boldsymbol{p} - \boldsymbol{p}')$ with $\omega_{\boldsymbol{p}} = (\boldsymbol{p}^2 + m^2)^{1/2}$; see Eq. (9)).

The first real difficulty arises when we try to introduce a (conjugate) position representation. In the relativistic theory we cannot introduce localized particle position states $|\boldsymbol{x}\rangle$ as

---

Footnote 4 continued

the particle is the same as the antiparticle. If antiparticles vanish mysteriously when you take the NRQM limit, then, of course, real scalar fields cannot have an NRQM limit. As we shall see, this is incorrect.

[5] Most of the intuition you develop about path integrals is based on the quadratic momentum dependence of the Hamiltonian, making this intuition pretty much useless in the study of even a relativistic free particle.

[6] There is a folklore belief that, when you take the $c \to \infty$ limit of the Feynman propagator $G_R(x_2, x_1)$, you will get the non-relativistic propagator $G_{\mathrm{NR}}(x_2, x_1)$. As I will show in Sect. 5.3, this is again *not* true without extra, ad hoc, assumptions.

[7] I will define a "free particle" to be one for which neither the Lagrangian nor the Hamiltonian depends on the spacetime coordinates explicitly and $H(\boldsymbol{p}) = H(|\boldsymbol{p}|)$. Throughout this paper we will be concerned only with a free particle. As you will see, such a system itself creates enough problems!

eigenstates of a position operator because no sensible position operator can be constructed. We can still attempt to define states $|\boldsymbol{x}\rangle$, *labeled by* spatial coordinates $\boldsymbol{x}$, as Fourier (or Fourier-like) transforms of the momentum eigenstates $|\boldsymbol{p}\rangle$, but with a relativistically invariant integration measure. This leads to a Lorentz invariant propagator for the system, given by[8]

$$G_+(x_2, x_1) \equiv \int \frac{d^n \boldsymbol{p}}{(2\pi)^n} \frac{1}{2\omega_{\boldsymbol{p}}} \exp(-ip.x), \quad x \equiv x_2 - x_1 \tag{3}$$
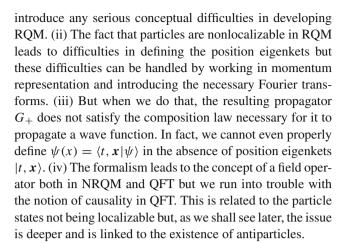
with the D-dimensional (spacetime) momentum space representation:[9]

$$G_+(p) \equiv \int d^D x\, G(x) e^{ip_a x^a} = \delta(p^2 - m^2)\theta(p^0),$$
$$D = n + 1. \tag{4}$$

But the trouble is that the states $|\boldsymbol{x}\rangle$ we have defined (and used to construct $G_+(x_2, x_1)$), do *not* represent localized particles. The amplitude $\langle \boldsymbol{x}|\boldsymbol{y}\rangle$ will not be a Dirac delta function $\delta(\boldsymbol{x} - \boldsymbol{y})$. So, even though defining $|\boldsymbol{x}\rangle$ as Fourier-like transform of $|\boldsymbol{p}\rangle$ allows us to define a Lorentz invariant propagator for the system $G_+(x_2, x_1)$, there is no way of introducing a relativistic wave function in the coordinate representation, $\psi(x) = \langle t, \boldsymbol{x}|\psi\rangle$, in the absence of position eigenstates $|t, \boldsymbol{x}\rangle$. In fact, the propagator $G_+(x_2, x_1)$ does not satisfy the correct composition law or the limiting behavior which are necessary for it to "propagate" a wave function.

So the straightforward Hamiltonian approach does not lead to an RQM such that we can obtain the NRQM as a limiting case. The utility of this discussion, for our purpose, is different. In Sect. 3, I show how the above description leads to a natural notion of (non-Hermitian) field operators *both in NRQM and RQM*. Here we see the first glimpse of an approach in which a natural transition from QFT to NRQM could be possible entirely in terms of field operators. We do not use the position *operator* $\hat{x}^\alpha$ at all and both t and $\boldsymbol{x}$ remain c-number labels, even after we have obtained the NRQM. The propagator obtained in Sect. 2 can be expressed in terms of the field operators, again, both in NRQM and in QFT. In the relativistic case, the field operators are Lorentz invariant but they do not commute on space-like surfaces. Hence they cannot be used to construct physical observables directly. (This requires some more work and leads to the notion of antiparticle *both in QFT and NRQM*; see Sect. 6.)

The discussion in Sects. 2 and 3 tells us that: (i) Lorentz invariance or the square root in the Hamiltonian does not

introduce any serious conceptual difficulties in developing RQM. (ii) The fact that particles are nonlocalizable in RQM leads to difficulties in defining the position eigenkets but these difficulties can be handled by working in momentum representation and introducing the necessary Fourier transforms. (iii) But when we do that, the resulting propagator G_+ does not satisfy the composition law necessary for it to propagate a wave function. In fact, we cannot even properly define $\psi(x) = \langle t, \boldsymbol{x}|\psi\rangle$ in the absence of position eigenkets $|t, \boldsymbol{x}\rangle$. (iv) The formalism leads to the concept of a field operator both in NRQM and QFT but we run into trouble with the notion of causality in QFT. This is related to the particle states not being localizable but, as we shall see later, the issue is deeper and is linked to the existence of antiparticles.

In Sect. 4 we look at the same (free particle) system, described by a Hamiltonian $H(|\boldsymbol{p}|)$, from the path integral perspective. In Sect. 4.1, I show how the *Hamiltonian* path integral is indeed straightforward to evaluate for such systems – even for the relativistic case with a square root Hamiltonian. If you use the standard measure $d^n x\, d^n \boldsymbol{p}$ in the Hamiltonian path integral, you get the correct answer in NRQM; but, in the case of RQM, you get a propagator – called the Newton–Wigner propagator – which is *not* Lorentz invariant. It is possible to tinker with the path integral measure – taking a cue from our discussion in Sect. 2 – and arrange matters so that the resulting propagator *is* Lorentz invariant. This procedure again leads to the same propagator $G_+(x_2, x_1)$ obtained earlier. This also means that we inherit all the difficulties encountered earlier.

In Sect. 4.2, I study the same system using a *Lagrangian* path integral. Again, there is a natural way of defining the measure for this path integral which leads to the correct result in NRQM. The same procedure, when applied to the relativistic Lagrangian, leads to nonsense – that is, the path integral does not exist for any choice of the measure. The fact that, for the relativistic particle, the Hamiltonian path integral exists while the Lagrangian path integral does not can be traced to the structure of the Hamiltonian. One can write down a general condition which must be satisfied by the Hamiltonian if the Lagrangian and Hamiltonian approaches have to lead to the same result. The square root Hamiltonian of the relativistic particle violates this condition. This is probably the only occasion in which the square root in the Hamiltonian leads to a serious technical difficulty.

There is, however, another – rather elegant – procedure for defining the Lagrangian path integral for a relativistic particle. This makes use of the geometric interpretation of the relativistic action as the path length in the Euclidean space. You can then define the path integral in an Euclidean lattice and obtain a continuum limit using a natural regularization. I do this in Sect. 5 and show that the resulting propagator $G_R(x_2, x_1)$ is the standard Feynman propagator in QFT with the Fourier space representation:

---

[8] Recall that I denote by $p.x$ the on-shell dot product in which $p_0 = (\boldsymbol{p}^2 + m^2)^{1/2}$, while $p_a x^a$ will denote the off-shell dot product.

[9] This is built from the so-called "positive frequency" modes of the Klein–Gordan (KG) equation and has to be distinguished from the Feynman propagator we will come across later on. We shall develop all these results in detail in the coming sections.

$$G_R(p) = \int d^D x \, G_R(x) \exp(ip_a x^a) = \frac{i}{(p^2 - m^2 + i\epsilon)}. \tag{5}$$

In Sect. 5.2, I show that this particular path integral approach for the relativistic case is very similar to the path integral based on the Jacobi action for a non-relativistic free particle. This mathematical identification clarifies several peculiar features of the Feynman propagator. I also discuss briefly some aspects of reparametrization invariance and its connection with the Jacobi action.

Obtaining the Feynman propagator from a path integral prescription is gratifying but this does not again help in our task of obtaining NRQM from QFT. In Sect. 5.3, I discuss the non-relativistic limit of $G_R(x_2, x_1)$ and show that it does *not* reduce to the propagator $G_{NR}(x_2, x_1)$ of NRQM. So while the lattice regularization provides a natural way of obtaining $G_R(x_2, x_1)$, it does not help us in obtaining the NRQM limit in a seamless manner. Once again, we cannot use $G_R(x_2, x_1)$ to propagate a relativistic wave function because $G_R(x_2, x_1)$ does not obey the correct composition law and does not have the appropriate limit. In Sect. 5.5, I provide a brief discussion of the different composition laws obeyed by relativistic and non-relativistic propagators and how the relativistic composition law goes over to a non-relativistic one in the $c \to \infty$ limit. This discussion clarifies several issues discussed in the literature.

In Sect. 2, we obtain $G_+(x_2, x_1)$ as a matrix element of a time evolution operator provided the states $|x\rangle$ are defined via Fourier transform from the eigenkets $|p\rangle$ of the momentum operator. On the other hand, $G_R(x_2, x_1)$ is obtained in Sect. 5 from a lattice regularization procedure, applied to the path integral, and it is not clear whether it is also a matrix element of the time evolution operator. Strictly speaking, it is not. However, it is possible to express it as such a matrix element using a particular integral representation of the time evolution operator. I do this in Sect. 5.4 and show how this approach connects with the discussion in Sect. 5.2.

These results show how difficult it is to obtain the NRQM from QFT in a straightforward manner. We run into difficulties both in the Hamiltonian approach and in the path integral approach. The lattice regularization of the relativistic path integral does lead to the QFT propagator $G_R(x_2, x_1)$. But this propagator does not have a single-particle NRQM limit. This is to be expected because $G_R(x_2, x_1)$ contains information as regards *both* particles and antiparticles. In the NRQM limit, it should therefore represent the dynamics of *both* the particle and the antiparticle rather than just a single particle. I show how this result arises – thereby answering the question raised in the subtitle of this paper! – in the last two sections.

Sections 6 and 7 identify the necessary ingredients for the NRQM to arise in the appropriate limit of QFT. This

is done by using a *pair* of field operators rather than a single relativistically invariant operator. Such a pair restores microscopic causality in QFT and collectively describes a particle-antiparticle system. This behavior survives in the NRQM limit and we obtain the Schroedinger equation for *two* field operators, one describing the particle and the other describing the antiparticle. They coexist on an equal footing in the NRQM limit.

So, I have good news and bad news. Good news is that one *can* obtain NRQM, as a limiting case of QFT, if – but only if – we interpret NRQM in terms of a field operator satisfying the Schroedinger equation *à la* (what is usually called, quite misleadingly, as) the "second quantized" approach. The bad news is that you *cannot* get the standard formalism (viz. the stuff we teach kids in QM101, in which x^α and p_β are treated as operators and $\psi(t, x) = \langle t, x | \psi \rangle$ is a "wave function" etc.) as a natural limiting case of QFT. Section 8 discusses some of the broader implications of this result.

While the main focus of this paper is on the conceptual issues (and it does clarify and highlight several of them), there are also many interesting results of technical nature which either do not exist in the previous literature or not adequately discussed. I mention below some of them:

(a) The Hamiltonians for both the relativistic and the non-relativistic (free) particle depends only on their momentum. Section 2 discusses such systems, for which $H(p, x) = H(|p|)$, in a *unified* manner and identifies the reasons why, in spite of this simplicity, we do not have an RQM but we have an NRQM. The unified, focused, discussion should have found a place in textbooks but it had not.

(b) The most natural way of defining a path integral, either from a Hamiltonian $H(p)$ or from a Lagrangian $L(\dot{x})$, is by time slicing. (We look for other "sophisticated" methods only when this approach fails but alas, often, without investigating *why* exactly it failed!). Section 4.1 explains what happens (or goes wrong) when you attempt time slicing with the Hamiltonian for a relativistic particle; I have not seen such an explicit discussion, e.g., about the issues regarding choice of measure, see Eq. (65), in the published literature. Section 4.2 takes up the corresponding question in the case of the Lagrangian path integral. I show that there is a natural way of defining the time-sliced path integral leading to Eq. (72) and use this to clearly contrast the NR case with the relativistic case. I have not seen such a discussion – leading to e.g., Eq. (79) and the discussion in the two paragraphs following Eq. (79) – in the literature.

(c) One consequence of the above analysis is the following: It clearly shows that the Hamiltonian and the Lagrangian time-slicing procedures are not equivalent – another fact which is inadequately stressed in the literature. I also

identify the formal condition, Eq. (80), for their equivalence which I have not seen in the literature, at least not in this context (though it might exist buried somewhere in the literature on formal path integral techniques).

(d) Much of the discussion in different subsections of Sect. 5 is new. In particular, the discussion in Sect. 5.1 leading to e.g., to the interpretation in Eq. (97), the NR limit of lattice regularization in Eqs. (104)–(112), comments in the last paragraph of Sect. 5.2 leading to Eq. (135) are either entirely new or highlights aspects inadequately discussed in the literature.

(e) Section 5.3 shows that you cannot get the NR propagator from the $c \to \infty$ limit of the Feynman propagator. Again, I have not seen an explicit discussion of this (correct) result in the textbook literature. The result in Sect. 5.4 is new and clarifies the structure of the G_F from an alternative point of view.

(f) Section 6 emphasizes the fact that the standard KG field is built from *two* fields which in the NR limit represent the particle and the antiparticle. This, by itself, is not new and exists in several textbooks including my own [2]. But it assumes importance in the context of Eq. (194) which *I claim nobody understands*, in spite of it being the key equation in QFT, allowing the formalism to work. The fact that the path integral for, ostensibly, a single relativistic particle actually describes the propagation of *two* particles is the key issue here and the discussion in Sect. 6 provides the backdrop for it.

Of course, these technical results are just the trees in the wood of conceptual discussion and, hopefully, the reader will not miss the latter for the former.

2 Quantum theory of a system with the Hamiltonian $H(p, x) = H(|p|)$

The classical dynamics of a free particle is completely described by the action which has no explicit dependence on the space or time coordinates:

$$A = \int dt\, L(\dot{x}) = \int dt[p \cdot \dot{x} - H(p)] \qquad (6)$$

in terms of a well-defined Lagrangian $L(\dot{x}) = L(|\dot{x}|)$ or a Hamiltonian $H(p) = H(|p|)$. In the case of a *non-relativistic* free particle we take:

$$L_{\mathrm{NR}}(\dot{x}) = \frac{1}{2}m\dot{x}^2; \qquad H_{\mathrm{NR}}(p) = \frac{p^2}{2m}, \qquad (7)$$

while, for the *relativistic* free particle, we have[10]

$$L_R(\dot{x}) = -m(1 - \dot{x}^2)^{1/2}; \qquad H_R(p) = (p^2 + m^2)^{1/2}. \quad (8)$$

In either case, the Lagrangian and the Hamiltonian are independent of x and we can deal with both of them at one go. The classical equations of motion are easy to solve leading to $p = p_0 = $ constant, $x^\alpha(t) = F^\alpha t + x^\alpha(0)$, where $F^\alpha \equiv (\partial H / \partial p_\alpha) = $ constant. That is the end of the story.

What about the quantum theory? If one does not bring in any extra symmetry considerations, then the quantum theory of any system with $H = H(|p|)$ is also trivial in the Heisenberg picture. We upgrade the position and momentum to operators satisfying the commutation rule $[x^\alpha, p_\beta] = i\delta^\alpha_\beta$, which can be concretely implemented – in the space of normalizable complex functions – in the *momentum* representation with $\hat{x}^\alpha = i\partial/\partial p_\alpha$. Since the Hamiltonian commutes with momentum, $\hat{p}(t) = \hat{p}(0)$. It is trivial to integrate the operator equation for x^α and obtain $\hat{x}^\alpha(t) = \hat{F}^\alpha t + \hat{x}^\alpha(0)$ where $\hat{F}^\alpha \equiv (\partial \hat{H} / \partial \hat{p}_\alpha) = $ constant. Since we have solved the operator equations, we can answer any question as regards the quantum dynamics. Obviously, this procedure should work for $H_{\mathrm{NR}}(p) = p^2/2m$ as well as for $H_R(p) = (p^2 + m^2)^{1/2}$.

So it is not the form of the Hamiltonian which creates problems when we try to construct relativistic quantum mechanics (RQM) of a free, single, particle. But we do know that combining principles of special relativity and quantum theory does require more drastic modifications of the description and, in fact, we cannot have a viable, single-particle quantum theory based on, say, a relativistically invariant wave function. The question arises as to why this is the case.

When you move from NRQM to RQM, there are *two* new ingredients which come in. First, the Hamiltonian for a free particle changes from $H_{\mathrm{NR}}(p) = p^2/2m$ to $H_R = +(p^2 + m^2)^{1/2}$ with corresponding changes in the dynamical equations. Second, we want the physics to respect Lorentz invariance rather than Galilean invariance. As we have seen above, the square root structure of the Hamiltonian does not create any new conceptual issues when we use the momentum representation and Heisenberg picture.[11] The next suspect, of course, is the requirement of Lorentz invariance. As we shall see, the issue of maintaining Lorentz invariance requires having the correct, relativistically invariant, inte-

---

[10] The Hamiltonian for a classical relativistic particle is positive definite and the square root is taken with positive sign in $H(p) = +(p^2 + m^2)^{1/2}$. This is the classical system we want to quantize – not a strange one with $H(p) = \pm(p^2 + m^2)^{1/2}$, which has no classical meaning.

[11] If you attempt to write the Schroedinger equation with H_R in the coordinate representation, the square root in the Hamiltonian will make the equation nonlocal. But then, if you write the Schroedinger equation for a non-relativistic particle moving under the action of a non-polynomial potential $V(x)$ in the *momentum* representation, you will again get a nonlocal Schroedinger equation. So this by itself is not a conceptual difficulty; but merely a technical nuisance.

gration measure in the momentum space when we describe, say, the momentum eigenstates of particles. Roughly speaking, you can ensure that a classical theory is relativistically invariant, if you ensure that the dynamical equations are relativistically invariant. But in quantum theory, you need to ensure that *both* the dynamical equations (for the operators in Heisenberg picture, say) as well as the description of quantum states in the Hilbert space are relativistically invariant.[12] The first requirement – viz. relativistic invariance of dynamical equations – can be ensured by using a relativistically invariant action or Hamiltonian; but the second requirement does not have a direct analog in *classical* relativistic mechanics. We will see that this requirement is the root cause of several nontrivial features in QFT. We will now see in some detail the mathematical consequences of these requirements.

2.1 Propagators in momentum and coordinate spaces

Since a Hermitian momentum operator has to exist for the proper definition of $H(\boldsymbol{p})$, we start by introducing a complete set of orthonormal momentum eigenkets, $|\boldsymbol{p}\rangle$, which must exist for any system described by a Hamiltonian of the form $H(\boldsymbol{p})$, including NRQM and RQM. We would then like $\langle \boldsymbol{p}'|\boldsymbol{p}\rangle$ to be proportional to $\delta(\boldsymbol{p} - \boldsymbol{p}')$. This works in NRQM but the integration over $d^n \boldsymbol{p}\,\delta(\boldsymbol{p} - \boldsymbol{p}')$ is not Lorentz invariant. The relativistically invariant measure for momentum integration is $d\Omega_{\boldsymbol{p}} \equiv d^n \boldsymbol{p}/(2\pi)^n (1/\Omega_{\boldsymbol{p}})$ with $\Omega_{\boldsymbol{p}} = 2\omega_{\boldsymbol{p}}$. So we need to postulate:

$$\langle \boldsymbol{p}'|\boldsymbol{p}\rangle = (2\pi)^n \Omega_{\boldsymbol{p}} \,\delta(\boldsymbol{p} - \boldsymbol{p}'), \qquad d\Omega_{\boldsymbol{p}} \equiv \frac{d^n \boldsymbol{p}}{(2\pi)^n} \frac{1}{\Omega_{\boldsymbol{p}}}, \quad (9)$$

so that $\langle \boldsymbol{p}'|\boldsymbol{p}\rangle d\Omega_{\boldsymbol{p}} = \delta(\boldsymbol{p}' - \boldsymbol{p})d^n \boldsymbol{p}$. In NRQM we can take $\Omega_{\boldsymbol{p}}$ to be a constant, or even unity; but in RQM the Lorentz invariance of the measure for momentum integration $d\Omega_{\boldsymbol{p}}$ requires the factor $\Omega_{\boldsymbol{p}} = 2\omega_{\boldsymbol{p}}$. By keeping the choice of $\Omega_{\boldsymbol{p}}$ unspecified in the algebraic expressions, we take care of the two cases at one go; further, in the non-relativistic limit, $\omega_{\boldsymbol{p}}$ can be approximated by the constant m allowing us to take the limit seamlessly. With this definition, the resolution of unity and the consistency condition on the momentum eigenkets become

$$1 \equiv \int d\Omega_{\boldsymbol{p}'}|\boldsymbol{p}'\rangle\langle \boldsymbol{p}'|; \qquad |\boldsymbol{p}\rangle \equiv \int d\Omega_{\boldsymbol{p}'}|\boldsymbol{p}'\rangle\langle \boldsymbol{p}'|\boldsymbol{p}\rangle. \quad (10)$$

These relations can be taken care of by the choices in Eq. (9). In the integration measure as well as in the Dirac delta function, we have introduced a factor $\Omega_{\boldsymbol{p}}$, which, of course, cancels out in the right-hand side of the second relation in Eq. (10).

---

[12] It is possible to address some of these issues, very formally, in terms of the structure of the Lorentz group and Galilean group and their interrelationship. We will, however, adopt a more transparent and down-to-earth approach in this paper.

Given these momentum eigenstates, we can define a natural *momentum space* propagator by the rule:

$$G(t_b, \boldsymbol{p}_b; t_a, \boldsymbol{p}_a) \equiv \langle \boldsymbol{p}_b|e^{-it\hat{H}(\boldsymbol{p})}|\boldsymbol{p}_a\rangle$$
$$= (2\pi)^n \Omega_{\boldsymbol{p}_b} \,\delta(\boldsymbol{p}_a - \boldsymbol{p}_b) \exp{-itH(\boldsymbol{p}_b)}; \quad (11)$$

where $t \equiv t_b - t_a$. Given any arbitrary state $|\phi\rangle$ in the Hilbert space we can "propagate" the complex function $\phi(t_a, \boldsymbol{p}_a) \equiv \langle \boldsymbol{p}|\phi\rangle$ by this propagator:

$$\phi(t_b, \boldsymbol{p}_b) = \int d\Omega_{\boldsymbol{p}_a} G(t_b, \boldsymbol{p}_b; t_a, \boldsymbol{p}_a)\phi(t_a, \boldsymbol{p}_a)$$
$$= \phi(t_a, \boldsymbol{p}_b) \exp{-itH(\boldsymbol{p}_b)}. \quad (12)$$

So the momentum space evolution is just a change in phase. Since momentum operator generates translation in space, it seems natural to introduce a *position* space propagator by the definition:

$$G(t_b, \boldsymbol{x}_b; t_a, \boldsymbol{x}_a) \equiv \int d\Omega_{\boldsymbol{p}_a} d\Omega_{\boldsymbol{p}_b} G(t_b, \boldsymbol{p}_b; t_a, \boldsymbol{p}_a)$$
$$\times \exp{i(\boldsymbol{p}_b \cdot \boldsymbol{x}_b - \boldsymbol{p}_a \cdot \boldsymbol{x}_a)}. \quad (13)$$

Using Eq. (11) in Eq. (13) and performing the integrations, we get the propagator, $G(x) \equiv G(t_b, \boldsymbol{x}_b; t_a, \boldsymbol{x}_a)$ where $x = x_b - x_a$, for both NRQM and RQM at one go, in the form

$$G(x) = \int d\Omega_{\boldsymbol{p}} \exp(-ip \cdot x)$$
$$= \int \frac{d^n \boldsymbol{p}}{(2\pi)^n} \frac{1}{\Omega_{\boldsymbol{p}}} \exp(-ip \cdot x) \quad (14)$$

where we have introduced the four-component object (in both NRQM and RQM) by $p^a = (H(\boldsymbol{p}), \boldsymbol{p})$, which, of course, is a genuine four-vector in RQM and just a convenient notation in NRQM. For later reference, note that the standard spatial Fourier transform (defined with the measures $d^n \boldsymbol{x}$ and $d^n \boldsymbol{p}/(2\pi)^n$) of this propagator is given by

$$G_{\boldsymbol{p}}(t) \equiv \int d^n \boldsymbol{x}\, G(t, \boldsymbol{x})\, e^{-i\boldsymbol{p}\cdot\boldsymbol{x}} = \frac{1}{\Omega_{\boldsymbol{p}}} e^{-itH(\boldsymbol{p})}. \quad (15)$$

Let us now consider the two cases, NRQM and RQM. In NRQM we get

$$G_{\mathrm{NR}}(x) = \int \frac{d^n \boldsymbol{p}}{(2\pi)^n} \exp\left[i\left(\boldsymbol{p} \cdot \boldsymbol{x} - \frac{p^2}{2m}t\right)\right]$$
$$= \left(\frac{m}{2\pi it}\right)^{n/2} \exp\left(\frac{im|\boldsymbol{x}|^2}{2t}\right), \quad (16)$$

and in RQM we have,[13] with $x^2 \equiv x_a x^a$,

$$G_+(x) \equiv \int \frac{d^n \boldsymbol{p}}{(2\pi)^n} \frac{1}{2\omega_{\boldsymbol{p}}} \exp(-ip.x) = F(x^2), \qquad (17)$$

which is clearly Lorentz invariant. For space-like separations, F can be expressed in terms of a Bessel function and decays exponentially; for time-like separations, it can be expressed in terms of a Hankel function and oscillates; it has a singular behavior on the light cone (see, e.g., [2]). So obtaining a Lorentz invariant propagator is not an issue at all. If we take the $c \to \infty$ limit of $G_+(x)$, we get

$$\lim_{c \to \infty} G_+(x) = \frac{e^{-i(mc^2)t}}{2m} \left[G_{\mathrm{NR}} - \frac{i\hbar}{mc^2} \frac{\partial G_{\mathrm{NR}}}{\partial t} + \cdots \right]. \qquad (18)$$

In this expression, the overall factor $(1/2m)$ is irrelevant; the factor $e^{-i(mc^2)t}$ is unavoidable because the rest energy mc^2 will always contribute to the phase. The second and higher order terms within the square bracket in Eq. (18) vanish in the $c \to \infty$ limit. So one can think of the non-relativistic propagator being recovered in the limit:

$$\lim_{c \to \infty} [(2m)e^{i(mc^2)t}]G_+(x) = G_{\mathrm{NR}}, \qquad (19)$$

which seems reasonable. So far, so good.

2.2 The problems in defining localized particle states

We would, however, like to think of this real space propagator, *defined* though the Fourier transform in Eq. (13) to be the same as the matrix element of the time evolution operator:

$$G(x_b, x_a) = G(t, \boldsymbol{x}) = \langle \boldsymbol{x}_b | e^{-it\hat{H}(\boldsymbol{p})} | \boldsymbol{x}_a \rangle \qquad (20)$$

for some suitable states $|\boldsymbol{x}\rangle$. To do this we need to introduce the states $|\boldsymbol{x}\rangle$ labeled by the spatial coordinates. In NRQM they could be thought of as the eigenkets of the operator $\hat{\boldsymbol{x}}(0)$. For a more general system described by an arbitrary $H(\boldsymbol{p})$ like, for e.g. in RQM, we do not have the natural notion of such a position operator. But we can take a cue from the previous results and use the property that the momentum operator is the generator of spatial translations (which holds both in NRQM and RQM) to define $|\boldsymbol{x}\rangle$ along the following

lines:[14]

$$|\boldsymbol{x}\rangle \equiv e^{-i\boldsymbol{x}\cdot\hat{\boldsymbol{p}}} |0\rangle = \int d\Omega_{\boldsymbol{p}} e^{-i\boldsymbol{p}\cdot\boldsymbol{x}} C_{\boldsymbol{p}} |\boldsymbol{p}\rangle,$$

$$C_{\boldsymbol{p}} \equiv \langle \boldsymbol{p}|0\rangle; \quad \langle \boldsymbol{p}|\boldsymbol{x}\rangle = C_{\boldsymbol{p}} e^{-i\boldsymbol{x}\cdot\boldsymbol{p}}. \qquad (21)$$

This defines $|\boldsymbol{x}\rangle$ in terms of a single function $C_{\boldsymbol{p}}$. Inserting a complete set of momentum eigenstates in the matrix element in Eq. (20), and using the last relation in Eq. (21), we can evaluate the propagator explicitly in terms of $C_{\boldsymbol{p}}$. We get

$$G(x) = \int d\Omega_{\boldsymbol{p}} |C_{\boldsymbol{p}}|^2 \exp(-ip.x) \qquad (22)$$

where we have again defined the four-component object $p^a = (H(\boldsymbol{p}), \boldsymbol{p})$ taking care of both NRQM and RQM.

In NRQM, it is natural to take the measure in the momentum space integration with $\Omega_{\boldsymbol{p}} = $ constant; similarly, we can also set $C_{\boldsymbol{p}} = 1$. With these choices and using $H_{\mathrm{NR}} = \boldsymbol{p}^2/2m$ in Eq. (22), we immediately obtain the NRQM propagator given by Eq. (16). In RQM, we want to obtain a Lorentz invariant propagator. In Eq. (22), the measure $d\Omega_{\boldsymbol{p}}$ and the function $\exp(-ip \cdot x)$ are Lorentz invariant. Therefore, the propagator will be Lorentz invariant if we take $C_{\boldsymbol{p}} = $ constant. It is conventional to scale things so that $C_{\boldsymbol{p}} = 1$. Then the propagator is given by the expression in Eq. (17). We have thus arrived at a Lorentz invariant propagator for RQM which can also be interpreted as the matrix element of the time evolution operator through Eq. (20). Unfortunately, the situation is not so simple when we study it more closely.

To begin with, note that the only difference between the relativistic and non-relativistic propagators is in the $(1/\Omega_{\boldsymbol{p}})$ factor, which we can take to be a constant (or even unity) in NRQM but which is $(1/2\omega_{\boldsymbol{p}})$ in QFT. As we shall see, this makes all the difference. From the definition of $|\boldsymbol{x}\rangle$ in Eq. (21), it follows that

$$\langle \boldsymbol{y}|\boldsymbol{x}\rangle = \int d\Omega_{\boldsymbol{p}} e^{-i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{y})} |C_{\boldsymbol{p}}|^2. \qquad (23)$$

If you want localized particle positions, this expression should be proportional to a Dirac delta function. This in turn requires $|C_{\boldsymbol{p}}|^2 = 2\omega_{\boldsymbol{p}}$ to give $d\Omega_{\boldsymbol{p}} |C_{\boldsymbol{p}}|^2 = [d^n \boldsymbol{p}/(2\pi)^n]$. But we get a Lorentz invariant propagator from Eq. (22) only if $|C_{\boldsymbol{p}}|^2 = $ constant in Eq. (22)! So, while the propagator defined through Eq. (20) can be made Lorentz invariant, we do not know *what* it propagates because $|\boldsymbol{x}\rangle$ do not represent localized particle states! (The difficulty in localizing particles states in RQM is discussed extensively in the literature; see, e.g., [4–8].)

---

[13] While discussing the general expressions for the propagator, valid in both NRQM and RQM, we will denote it by $G(x)$ with no subscripts. The propagator in NRQM is unique and will be noted by G_{NR}. In RQM and QFT, we will encounter different types of propagators denoted with different subscripts. This particular one carries the subscript $+$, since it is made of positive frequency solutions of the KG equation.

[14] Here, as well as in most of the discussions which follow, we are interested in expressions at a given time t, taken to be $t = 0$, for convenience. The notion of a state $|\boldsymbol{x}\rangle = |t, \boldsymbol{x}\rangle$ such that $|\boldsymbol{x}\rangle \equiv |0, \boldsymbol{x}\rangle$ will be introduced later in Eq. (27).

Furthermore, with this Lorentz invariant choice $C_p = 1$ we also have the result

$$
\begin{aligned}
\int d^n \boldsymbol{x} \, |\boldsymbol{x}\rangle\langle\boldsymbol{x}|\boldsymbol{p}\rangle &= \int d^n \boldsymbol{x} \, e^{i\boldsymbol{p}\cdot\boldsymbol{x}} |\boldsymbol{x}\rangle \\
&= \int d^n \boldsymbol{x} \, e^{i\boldsymbol{p}\cdot\boldsymbol{x}} \int d\Omega_{q'} \, e^{-i\boldsymbol{q}\cdot\boldsymbol{x}} |\boldsymbol{q}\rangle \\
&= \int d^n q \, \frac{1}{\Omega_q} \delta(\boldsymbol{p}-\boldsymbol{q}) |\boldsymbol{q}\rangle = \frac{1}{\Omega_p} |\boldsymbol{p}\rangle.
\end{aligned}
\tag{24}
$$

So we cannot use the states $|\boldsymbol{x}\rangle$ for the resolution of identity. Equation (24) also shows that it is the combination $\Omega_p d^n \boldsymbol{x}$ rather than $d^n \boldsymbol{x}$ which behaves better. For example, while the measure of integration $d^n \boldsymbol{x}$ is not Lorentz invariant, the combination $\Omega_p d^n \boldsymbol{x}$ is. (We will discuss this aspect in greater detail later on.) In the case of $d^n \boldsymbol{p}$, we could work from the Lorentz invariant combination $d^4 p \delta(p^2 - m^2)\theta(p^0) \propto d^n \boldsymbol{p}/2\omega_p$ but there is no natural analog[15] for that in the case of $d^n \boldsymbol{x}$. The best one can do is to write, for any state $|\psi\rangle$, the relation

$$
\begin{aligned}
|\psi\rangle &= \int d\Omega_p |\boldsymbol{p}\rangle\langle\boldsymbol{p}|\psi\rangle \\
&= \int d\Omega_p \int [d^n \boldsymbol{x} \, \Omega_p] |\boldsymbol{x}\rangle\langle\boldsymbol{x}|\boldsymbol{p}\rangle\langle\boldsymbol{p}|\psi\rangle \\
&= \int \frac{d^n \boldsymbol{p} \, d^n \boldsymbol{x}}{(2\pi)^n} |\boldsymbol{x}\rangle\langle\boldsymbol{x}|\boldsymbol{p}\rangle\langle\boldsymbol{p}|\psi\rangle,
\end{aligned}
\tag{25}
$$

which is Lorentz invariant if the left-hand side is. So there is some kind of resolution of identity in *phase* space:

$$
1 = \int \frac{d^n \boldsymbol{p} \, d^n \boldsymbol{x}}{(2\pi)^n} |\boldsymbol{x}\rangle\langle\boldsymbol{x}|\boldsymbol{p}\rangle\langle\boldsymbol{p}|
\tag{26}
$$

but not in normal space. We will come across this combination again later, while computing phase space path integrals. Note, for future reference, that the natural extension of $|\boldsymbol{x}\rangle = |0, \boldsymbol{x}\rangle$ for $t \neq 0$ is defined as the state $|x\rangle = |t, \boldsymbol{x}\rangle$ through the relation

$$
|x\rangle \equiv |t, \boldsymbol{x}\rangle \equiv e^{iHt} |\boldsymbol{x}\rangle = \int d\Omega_p \, e^{ip.x} |\boldsymbol{p}\rangle
\tag{27}
$$

where $H(\boldsymbol{p})$ is the Hamiltonian.[16]

The propagator we have obtained has another nice property which arises directly from the definition in Eq. (11). It satisfies the *first order* differential equation

$$
(i\partial_t - H(\boldsymbol{p})) \, G = 0
\tag{28}
$$

for any $H(\boldsymbol{p})$. In the specific case of the relativistic free particle, the structure of Eq. (17) tells you that it *also* satisfies the equation

$$
[\partial_a \partial^a + m^2] G_+(x) = 0.
\tag{29}
$$

The zeros in the right-hand sides of Eqs. (28) and (29) are closely related to the fact that the definition in Eq. (20) – as well as the form of the final propagator – is valid for both $t > 0$ and $t < 0$. *Nowhere did we assume that $t > 0$ to obtain the form of the propagator.* The time evolution operator in quantum theory $U(t_2, t_1) \equiv \exp[-iH(t_2 - t_1)]$ evolves a state from $t = t_1$ to $t = t_2$ irrespective of the time ordering of t_2 and t_1; that is, this is a valid evolution operator for *both* $t_2 > t_1$ and $t_2 < t_1$. For example, in NRQM, given a wave function $\psi(t, \boldsymbol{x})$ we can determine the wave function at all the earlier times and later times.[17] Therefore the expression for the propagator in NRQM, defined as the matrix element $\langle \boldsymbol{x}_2 | U(t_2, t_1) | \boldsymbol{x}_1 \rangle$, is valid for both $t_2 > t_1$ and $t_2 < t_1$.

Sometimes it is convenient to *define* another propagator by multiplying G by a theta function in time, getting $U(x_2, x_1) = \theta(t)G(x_2, x_1)$, which will satisfy the differential equation

$$
(i\partial_{t_2} - H) \, U = i\delta(t_2 - t_1)\langle \boldsymbol{x}_2 | \boldsymbol{x}_1 \rangle.
\tag{30}
$$

The right-hand side will reduce to $i\delta(x_2 - x_1)$ in NRQM but not in the relativistic theory. When we bring in Lorentz invariance, we run into trouble regarding the time ordering. The notion of, say, $t_2 > t_1$ is well-defined only if the events x_2 and x_1 are separated by a time-like interval. When the events are separated by a space-like interval, we can always choose a Lorentz frame such that $t_2 = t_1$ and hence $G(x_2, x_1) = \langle \boldsymbol{x}_2 | \boldsymbol{x}_1 \rangle$. If $G(x_2, x_1)$ does not vanish for space-like intervals, then multiplying $G(x_2, x_1)$ by $\theta(t_2 - t_1)$ will not lead to a Lorentz invariant construct.

2.3 Propagator does not propagate the wave functions

The reason why this propagator $G_+(x)$ (in spite of (i) being defined as a time evolution operator for the relativistic Hamiltonian through Eq. (20) and (ii) being Lorentz invariant) cannot be used to define a single-particle RQM is the following: We cannot use it to propagate a wave function with standard probabilistic interpretation in real space. To see this let us recall how this becomes feasible in NRQM. The dynamics of a free particle in NRQM can be described using the propagator $G_{NR}(x_b, x_a)$, which relates the Schroedinger wave function at two different times through the relation

---

[15] Taking a cue from momentum space, one can redefine the integration measure as $d^4 x \delta(x^2 - \mu^2)$ where μ specifies the space-like hypersurface $t^2 - \boldsymbol{x}^2 = \mu^2$. I have explored this possibility [9] but it leads to problems.

[16] The sign in the exponential is correct and gives $\langle t, \boldsymbol{x}| = \langle 0, \boldsymbol{x}| \exp(-iHt)$, which is the correct relation; see, e.g., 1.2.2 of Ref. [2].

[17] The same results hold even in a relativistic theory where $H(t_2 - t_1)$ will be replaced by an integral of $d\Sigma^a P_a$ over a space-like hypersurface Σ of the four-momentum P_a and the evolution proceeds from one space-like hypersurface to another.

$$\psi(x_b) = \int d^n \boldsymbol{x}_a \, G_{\text{NR}}(x_b, x_a) \psi(x_a). \tag{31}$$

This provides the physical interpretation for $G_{\text{NR}}(x_b, x_a)$ as the amplitude for the particle to propagate from the event $\mathcal{A}$ to the event $\mathcal{B}$. One can immediately draw two key conclusions from the existence of a relation like Eq. (31):

(1) Consistency of Eq. (31) in the limit $t_b \to t_a$ tells you that $G_{\text{NR}}(x_b, x_a)$ must satisfy the boundary condition

$$\lim_{t_b \to t_a} G_{\text{NR}}(x_b, x_a) = \delta(\boldsymbol{x}_b - \boldsymbol{x}_a). \tag{32}$$

(2) The propagator must satisfy the transitivity condition (also called the composition law) given by

$$G_{\text{NR}}(x_b, x_a) = \int d^n \boldsymbol{x}_1 \, G_{\text{NR}}(x_b, x_1) \, G_{\text{NR}}(x_1, x_a). \tag{33}$$

This is an extremely stringent condition on the form of the propagator $G_{\text{NR}}(x_b, x_a)$. In the case of a free particle, $G_{\text{NR}}(x_b, x_a)$ must be a function of $x_b - x_a$ alone. It is then straightforward to show (see page 5 of [2]) that the spatial Fourier transform $G_{\text{NR}}(t, \boldsymbol{p})$ must have the form

$$G_{\text{NR}}(t, \boldsymbol{p}) \equiv \int d^n \boldsymbol{x} \, G_{\text{NR}}(t, \boldsymbol{x}) \, \exp(-i \boldsymbol{p} \cdot \boldsymbol{x})$$
$$= \exp[-it F(\boldsymbol{p})]. \tag{34}$$

That is, $G_{\text{NR}}(t, \boldsymbol{p})$, the propagator in momentum space, is a unit norm complex function with a phase that is linear in time.

Neither of these conditions, in Eq. (32), Eq. (33), hold for $G(x)$. The condition in Eq. (32) is violated because $G(0, \boldsymbol{x}_b - \boldsymbol{x}_a) = \langle \boldsymbol{x}_b | \boldsymbol{x}_a \rangle$ is not a Dirac delta function; this is the same issue of $|\boldsymbol{x}_b\rangle$ not representing a localized particle state. The condition in Eq. (33) is violated because the spatial Fourier transform of G, given by Eq. (15), is not of the form of Eq. (34). So the idea of "propagation of a wave function" in Eq. (31) does not work in RQM.

It is interesting to ask how Eq. (32) is reproduced in the non-relativistic limit. Using Eq. (28), we can rewrite Eq. (18), in the limit of $c \to \infty$, as

$$G_+(x) \approx \frac{e^{-i(mc^2)t}}{2m} \left[G_{\text{NR}} + \frac{\lambda_C^2}{2} \nabla^2 G_{\text{NR}} + \cdots \right],$$
$$\lambda_C \equiv \frac{\hbar}{mc}. \tag{35}$$

Taking the limit of $t_2 \to t_1$ we find

$$G_+(\boldsymbol{x}_2 - \boldsymbol{x}_1) \approx \frac{1}{2m} \left[\delta(\boldsymbol{x}_2 - \boldsymbol{x}_1) + \frac{\lambda_C^2}{2} \nabla^2 \delta(\boldsymbol{x}_2 - \boldsymbol{x}_1) + \cdots \right] \tag{36}$$

with a highly singular second term. This implies that

$$(2m) \int d\boldsymbol{x}_1 G_+(\boldsymbol{x}_2 - \boldsymbol{x}_1) \psi(\boldsymbol{x}_1) \approx \psi(\boldsymbol{x}_2) - \frac{\lambda_C^2}{2} \nabla^2 \psi(\boldsymbol{x}_2). \tag{37}$$

The second term is nonlocal and probes the wave function over a region of the size of the Compton wavelength λ_C. Clearly this non-localizability of the particle state is the cause for the trouble which vanishes in the $c \to \infty$ limit. So the propagator $G_+(x)$ cannot be used to propagate anything consistently in RQM.

One might think that the propagation equation (12) in momentum space should lead to a similar equation in real space in terms of the Fourier transform $\psi(t, \boldsymbol{x})$ of $\phi(t, \boldsymbol{p})$. This is indeed true, but the propagator which will appear in that expression is *not* the Lorentz invariant one, defined by Eq. (13). We could define the Fourier transform $\psi(t, \boldsymbol{x})$ of $\phi(t, \boldsymbol{p})$ with either the measure $d^n \boldsymbol{p}$ or with $d\Omega_{\boldsymbol{p}}$ and the two approaches lead to similar difficulties. The $\Omega_{\boldsymbol{p}}$ factors will come in the way when you try to translate Eq. (12) into something like Eq. (31) with $G(t_b, \boldsymbol{p}_b; t_a, \boldsymbol{p}_a)$ replaced exactly by $G(t_b, \boldsymbol{x}_b; t_a, \boldsymbol{x}_a)$. For example, if you define $\psi(t, \boldsymbol{x})$ with the Lorentz invariant measure

$$\psi(x_b) \equiv \int d\Omega_b \phi(t_b, \boldsymbol{p}_b) \exp(i \boldsymbol{p}_b \cdot \boldsymbol{x}_b) \tag{38}$$

and use Eq. (12) you will find that

$$\psi(x_b) = \int d^n \boldsymbol{x}_a \, K(x_b, x_a) \psi(x_a) \tag{39}$$

with

$$K(x_b; x_a) \equiv \int d\Omega_{p_a} d\Omega_{p_b} \left[\Omega_{p_b} G(t_b, \boldsymbol{p}_b; t_a, \boldsymbol{p}_a) \right]$$
$$\times \exp i (\boldsymbol{p}_b \cdot \boldsymbol{x}_b - \boldsymbol{p}_a \cdot \boldsymbol{x}_a)$$
$$= \int \frac{d^n \boldsymbol{p}}{(2\pi)^n} \, \exp(-ip.x) = 2i \frac{\partial G_+}{\partial t}. \tag{40}$$

This $K(x_b; x_a)$ does propagate ψ but it is not Lorentz invariant. As you can see, the extra factor of Ω_{p_b} in the integrand ensures that $K(x_b; x_a)$ reduces to a Dirac delta function when $t \to 0$, ensuring the consistency with Eq. (39). The combination $K(x_b; x_a) d^n \boldsymbol{x}_a$ behaves as a Lorentz scalar though neither $K(x_b; x_a)$ nor $d^n \boldsymbol{x}_a$ individually is, thereby allowing us to define ψ as a Lorentz scalar. Thus we can define a propagation relation only with a propagator which is not Lorentz invariant.[18] The propagator $K(x_b, x_a)$ is sometimes

---

[18] In NRQM, we can treat both momentum eigenstates $|\boldsymbol{p}\rangle$ and position eigenstates $|\boldsymbol{x}\rangle$ on an equal footing, while in the RQM momentum eigenstate $|\boldsymbol{p}\rangle$ acquires a preferred status. Notice, however, that even in textbook NRQM, there is one key difference between these descriptions. The probability density $\rho \equiv |\langle \boldsymbol{x} | \psi \rangle|^2$ in position space satisfies a continuity equation $\partial_t \rho + \nabla \cdot \boldsymbol{j} = 0$ while we do not have a corresponding continuity equation for the probability density $\bar{\rho} \equiv |\langle \boldsymbol{p} | \psi \rangle|^2$ in momentum space. This is hardly emphasized in the text books.

called the Newton–Wigner propagator. (For a small sample of literature dealing with Newton–Wigner states and related topics, see [15–28].)

This is the propagator you get if you forget all about Lorentz invariance and study a system with the Hamiltonian $H = (p^2 + m^2)^{1/2}$ as though you are doing NRQM with this Hamiltonian. In this case, we will be working with $\Omega_p = 1$ in Eq. (9) and will take $C_p = 1$ in Eq. (21). Equation (20) will then lead to $K(x_b, x_a)$. We will also recover the standard resolution of the identity for the states $|x\rangle$ in Eq. (24), because we have set $\Omega_p = 1$. Everything will proceed exactly as in NRQM except for the fact that $p_0 = (p^2 + m^2)^{1/2}$ in Eq. (40). This propagator will satisfy the standard composition law and the boundary condition in Eqs. (32) and (33), which is, of course, necessary for a propagation law of the form Eq. (39) to hold. Finally, if you take the $c \to \infty$ limit, $K(x_b, x_a)$ will reduce to $G_{NR}(x_b, x_a)$ (except for the understandable factor $\exp(-imt)$). So the square root in the Hamiltonian is of no real consequence in developing a quantum theory, *if* you are willing to sacrifice Lorentz invariance. Needless to say, this is too high a price to pay.

The fact that spatial integration with the measure $d^n x$ is not Lorentz invariant also means that a relation like Eq. (33) has no hope of surviving in a Lorentz invariant theory if the propagators are Lorentz invariant. The standard procedure to define invariant spatial integration is to use a (variant of a) combination like $d\Sigma^a F_1 \partial_a F_2 = d^n x\, F_1 \partial_0 F_2$ for two scalar functions F_1, F_2. This, however, does not help us to define a wave function for a relativistic particle. But it again raises the question as to how the correct composition law in Eq. (33) is recovered in the non-relativistic limit; we will discuss this issue in Sect. 5.5.

Some of these ideas involving the states $|k\rangle$ and $|x\rangle$ are usually expressed by introducing a one-particle "wave function" which, as we know, is not a useful notion. Nevertheless, to connect with previous literature, let me briefly mention how this comes about. Consider a state $|\Psi\rangle$ defined in terms of a function $F(k)$ by

$$|\Psi\rangle \equiv \int d\Omega_k F(k)|k\rangle. \tag{41}$$

We clearly have $F(p) = \langle p|\Psi\rangle$. Given the definition of $|x\rangle$ in Eq. (27), we see that

$$\langle x|\Psi\rangle = \int d\Omega_k\, e^{-ikx} F(k) = \bar{F}(x). \tag{42}$$

It is easy to show that this function $\bar{F}(x)$ satisfies the relativistic Schroedinger equation

$$i\partial_t \bar{F}(x) = (-\nabla^2 + m^2)^{1/2} \bar{F}(x) = \hat{H}(\hat{p})\bar{F}(x). \tag{43}$$

By acting on the left-hand side with $i\partial_t$ again, we see that $\bar{F}(x)$ *also* satisfies the Klein–Gordon equation $(\Box + m^2)\bar{(x)} = 0$. The fact that $\bar{F}(x)$, which is analogous to a

single-particle wave function, and the operator $A(x)$ both satisfy the Klein–Gordon equation sometimes creates (avoidable) confusion in the literature.

Because of the $2\omega_k$ factor in the measure $d\Omega_k$, $\bar{F}(x)$ is not a straightforward Fourier transform of $F(k)e^{-i\omega_k t}$ in RQM. This is also reflected in the fact that while $|\Psi\rangle$ has a straightforward expansion in terms of $|k\rangle$, the corresponding expansion is non-local when we attempt it[19] in terms of $|x\rangle$. The norm of the state $|\Psi\rangle$ can be expressed in two equivalent ways:

$$\int d\Omega_k\, F^*(t, k)\, F(t, k) = i \int d\Sigma^a\, \bar{F}^*(x)\, \overleftrightarrow{\partial_a}\, F(x), \tag{44}$$

which shows that it is fairly natural in the momentum space but involves what is called the Klein–Gordon inner product in real space.

3 Fields from propagators in NRQM and RQM

The fact that the relativistic propagator does not propagate a wave function, while the non-relativistic propagator does, leads to the first point of departure between the two. Even though a useful notion of wave function fails to exist in the relativistic case, the propagator does lead to a natural notion of field *operators* (not c-number wave functions) in both NRQM and RQM. They can be introduced in a unified way, and as we shall see later, actually facilitate a seamless transition from QFT to NRQM. This section introduces this idea, which we will explore further in Sect. 6.

To do this, recall that the $|p\rangle$ represents the state with a single particle having a momentum p and energy $H(p)$ both in NRQM and RQM. When a particle is in an external field or when its interacts with other particles, it could evolve from, say, a state $|p_1\rangle$ to $|p_2\rangle$. Such a process can be equivalently thought of as annihilating a particle in state $|p_1\rangle$, leading to a no-particle state, which we will denote by $|0\rangle$, followed by the creation of a particle in $|p_2\rangle$ from $|0\rangle$. To specify these processes, we can introduce a pair of operators A_p and $A_p^\dagger$ ("creation" and "annihilation" operators) which obey the following relations:

$$\left[A_p, A_q^\dagger\right] = (2\pi)^n\, \Omega_p \delta(p - q),$$
$$A_p|0\rangle = 0, \qquad |p\rangle \equiv A_p^\dagger|0\rangle. \tag{45}$$

The first relation defines the commutator structure of the creation and annihilation operators in the momentum space with the Dirac delta function in the right-hand side defined with the invariant measure containing the factor Ω_p. The second relation defines the unique no-particle state $|0\rangle$ as the one

---

[19] Recall our notation $|x\rangle \equiv |x\rangle_{t=0} = |0, x\rangle$. We will use the same notation for all physical quantities.

annihilated by A_p for all p. The third relation constructs the momentum eigenstate from $|0\rangle$ by the action of the creation operator. All these work *both* in NRQM and RQM. Combining Eqs. (27) and (45) we find that $|x\rangle$ can be expressed in the form

$$|x\rangle = \int d\Omega_p A_p^\dagger e^{ip.x}|0\rangle \equiv A^\dagger(x)|0\rangle \qquad (46)$$

where we have defined the operator

$$A(x) \equiv \int d\Omega_p A_p e^{-ip.x}; \qquad A^\dagger(x) \equiv \int d\Omega_p A_p^\dagger e^{ip.x}. \qquad (47)$$

So we find that the state $|x\rangle$ can be obtained from the state $|0\rangle$ by the action of a non-Hermitian "field operator" $A^\dagger(x)$ both in NRQM and in RQM. The propagator we obtained earlier can now be expressed in the form

$$\langle x_2|x_1\rangle = \langle t_2, \boldsymbol{x}_2|t_1, \boldsymbol{x}_1\rangle = \langle 0|A(x_2)A^\dagger(x_1)|0\rangle$$
$$= \int d\Omega_p e^{-ip.x} = G_+(x_2; x_1) \qquad (48)$$

with the four-component object $(\boldsymbol{p}, H(\boldsymbol{p}))$. Again this relation is valid both in NRQM and RQM allowing seamless limiting process.

The difference between NRQM and QFT is in the interpretation of the amplitude in the left-hand side in Eq. (48). In NRQM, the state $|t_1, \boldsymbol{x}_1\rangle$ can be defined as the eigenstate of the position operator $\hat{\boldsymbol{x}}(t_1)$ at time t_1 with eigenvalue $\boldsymbol{x}_1$; that is, $\hat{\boldsymbol{x}}(t_1)|t_1, \boldsymbol{x}_1\rangle = \boldsymbol{x}_1|t_1, \boldsymbol{x}_1\rangle$. Such an interpretation is not possible in RQM, since we do not have a suitable position operator and the states like $|x\rangle$ has to be built from $|\boldsymbol{p}\rangle$ by Fourier transform tricks. We also have the equal time result:

$$\langle t_2, \boldsymbol{x}_2|t_2, \boldsymbol{x}_1\rangle = \int d\Omega_p e^{i\boldsymbol{p}\cdot(\boldsymbol{x}_2-\boldsymbol{x}_1)}, \qquad (49)$$

which is a Dirac delta function in NRQM but not in RQM, because of the $2\omega_p$ factor in the measure, leading to the issue of the non-localizability of particle position.
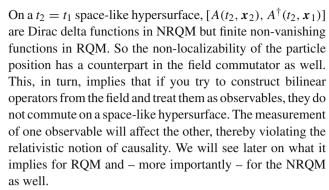
It is trivial to see that the field operator defined in Eq. (47) *always* obeys the *first order* differential equation:

$$[i\partial_t - H(\boldsymbol{p})]A = 0, \qquad [-i\partial_t - H(\boldsymbol{p})]A^\dagger = 0, \qquad (50)$$

including *both* in NRQM and in RQM. In NRQM, it is just the Schroedinger equation. If $H = (\boldsymbol{p}^2 + m^2)^{1/2}$ the field operator *also* obeys the Klein–Gordon equation $\Box A(x) = 0 = \Box A^\dagger(x)$.

A straightforward computation, using Eqs. (47) and (45), shows that the field obeys the commutation rule

$$[A(x_2), A^\dagger(x_1)] = \int d\Omega_p \int d\Omega_q e^{-ipx_2} e^{iqx_1}[A_p, A_q^\dagger]$$
$$= \int d\Omega_p e^{-ipx} \equiv G_+(x_2; x_1) = \langle x_2|x_1\rangle. \qquad (51)$$

On a $t_2 = t_1$ space-like hypersurface, $[A(t_2, \boldsymbol{x}_2), A^\dagger(t_2, \boldsymbol{x}_1)]$ are Dirac delta functions in NRQM but finite non-vanishing functions in RQM. So the non-localizability of the particle position has a counterpart in the field commutator as well. This, in turn, implies that if you try to construct bilinear operators from the field and treat them as observables, they do not commute on a space-like hypersurface. The measurement of one observable will affect the other, thereby violating the relativistic notion of causality. We will see later on what it implies for RQM and – more importantly – for the NRQM as well.

Some of the unnaturalness in the above expressions can be taken care of by sacrificing manifest Lorentz invariance. For the sake of completeness we will briefly describe these constructs and their relationship to the Newton–Wigner position operator. This is usually done by introducing a different set of creation and annihilation operators $a_k, a_k^\dagger$ through the relation $[(2\pi)^n 2\omega_k]^{1/2} a_k \equiv A_k$ etc. A comparison with Eq. (45) shows that these operators obey the simpler commutation rule

$$\left[a_k, a_p^\dagger\right] = \delta(\boldsymbol{k} - \boldsymbol{p}), \qquad (52)$$

which is *not* Lorentz invariant. If we also define f_k by the corresponding rule, $[(2\pi)^n 2\omega_k]^{1/2} f_k \equiv F_k$, we can write the state $|\Psi\rangle$ in Eq. (41) in the form

$$|\Psi\rangle = \int d^n\boldsymbol{k}\, f(\boldsymbol{k})\, a_k^\dagger|0\rangle. \qquad (53)$$

We can also define the fields $a(x), a^\dagger(x)$ in terms of $A(x), A^\dagger(x)$ in an analogous fashion. While the relationship between $F(t, \boldsymbol{k})$ and $f(t, \boldsymbol{k})$ is a simple scaling in momentum space, the corresponding relationship between $\bar{F}(t, \boldsymbol{x})$ and $\bar{f}(t, \boldsymbol{x})$ is much more complicated in real space and is given by

$$\bar{f}(t, \boldsymbol{x}) = \int d^n\boldsymbol{x}'\, Q(\boldsymbol{x}, \boldsymbol{x}')\, \bar{F}(t, \boldsymbol{x}') \qquad (54)$$

where

$$Q(\boldsymbol{x}, \boldsymbol{x}') = \int d\Omega_k\, (2\omega_k)^{3/2}\, e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')}. \qquad (55)$$

One reason people like to work with $a(x)$ and $a^\dagger(x)$ is that it allows defining a set of states $|\boldsymbol{x}\rangle_{\text{NW}}$ as eigenstates of a position operator called the Newton–Wigner position operator. We define $|\boldsymbol{x}\rangle_{\text{NW}}$ through the relation $|\boldsymbol{x}\rangle_{\text{NW}} \equiv a^\dagger(\boldsymbol{x})|0\rangle$. It is then straightforward to verify that these states are eigenstates of an operator $\hat{\boldsymbol{x}}_{\text{NW}}$, that is, $\hat{\boldsymbol{x}}_{\text{NW}}|\boldsymbol{x}\rangle_{\text{NW}} = \boldsymbol{x}|\boldsymbol{x}\rangle_{\text{NW}}$ where the Newton–Wigner position operator $\hat{\boldsymbol{x}}_{\text{NW}}$ is defined as

$$\hat{\boldsymbol{x}}_{\text{NW}} \equiv \int d^n\boldsymbol{x}\, a^\dagger(\boldsymbol{x})\, \boldsymbol{x}\, a(\boldsymbol{x}) = \int d^n\boldsymbol{p}\, a^\dagger(\boldsymbol{p}) \left(i\frac{\partial}{\partial\boldsymbol{p}}\right) a(\boldsymbol{p}). \qquad (56)$$

This appears to be a natural definition both in position space and in momentum space (where $\boldsymbol{x}$ is replaced by $i\partial/\partial\boldsymbol{p}$), but – as we have stressed several times – it is not Lorentz invariant. If we try to re-express it in terms of Lorentz invariant operators $A_{\boldsymbol{p}}$, $A_{\boldsymbol{p}}^{\dagger}$ and the Lorentz invariant integration measure $d\Omega_{\boldsymbol{p}}$, then we get fairly complicated expressions given by

$$\hat{\boldsymbol{x}}_{\text{NW}} = \int d\Omega_{\boldsymbol{p}}\, A_{\boldsymbol{p}}^{\dagger} \left[i\left(\frac{\partial}{\partial\boldsymbol{p}} - \frac{\boldsymbol{p}}{2\omega_{\boldsymbol{p}}^2} \right) \right] A_{\boldsymbol{p}}$$
$$= \int d^n\boldsymbol{x}\, A^{\dagger}(\boldsymbol{x}) \left[\boldsymbol{x} + \frac{\nabla}{2(m^2 - \nabla^2)^{1/2}} \right] A(\boldsymbol{x}), \quad (57)$$

which are obviously not Lorentz invariant. These features once again stress the fact that a single-particle description of RQM is not easy to obtain.

3.1 Aside: some general comments

I have taken a particular approach to demonstrating the problems which arise when one attempts to introduce a Lorentz invariant, single-particle description in RQM with a natural definition of probability. Given the importance of this issue, it is not surprising that many people have attempted to do it from many other perspectives in the past. Each of them requires making some compromise and it is only fair to say that none of them appear natural. This is in fact the major reason why people adhere to the standard interpretation of QFT, in which one no longer attempts an interpretation in terms of a "relativistic wave function". Further discussion in this paper will confirm this point of view.

But before I proceed, it is probably worthwhile to make some general comments about these attempts, which will further clarify the situation. The basic point is strikingly simple: In NRQM you can treat: (i) the momentum operator in the position basis $\hat{p}_\alpha = -i\partial/\partial x^\alpha$ and (ii) the position operator in momentum basis $\hat{x}^\alpha = i\partial/\partial p_\alpha$ on an equal footing. This is because both are unconstrained variables (in a sense which will become clear in a moment) and the corresponding measures of integration are identical in form, being proportional to $d^D\boldsymbol{x}$ and $d^D\boldsymbol{p}$. A natural generalization to RQM will be[20] to use the momentum operator in the position basis being $\hat{p}_a = i\partial/\partial x^a$ and the position operator in momentum basis being $\hat{x}^a = -i\partial/\partial p^a$ (in our mostly negative signature). The essential problem is that the four-momentum is a constrained variable, satisfying the condition $p^a p_a = m^2$, while the four-coordinate x^a has no such constraint. This also implies a key difference between the measures of integration in coordinate and momentum spaces. As long as the mass m

is treated as a Lorentz invariant, scalar constant, this asymmetry will always surface up somewhere in the formalism. As soon as we do this, we also have to treat the coordinate time $\hat{x}^0$ as an operator with all sorts of interpretation issues. One invariably pays a price for such attempts, for example, in the form of having to make m a variable, dynamical entity, rather than retaining it as a parameter, which happens, e.g., in approaches like [10].

Other attempts to handle this issue demand working with an ensemble of particles (see, e.g., [11,12] for a sample) – rather than a single-particle theory – with several peculiar interpretation issues. In addition, it being a many-particle description one runs into difficulties in defining a center-of-mass with expected properties. Moreover, the entire formalism lacks naturalness and one wonders whether this is a remedy worse than the disease. We again see that a strictly single-particle description with a constant mass parameter is not easy to obtain.

There is actually a fundamental reason why such issues arise and one is forced away from a constant mass description (see, e.g., [13]), which I will describe very briefly. Let us assume there exist an operator X^a and quantum states $|\psi\rangle$ etc. such that $\langle\psi|X^a|\psi\rangle = x^a$ are the coordinates of a localized event. (I temporarily use capital letters to denote operators to avoid the clutter of adding 'hats'). Then, using the facts that: (a) a Lorentz–Poincaré transformation is to be implemented in the Hilbert space by a unitary operator, and (b) knowing the transformation rule for the coordinates x^a, one can determine the commutation rules of the position and momentum operators. We will then find that the position operator X^a does *not* commute with the operator corresponding to the Casimir invariant $P^2 \equiv P^a P_a \equiv M^2$. In fact, you get $[X^a, M^2] = -2i P^a$, which can lead to all sorts of trouble. For example, working in the subspace which excludes zero-mass states, we can rewrite this relation as $[X^a, M] = -2i P^a/M$, which will lead to the uncertainty relation (with c-factors reintroduced) $\Delta X^a \Delta(Mc) \geq (\hbar/2)|\langle P^a/Mc\rangle|$. In a single-particle description, we necessarily have $\Delta(Mc) = 0$, violating this bound. We now see why a single-particle description cannot coexist with an operator X^a with standard Lorentz transformation properties. This is the fundamental reason why many previous attempts have to tinker with the mass parameter and either make it a dynamical variable or introduce a many-particle description.

Another possible "way-out" is to tinker with the notion of localization itself, one possibility being to work with hyperplane-dependent states [13]. It is difficult to think of these as localized states around an event and the description is definitely not the most natural one. I merely quote this to show that you need to pay a price one way or another; either the mass becomes a dynamical variable or one needs a more liberal view of what localization means. These attempts also run into trouble [14] with the natural notion of causality

---

[20] For example, we will see later that in the Schwinger proper-time approach one can work with the worldline $x^a(s)$. This provides a natural backdrop for introducing the operators $\hat{x}^a$ etc.

based on the idea that the association of an operator with a spacetime region implies that one can measure it by performing operations confined to that region. In fact, as we shall see later, it is the consistency with micro-causality and Lorentz invariance which makes the single-particle description extremely difficult to come by.

4 Propagators from path integrals

Let us now consider the above results from the path integral perspective, which is expected to provide an intuitive connection between the classical and quantum mechanics. The path integral formalism also has the advantage that we can work with c-number functions rather than with operators, state vectors etc. If the classical physics of the system is described by an action A, specified as a functional of the relevant paths, then $G(x_b, x_a)$ is expected to arise from a sum over all paths connecting the events $\mathcal{A}$ and $\mathcal{B}$, with $\exp(iA)$ being the amplitude for each path. (The relativistic path integral has been studied in several papers in the literature; see, e.g., [15,17,28–34].)

There are *three* forms for the action functional which we will concentrate on. The first one is the Hamiltonian form of the action:

$$A_{\boldsymbol{p}}[\boldsymbol{p}(t), \boldsymbol{x}(t)] \equiv \int_a^b \mathrm{d}t[\boldsymbol{p} \cdot \dot{\boldsymbol{x}} - H(\boldsymbol{p})] \tag{58}$$

where the action $A_{\boldsymbol{p}}$ is a functional of $\boldsymbol{p}(t)$ and $\boldsymbol{x}(t)$, which are treated as independent. The second one is the (more familiar) Lagrangian form of the action:

$$A_{\boldsymbol{x}}[\boldsymbol{x}(t)] = \int_a^b \mathrm{d}t\, L(\dot{\boldsymbol{x}}) \tag{59}$$

in which the action $A_{\boldsymbol{x}}$ is a functional of just $\boldsymbol{x}(t)$. Finally, we can also define a Jacobi action for our system, which is quite different from either of these. It requires a separate treatment which we will take up in Sect. 5.2.

In terms of either $A_{\boldsymbol{p}}[\boldsymbol{p}(t), \boldsymbol{x}(t)]$ or $A_{\boldsymbol{x}}[\boldsymbol{x}(t)]$, the path integral propagator is formally defined by

$$G(x_b, x_a) = \sum_{\boldsymbol{x}(t)} \exp(iA_{\boldsymbol{x}}),$$

$$G(x_b, x_a) = \sum_{\boldsymbol{x}(t), \boldsymbol{p}(t)} \exp(iA_{\boldsymbol{p}}). \tag{60}$$

Of the two, the Lagrangian path integral has an obvious intuitive appeal. In contrast, the "sum over paths" in phase space lacks a simple interpretation because, classically, a single point in phase space determines the trajectory. Also note that, in the Lagrangian path integral, the paths are continuous but not the momenta, while in the Hamiltonian path integral the paths are also discontinuous making the physical picture harder to interpret. So the meaning of the Hamiltonian path

integral is not as straightforward as that of the Lagrangian path integral.

If we are ensured that these path integrals lead to the same propagator (as they do in NRQM) one would have preferred the Lagrangian path integral, at least as a formal expression.[21] Unfortunately the Hamiltonian and Lagrangian path integrals are not guaranteed to lead to the same result. In fact, we will see that the most natural definition for the Lagrangian path integral does *not* work in the case of a relativistic particle, while the Hamiltonian path integral can be *made to* work with some extra tinkering of the measure. We will now examine both, starting from the Hamiltonian path integral.

4.1 Propagator from Hamiltonian path integral

Let us work out the Hamiltonian path integral for the "free particle" with $H = H(\boldsymbol{p})$ taking care of both NRQM and RQM at one go. The standard procedure which we will adopt involves the following steps.

(i) We discretize the time interval $t_b - t_a$ into N intervals of size ϵ such that $N\epsilon = t_b - t_a$. At the end of the computation we take the limit of $N \to \infty, \epsilon \to 0$, keeping the product $N\epsilon = t_b - t_a$ a constant.

(ii) We discretize the action and treat it as a function of $(\boldsymbol{p}_j, \boldsymbol{x}_j)$ where $j = 0, 1, 2, \ldots, N$, with the identifications $\boldsymbol{x}_0 = \boldsymbol{x}_a, \boldsymbol{x}_N = \boldsymbol{x}_b$ defining the end points. This discretized action is given by

$$\begin{aligned} A_{\boldsymbol{p}} &= \sum_{j=1}^N \left[\boldsymbol{p}_j \cdot (\boldsymbol{x}_j - \boldsymbol{x}_{j-1}) - \epsilon H(\boldsymbol{p}_j)\right] \\ &= \sum_{j=1}^{N-1} \left(\boldsymbol{p}_j - \boldsymbol{p}_{j+1}\right) \cdot \boldsymbol{x}_j \\ &\quad + \boldsymbol{p}_N \cdot \boldsymbol{x}_N - \boldsymbol{p}_1 \cdot \boldsymbol{x}_a - \epsilon \sum_{j=1}^N H(\boldsymbol{p}_j). \end{aligned} \tag{61}$$

As we will see, the second form of $A_{\boldsymbol{p}}$ is more convenient for the computation.

(iii) The sum over paths is treated as integrations over $(\boldsymbol{p}_j, \boldsymbol{x}_j)$. The $\boldsymbol{x}_j$ integrations are over $j = 1, 2, \ldots, N-1$, keeping the end points fixed, so that there are $N-1$ integrals to do. The $\boldsymbol{p}_j$ integrations are over $j = 1, 2, \ldots, N$ so that there is one extra momentum integration.

The crucial question, of course, is the choice of measure for the integration. The natural choice is to use just $d\bar{\Gamma} = d^n x \, d^n p/(2\pi)^n$. In this case, the propagator is defined by the integrals over the discretized action, given by the second equation in Eq. (61):

$$G = \int d\bar{\Gamma}_1 \cdots d\bar{\Gamma}_{N-1} \int \frac{d^n p_N}{(2\pi)^n} \, \exp i A_p. \tag{62}$$

Note that this choice will lead to the surviving momentum integration (because there are N momentum integrations but only $N-1$ position integrations) to appear with the measure $d^n p/(2\pi)^n$. At each intermediate step, the integration over dx_n leads to a Dirac delta function on the momentum. (This is the advantage of using the second expression in Eq. (61).) On integrating over the momenta, only the contribution from one end point survives (since there is no corresponding x integration) leading to the propagator:

$$G(x) = \int \frac{d^n p}{(2\pi)^n} \, e^{-ip \cdot x} \tag{63}$$

defined again using the four-component object $p_a = (H, p)$. This leads to the standard propagator G_{NR} in Eq. (16) in NRQM. But in the case of RQM, the surviving integration over $d^n p_N/(2\pi)^n$ will break the Lorentz invariance, leading to the Newton–Wigner propagator encountered earlier in Eq. (40):

$$K(x) = \int \frac{d^n p}{(2\pi)^n} \, e^{-ip \cdot x} = 2i \frac{\partial}{\partial t_b} G_+(x_b, x_a). \tag{64}$$

This Newton–Wigner propagator is obviously not Lorentz invariant and is built from positive frequency solutions of Klein–Gordon equation. This situation is completely analogous to NRQM; the price we have paid is the lack of Lorentz invariance which, unfortunately, is too high.

If we want a Lorentz invariant propagator the final momentum integration measure has to be $d\Omega_p = d^n p/(2\pi)^n (1/\Omega_p)$. But this will lead to a wrong result in the intermediate integrals, if it is used with $d^n x$. To solve this problem, we are forced to tinker with the choice of measure and choose it to be

$$d\Gamma = \left[d^n x \, \Omega_n \right] \left[d^n p \, \Omega_n^{-1} \right] (2\pi)^{-n}. \tag{65}$$

At each intermediate step, this is the same as the original choice $d\bar{\Gamma} = dx_n \, dp_n/(2\pi)^n$ (since the Ω_n factors cancel) but the surviving momentum integration will come with an invariant measure. With this choice, the propagator is now defined by the integrals over the discretized action, with

$$G = \int d\Gamma_1 \cdots d\Gamma_{N-1} \int d\Omega_N \, \exp i A_p. \tag{66}$$

At each intermediate step, the integration over dx_n again leads to a Dirac delta function on the momentum. On integrating over the momenta, only the contribution from one end

point survives (since there is no corresponding x integration) leading to the final result:

$$G = \sum_p \sum_x e^{i A_p} = \int d\Omega_{p_a} \, e^{-i p_a \cdot x} = G_+, \tag{67}$$

which matches with the result in Eq. (14) obtained from the Hamiltonian procedure.

A somewhat more intuitive way of obtaining these results is as follows: Rewrite the Hamiltonian form of the action by eliminating $\dot{x}$:

$$A_p = p \cdot x \Big|_a^b - \int_a^b dt \left[x \cdot \dot{p} + H(p) \right]. \tag{68}$$

We then define the measure for the sum over $x(t)$ such that it gives a Dirac delta function of $\dot{p}$. Then the path integral becomes

$$G = \sum_p \sum_x e^{i A_p} = \sum_p \delta(\dot{p}) e^{i(p_b \cdot x_b - p_a \cdot x_a)} \, e^{-i \int dt \, H}. \tag{69}$$

The existence of a delta function tells you that in the sum p (and thus $H(p)$) remains constant, which immediately leads to the result in Eq. (67).

Clearly, a nontrivial choice of measure – which is not easy to justify from first principles – was needed to get the correct result. The final, surviving momentum integral has to come with the measure $d\Omega_p$ to give a Lorentz invariant result but the intermediate integrations have to be over $dx_n \, dp_n$ to give the Dirac delta functions. This requires one to define the phase space measure by Eq. (65), which is the structure we were led to earlier, in Eq. (25). This is the first instance of our running into a measure problem and, of course, it does not arise in NRQM when $\Omega_p = 1$. Since the final answer is G_+ we will inherit all the issues discussed in Sect. 2.

4.2 Propagator from the Lagrangian path integral

There is a fairly general and natural procedure for defining the Lagrangian path integral by time slicing which works very well for the non-relativistic particle but fails for the relativistic particle. To see how this disaster comes about, we will next consider the discretized version of the Lagrangian path integral for both cases.

To compute the propagator $G(x_b, x_a)$ it is again convenient to divide the time interval $(t_b - t_a)$ into N equal parts of interval ϵ such that $N\epsilon = t_b - t_a$. In the interval (t_{n-1}, t_n) we will approximate the action by $A = \epsilon L(\dot{x}) = \epsilon L((x_n - x_{n-1})/\epsilon)$. The full propagator is obtained by multiplying the amplitudes for each of the infinitesimal intervals with the intermediate spatial coordinates integrated out. This would lead to an expression for the path integral of the form

$$\sum_{\mathbf{x}} \exp i \int L \mathrm{d}t = \int \prod_{k=1}^{(N-1)} d\mathbf{x}_k \, M(N, \epsilon) \, \exp i \epsilon L(\boldsymbol{\ell}/\epsilon);$$
$$\boldsymbol{\ell} \equiv (\mathbf{x}_n - \mathbf{x}_{n-1}) \tag{70}$$

where $M(N, \epsilon)$ is a measure which we hope to choose such that the continuum limit exists.

To evaluate this expression, it is convenient to work in the Euclidean sector. (We assume that we can obtain the Lorentzian result by analytic continuation at the end of the calculation.) Let us introduce the spatial Fourier transform of the discretized Euclidean amplitude $e^{-\epsilon L(\boldsymbol{\ell}/\epsilon)}$ by

$$e^{-\epsilon L(\boldsymbol{\ell}/\epsilon)} = \int d^n \mathbf{p} \, F(\mathbf{p}, \epsilon) \, e^{i\mathbf{p} \cdot \boldsymbol{\ell}}. \tag{71}$$

The intermediate integrations in Eq. (70) now lead to a series of Dirac delta functions allowing us to determine the spatial Fourier transform of the propagator in the form

$$G(\mathbf{p}) = C(N, \epsilon) \left[F(\mathbf{p}, \epsilon) \right]^N \tag{72}$$

where $C(N, \epsilon)$ takes care of the integration measure and other numerical constants. We now have to take the limit $\epsilon \to 0$, $N \to \infty$ with $N\epsilon = t$. If such a limit exists for a suitable choice of $C(N, \epsilon)$, then we have succeeded in defining the path integral. As we will see, this works for a non-relativistic particle, but not for a relativistic particle.

Let us first consider the non-relativistic case, for which the relevant Fourier transform in Eq. (71) is given by

$$F(\mathbf{p}) = \int d^n \boldsymbol{\ell} \, \exp \left(-i \mathbf{p} \cdot \boldsymbol{\ell} - \frac{m}{2\epsilon} \boldsymbol{\ell}^2 \right)$$
$$= \left(\frac{2\pi\epsilon}{m} \right)^{n/2} \exp \left(-\frac{\epsilon \mathbf{p}^2}{2m} \right). \tag{73}$$

Therefore, the Fourier transform of the discretized path integral is given by

$$G(\mathbf{p}) = C(N, \epsilon)(F)^N$$
$$= C(N, \epsilon) \left(\frac{2\pi\epsilon}{m} \right)^{nN/2} \exp \left[-\frac{\mathbf{p}^2}{2m}(N\epsilon) \right]. \tag{74}$$

We now see that the exponential factor has a finite limit when $N\epsilon = t_b - t_a$. The pre-factor can be made unity by choosing $C(N, \epsilon) = (2\pi\epsilon/m)^{-nN/2}$. We will then find the continuum limit of the propagator to be the one in Eq. (16). There are no surprises at all.

Let us next consider the relativistic case. The conventional action functional for a relativistic particle, analytically continued to the Euclidean sector, is given by

$$A_E = -m \int \sqrt{\delta_b^a \mathrm{d}x_a \mathrm{d}x^b} = -m \int_{t_1}^{t_2} \mathrm{d}t \sqrt{1 + \mathbf{v}^2}. \tag{75}$$

The relevant Fourier transform in Eq. (71) becomes

$$F(\mathbf{p}) = \int d^n \boldsymbol{\ell} \, \exp[-m(\epsilon^2 + \boldsymbol{\ell}^2)^{1/2} - i \mathbf{p} \cdot \boldsymbol{\ell}]$$

$$= \left(\frac{m}{2\pi} \right)^{1/2} \left(\frac{2\pi}{m} \right)^{n/2} \int_0^\infty \frac{d\mu}{\sqrt{\mu}}$$
$$\times \mu^{n/2} \exp \left(-\frac{\mu}{2m} \omega_p^2 - \frac{m}{2\mu} \epsilon^2 \right) \tag{76}$$

where $\omega_p^2 \equiv \mathbf{p}^2 + m^2$. The integral can be expressed in terms of McDonald functions, leading to

$$F(\mathbf{p}) = \left(\frac{2\pi}{m} \right)^{(n-1)/2} 2$$
$$\times \left(-\frac{m^2\epsilon^2}{\omega_p^2} \right)^{(n+1)/4} e^{-(i\pi/4)(n+1)} K_{-(n+1)/2}(\omega_p \epsilon). \tag{77}$$

We, however, only need its form for small ϵ; in this limit, this expression becomes

$$F(\mathbf{p}) = 2m(4\pi)^{(n-1)/2} \Gamma \left(\frac{n+1}{2} \right) \left(\frac{1}{\omega_p^2} \right)^{(n+1)/2}, \tag{78}$$

which can also be obtained directly from Eq. (76). Therefore, the Fourier transform of the discretized path integral for the relativistic case is given by

$$G(\mathbf{p}) = C(N, \epsilon) \left[F(\mathbf{p}) \right]^N \propto \frac{C(N, \epsilon)}{(\mathbf{p}^2 + m^2)^{(n+1)N/2}}. \tag{79}$$

We again need to take the limit of $N \to \infty$, $\epsilon \to 0$ with $N\epsilon = t$ in this expression and obtain a finite result. It is clear that one cannot obtain a finite result for any choice of the measure $C(N, \epsilon)$. Therefore the straightforward approach to obtain the propagator fails.

The algebraic reason for the different results in the case on non-relativistic and relativistic cases can be traced to the structure of the integrands in Eqs. (73) and (76). Reintroducing the c-factors, as occurring in the combination $c\Delta t = c\epsilon$, we note that the discretized action in the relativistic case has the combination $mc(c^2\epsilon^2 + \boldsymbol{\ell}^2)^{1/2}$. If we first take the $c \to \infty$ limit in this expression, keeping ϵ finite – which is what we do to get the non-relativistic result – this gives $mc^2\epsilon + (1/2)m(\boldsymbol{\ell}^2/\epsilon)$ and the Fourier transform leads to the result in Eq. (73) except for a finite, irrelevant, phase $-imc^2t$, in the Lorentzian sector. But if you take the $\epsilon \to 0$ limit first, keeping c finite – which is what we do in the exact relativistic case – the action $mc(c^2\epsilon^2 + \boldsymbol{\ell}^2)^{1/2}$ becomes $mc|\boldsymbol{\ell}|$ leading to the result in Eq. (78). So the fact that $c\epsilon$ goes to either infinity or zero, depending on whether you take the $c \to \infty$ limit first or the $\epsilon \to 0$ limit first, makes all the difference.

There is another crucial feature which is worth mentioning. If you take the propagator in NRQM, given by Eq. (16), and consider its limit when the time interval $t = \epsilon \to 0$, you find that the argument of the exponential factor is precisely equal to the non-relativistic action; that is, in this limit the propagator has the factor $\exp[i\epsilon L(|\mathbf{x}_2 - \mathbf{x}_1|/\epsilon)]$. So, the

propagator for a finite interval can indeed be thought of[22] as arising from a product of infinitesimal propagators. But this result does not generalize to the relativistic propagator. The infinitesimal form of the relativistic propagator is *not* related in any simple manner to the exponential of the action for infinitesimally separated events. This is again closely related to the composition laws obeyed by the two propagators. The composition law in Eq. (33) can be iterated repeatedly allowing the G_{NR}, for a finite interval of time, to be expressed as an integral over the products of the propagators for infinitesimal time separations. Since the relativistic propagator does not obey this composition law, you cannot do this in a straightforward manner.

Thus, while the Lagrangian and Hamiltonian path integrals lead to the same result in the NRQM, they differ widely for a relativistic action. The standard approach leads to a nonsensical result in the case of the Lagrangian path integral, while the Hamiltonian path integral measure has to be chosen carefully to lead to a Lorentz invariant result.

Why do the two approaches lead to different results? The Lagrangian and Hamiltonian path integrals will lead to the same result only if – in the discretized version – the integrals over $\boldsymbol{p}$ in the Hamiltonian path integral lead to the corresponding (discretized) Lagrangian form of the action.[23] So this equivalence will hold only if the following condition holds:

$$\int d^n \boldsymbol{p}\, M(\boldsymbol{p})\, e^{i\boldsymbol{p}\cdot\boldsymbol{\ell} - i\epsilon H(\boldsymbol{p})} = f(\epsilon)\, e^{i\epsilon L(\boldsymbol{\ell}/\epsilon)} \quad (80)$$

where $M(\boldsymbol{p})$ is some measure in momentum space and $f(\epsilon)$ is a measure for the Lagrangian path integral. So if the functions $M(\boldsymbol{p})$ and $f(\epsilon)$ exist, then the two procedures will give the same final result. This happens for the non-relativistic action but not for the relativistic action.

The time-slicing procedure to define the (Hamiltonian or Lagrangian) path integral automatically selects a class of paths which satisfy the following condition: any path which is included in the sum cuts the intermediate time slices at only one point. That is, you only sum over paths which are always going forward (or always going backwards) in time. In either case, it seems reasonable to interpret the expression in Eq. (67) with a $\theta(t)$ [or a $\theta(-t)$] factor. But, as we mentioned earlier, $\theta(t) G_+(x)$ is not Lorentz invariant. In fact the whole idea of choosing paths which go only forward in time is not a Lorentz invariant criterion when the events x_2 and x_1 are separated by a space-like interval. We will see in the next

section that using the lattice regularization procedure to give meaning to the path integral bypasses these issues.

5 Lattice regularization of the path integral

So far we have seen that: (a) The Hamiltonian path integral can be made to give the propagator $G_+(x)$ with a specific choice of measure, while (b) the straightforward way of computing the Lagrangian path integral does not work. Interestingly enough, there is another way to define the Lagrangian path integral for the relativistic particle based on a geometric interpretation of the relativistic action functional. This is based on a lattice regularization procedure and leads to the Feynman propagator (with $x^2 = x_a x^a$):

$$
\begin{aligned}
G_R(x) &= \int \frac{d^D p}{(2\pi)^D} \frac{i e^{-i p_a x^a}}{(p^2 - m^2 + i\epsilon)} \\
&= \frac{m}{4\pi^2 i \sqrt{x^2}} K_1(im\sqrt{x^2}),
\end{aligned} \quad (81)
$$

which is more relevant to standard QFT than $G_+(x)$. I will briefly describe how this result comes about. (More details of this approach are available in e.g. Ref. [2, Section 1.6.2].)

We will again work in the Euclidean space of D dimensions, evaluate the path integral and analytically continue to the Lorentzian space at the end. The Euclidean action in Eq. (75) can be expressed in the form

$$A_E = -m \int_a^b (\mathrm{d}t^2 + dx^2)^{1/2} = -m \int_a^b d\ell \equiv -m\,\ell \quad (82)$$

where $\ell(x_b, x_a)$ is the length of a path connecting the events $\mathcal{A}$ and $\mathcal{B}$. Our aim is to give meaning to the sum over paths

$$G_R(\mathbf{x_2}, \mathbf{x_1}; m) = \sum_{\text{all } \mathbf{x}(s)} \exp -m\,\ell[\mathbf{x}(s)] \quad (83)$$

in the Euclidean sector, where $\ell(\mathbf{x_2}, \mathbf{x_1})$ is just the Euclidean length of a path, connecting $\mathbf{x_1}$ and $\mathbf{x_2}$. (We will use $\mathbf{x}$ to denote the position in D-dimensional Euclidean space, in contrast to $\boldsymbol{x}$, which was used earlier for the position in the $n = D - 1$ dimensional space in Lorentzian spacetime. We will also label the $D = n + 1$ axes as $(x^1, x^2, \ldots x^j, \ldots x^D)$ with no x^0 axis.) This sum can be given a meaning through the following limiting procedure.

Consider a lattice of points in a D-dimensional cubic lattice with a uniform lattice spacing of ϵ. We will work out G in the lattice and will then take the limit of $\epsilon \to 0$ with a suitable measure. To obtain a finite answer, we have to use an overall normalization factor $M(\epsilon)$ in Eq. (83) and treat m (which is the only parameter in the problem) as varying with ϵ in a specific manner; i.e. we will use a function $\mu(\epsilon)$ in place of m on the lattice and will reserve the symbol m for the

---

[22] This idea works even in the presence of a potential for a Lagrangian of the form $L = (1/2)m\dot{\boldsymbol{x}}^2 - V(\boldsymbol{x})$ as first noted by Dirac, thereby paving the way for the path integral description of QM.

[23] As it turns out, this does happen in the case of NRQM, creating the myth that somehow this should always happen. This is far from true and does not, in general, hold even for an arbitrary 'free particle' Hamiltonian $H = H(\boldsymbol{p})$, let alone in the presence of interactions.

parameter in the continuum limit.[24] Thus the sum over paths in the continuum limit is *defined* by the limiting procedure:

$$G_R(\mathbf{x_2}, \mathbf{x_1}; m) = \lim_{\epsilon \to 0} [M(\epsilon)\mathcal{G}_E(\mathbf{x_2}, \mathbf{x_1}; \mu(\epsilon))] \quad (84)$$

where $\mathcal{G}_E(\mathbf{x_2}, \mathbf{x_1}; \mu(\epsilon))$ is the sum defined on a finite lattice with spacing ϵ.

In a lattice the sum can be evaluated in a straightforward manner. Because of the translation invariance of the problem, $\mathcal{G}_E$ can only depend on $\mathbf{x_2} - \mathbf{x_1}$; so we can set $\mathbf{x_1} = 0$ and call $\mathbf{x_2} = \epsilon\mathbf{R}$ where $\mathbf{R}$ is a D-dimensional vector with integral components: $\mathbf{R} = (n_1, n_2, n_3 \cdots n_D)$. Let $C(N, \mathbf{R})$ be the number of paths of length $N\epsilon$ connecting the origin to the lattice point $\epsilon\mathbf{R}$. Since all such paths contribute a term $[\exp -\mu(\epsilon)(N\epsilon)]$ to Eq. (83), we get

$$\mathcal{G}_E(\mathbf{R}; \epsilon) = \sum_{N=0}^{\infty} C(N; \mathbf{R}) \exp(-\mu(\epsilon)N\epsilon). \quad (85)$$

It can be shown from elementary combinatorics (see, e.g., Sect. 1.6.2 of Ref. [2]) that the $C(N; \mathbf{R})$ satisfies the condition

$$F^N \equiv \left[\sum_{j=1}^{D} 2\cos k_j\right]^N = \sum_{\mathbf{R}} C(N; \mathbf{R})e^{i\mathbf{k}.\mathbf{R}}. \quad (86)$$

Therefore,

$$\sum_{\mathbf{R}} e^{i\mathbf{k}.\mathbf{R}}\mathcal{G}_E(\mathbf{R}; \epsilon) = \sum_{N=0}^{\infty}\sum_{\mathbf{R}} C(N; \mathbf{R})e^{i\mathbf{k}.\mathbf{R}} \exp(-\mu(\epsilon)N\epsilon)$$
$$= \sum_{N=0}^{\infty} e^{-\mu(\epsilon)\epsilon N} F^N = \left[1 - Fe^{-\mu(\epsilon)\epsilon}\right]^{-1}. \quad (87)$$

Inverting the Fourier transform, we get

$$\mathcal{G}_E(\mathbf{R}; \epsilon) = \int \frac{d^D\mathbf{k}}{(2\pi)^D} \frac{e^{-i\mathbf{k}.\mathbf{R}}}{(1 - e^{-\mu(\epsilon)\epsilon} F)}$$
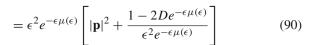$$= \int \frac{d^D\mathbf{k}}{(2\pi)^D} \frac{e^{-i\mathbf{k}.\mathbf{R}}}{(1 - 2e^{-\mu(\epsilon)\epsilon}\sum_{j=1}^{D}\cos k_j)}. \quad (88)$$

Converting to the physical length scales $\mathbf{x} = \epsilon\mathbf{R}$ and $\mathbf{p} = \epsilon^{-1}\mathbf{k}$ gives

$$\mathcal{G}_E(\mathbf{x}; \epsilon) = \int \frac{\epsilon^D d^D\mathbf{p}}{(2\pi)^D} \frac{e^{-i\mathbf{p}.\mathbf{x}}}{(1 - 2e^{-\mu(\epsilon)\epsilon}\sum_{j=1}^{D}\cos p_j\varepsilon)}. \quad (89)$$

This is an exact result in the lattice and we now have to take the limit $\epsilon \to 0$ in a suitable manner to keep the limit finite. As $\epsilon \to 0$, the denominator of the integrand becomes

$$1 - 2e^{-\epsilon\mu(\epsilon)}\left(D - \frac{1}{2}\epsilon^2|\mathbf{p}|^2\right)$$

---

[24] Purely from dimensional analysis, we would expect the mass parameter $\mu(\epsilon)$ to scale inversely as lattice spacing; we will see that this is what happens.

$$= \epsilon^2 e^{-\epsilon\mu(\epsilon)}\left[|\mathbf{p}|^2 + \frac{1 - 2De^{-\epsilon\mu(\epsilon)}}{\epsilon^2 e^{-\epsilon\mu(\epsilon)}}\right] \quad (90)$$

so that we get, for small ϵ,

$$\mathcal{G}_E(\mathbf{x}; \epsilon) \simeq \int \frac{d^D\mathbf{p}}{(2\pi)^D} \frac{A(\epsilon)e^{-i\mathbf{p}.x}}{|\mathbf{p}|^2 + B(\epsilon)} \quad (91)$$

where $A(\epsilon) = \epsilon^{D-2}e^{\epsilon\mu(\epsilon)}$ and $B(\epsilon) = (1/\epsilon^2)[e^{\epsilon\mu(\epsilon)} - 2D]$. The continuum theory has to be defined in the limit of $\epsilon \to 0$ with some measure $M(\epsilon)$; that is, we want to choose $M(\epsilon)$ such that the limit

$$G(\mathbf{x}) = \lim_{\epsilon \to 0} \{M(\epsilon)\mathcal{G}_E(\mathbf{x}; \epsilon)\} \quad (92)$$

is finite. It is easy to see that we only need to demand near $\epsilon \approx 0$, the validity of the conditions:

$$\mu(\epsilon) \approx \frac{\ln 2D}{\epsilon} + \frac{m^2}{2D}\epsilon \approx \frac{\ln 2D}{\epsilon}, \qquad M(\epsilon) = \frac{1}{2D}\frac{1}{\epsilon^{D-2}}. \quad (93)$$

With this choice, we get

$$G_R(\mathbf{x}) = \lim_{\epsilon \to 0}\mathcal{G}_E(\mathbf{x}; \epsilon)M(\epsilon) = \int \frac{d^D\mathbf{p}}{(2\pi)^D}\frac{e^{-i\mathbf{p}.x}}{|\mathbf{p}|^2 + m^2}, \quad (94)$$

which is the usual (Euclidean) Feynman propagator now obtained from a path integral using a lattice regularization. On analytic continuation to the Lorentzian sector, it gives the expression in Eq. (81). So we have succeeded in defining the relativistic path integral and evaluating it to give the Feynman propagator. We will now highlight several aspects of this approach.

5.1 Comments on the lattice regularization approach

The scaling of $\mu(\epsilon) = \ln 2D/\epsilon$ might appear quite strange and I will provide two alternative routes to this scaling which might demystify it a little bit. First one proceeds as follows: Let $\mathcal{N}(\ell)$ be the number of paths of length ℓ connecting the origin to the event $\mathbf{x}$ in the continuum limit. Then our propagator is given by

$$G(\mathbf{x}) = \int_0^{\infty} d\ell \, \mathcal{N}(\ell; \mathbf{x}) \, e^{-m\ell}. \quad (95)$$

This expression, which is the continuum analog of Eq. (85), is only a formal expression, since $\mathcal{N}(\ell; \mathbf{x})$ is divergent in the continuum limit. To give meaning to this equation we have to define $\mathcal{N}(\ell)$ on a lattice with spacing ϵ and take the appropriate limit after the integral is performed. We also need to replace m by the mass parameter $\mu(\epsilon)$ in the lattice. The Fourier transform of $\mathcal{N}_\epsilon(\ell)$ on the lattice is then given by Eq. (86). Switching to the continuum with the replacements $\mathbf{x} = \epsilon\mathbf{R}$ and $\mathbf{p} = \mathbf{k}/\epsilon$, it is easy to see that

$$\mathcal{N}_\epsilon(\ell; \mathbf{p}) \equiv \int d^D\mathbf{x} \, \mathcal{N}_\epsilon(\ell; \mathbf{x}) \, e^{i\mathbf{p}.\mathbf{x}} \simeq (2D - \epsilon^2\mathbf{p}^2)^{\ell/\epsilon} \quad (96)$$

where we have set $N \approx \ell/\epsilon$. Taking the Fourier transform of Eq. (95), using Eq. (96) and performing the integral over ℓ, we find that

$$G(\mathbf{p}; \epsilon) = -\frac{2D}{\epsilon} \left[\mathbf{p}^2 + \frac{2D}{\epsilon} \mu - \frac{2D}{\epsilon^2} \ln 2D \right]^{-1}. \qquad (97)$$

If we now assume that $\mu(\epsilon)$ scales as in the first equation of Eq. (93), the expression in the square bracket in Eq. (97) reduces to $\mathbf{p}^2 + m^2$. The overall factor in front can be taken care of by a suitable measure $M(\epsilon)$. You see that the $(\log 2D)/\epsilon$ scaling of $\mu(\epsilon)$ arises[25] due to the pre-factor $(2D)^{\ell/\epsilon}$ in Eq. (96).

The second approach to understanding the scaling $\mu\epsilon \approx \ln 2D$, which is of interest in its own right, is to think of the propagator as a solution to the KG equation with a delta function source and compare the versions in the continuum and in the lattice. Let us consider a path of N steps connecting the origin to a lattice site labeled by integer valued lattice points $\mathbf{n}$. Then the lattice propagator is given by the sum over all paths of the form

$$\mathcal{G}_{\mathbf{n}} = \sum_{\text{paths}} \exp(-m\epsilon N) \equiv \sum_{\text{paths}} K^N \qquad (98)$$

where $K = e^{-m\epsilon}$. We now interpret K as the probability (amplitude) for the particle to hop between two nearby cells of the lattice. This immediately allows us to write the recurrence relation to reach a specific lattice point $\mathbf{n}$ as

$$\mathcal{G}_{\mathbf{n}} = \delta_{\mathbf{0},\mathbf{n}} + K \sum_{j=1}^{D} (\mathcal{G}_{\mathbf{n}+\mathbf{m}_j} + \mathcal{G}_{\mathbf{n}-\mathbf{m}_j}) \qquad (99)$$

where $\mathbf{m}_j$ is the unit vector in the jth direction. This recurrence relation determines the lattice propagator. On the other hand, in the continuum limit the propagator satisfies the Klein–Gordon equation with a delta function source: The lattice version of this differential operator can easily be obtained by using the Taylor series relation

$$G(x + h) + G(x - h) - 2G(x) = h^2 G''(x) \qquad (100)$$

for each direction. Converting this relation into a lattice with lattice spacing ϵ, the discretized Klein–Gordon equation for the propagator becomes

---

[25] You might wonder whether one can give meaning to the Lagrangian path integral which led to the result in Eq. (79) by allowing m to depend on ϵ. Unfortunately this does not give the correct relativistic propagator. The best one can do is the following. If we assume that $m \to \infty$ as $\epsilon \to 0$, then Eq. (79) will reduce to an expression proportional to $\exp[-(p^2/2m^2)((n+1)t/\epsilon)]$ for a suitable choice of $C(N, \epsilon)$. The best we can do is to assume a scaling of the form $m^2(\epsilon) = [m_0(n+1)/\epsilon]$ where m_0 is a constant. This will lead to an expression proportional to $\exp[-(p^2/2m_0)t]$. This expression is finite and, rather curiously, results in the propagator for NRQM but not the correct one we want.

$$\frac{1}{\epsilon^2} \sum_{j=1}^{D} (\mathcal{G}_{\mathbf{n}+\mathbf{m}_j} + \mathcal{G}_{\mathbf{n}-\mathbf{m}_j} - 2\mathcal{G}_{\mathbf{n}}) - m^2 \mathcal{G}_{\mathbf{n}} = \delta_{\mathbf{0},\mathbf{n}}. \qquad (101)$$

This equation can be rewritten in the form

$$\left(\frac{1}{m^2\epsilon^2 + 2D} \right) \sum_{j=1}^{D} (\mathcal{G}_{\mathbf{n}+\mathbf{m}_j} + \mathcal{G}_{\mathbf{n}-\mathbf{m}_j}) + \delta_{\mathbf{0},\mathbf{n}} = \mathcal{G}_{\mathbf{n}}, \qquad (102)$$

where we have rescaled the Dirac delta function by ϵ^2 on the lattice. Comparing Eq. (99) with Eq. (102), we see that $\exp(m_0\epsilon)$ gets replaced by $(m^2\epsilon^2 + 2D)$ on the lattice. This is equivalent to the replacement of m_0 by $\epsilon^{-1} \ln 2D$ in the limit of $\epsilon \to 0$, which is precisely the mass renormalization we saw earlier.

How does it come that the Lagrangian path integral, originally evaluated with the time-slicing method, led to a meaningless expression (viz. Eq. (79)), while the lattice regularization method leads to the Feynman propagator? The reason has to do with the different kinds of paths which are summed over in the two approaches. When you define the path integral by time slicing, you implicitly assume that any path which is included in the sum cuts the intermediate time slices at only one point. That is, you only sum over paths which are always going forward (or always going backwards) in time. But when you sum over paths on the lattice, the paths can go back and forth in time. So the two sets of paths which are summed over are completely different and we have no reason to expect them to give the same answer.

This connection can be made more quantitative by examining a lattice regularization scheme for paths which go only forward in time in the Lorentzian sector. On analytic continuation to the Euclidean sector, they will go only forward in one of the axis, which we take to be the x^0 direction. (We will now label the $D = (n + 1)$ axes as $(x^0, x^1, \ldots x^n)$, restoring the x^0 axis which is treated as special.) Our aim is to see whether such a condition will lead to anything which resembles the non-relativistic propagator.

We know that a relativistic scalar field in the Euclidean sector will satisfy the Euclidean Klein–Gordon equation $(-\Box_E + m^2)\phi = 0$, while its non-relativistic counterpart $f(x)$, related to $\phi(x)$ by $\phi(x) \equiv e^{-mt} f(x)$, will satisfy the Euclidean Schroedinger equation $(\partial_t - (1/2m)\nabla^2)_E f = 0$. The latter is obtained from the former by *approximating* the second time derivative $\ddot{\phi}$ by $\ddot{\phi} \approx m^2\phi - 2me^{-mt} f$. In the momentum space, this involves replacement of $(p^2+m^2)_E \equiv \Omega^2 + \mathbf{p}^2 + m^2$ by $2mi\omega + \mathbf{p}^2$ where $\Omega \equiv \omega + im$ and we have ignored the ω^2 term in comparison with $m\omega$. This requires the denominator $(\Omega^2 + \mathbf{p}^2 + m^2)$ in the Euclidean relativistic propagator (written as a Fourier transform with respect to Euclidean time),

$$G_R = \int \frac{d\Omega \, d^n \mathbf{p}}{(2\pi)^D} \frac{e^{i(\Omega t + \mathbf{p} \cdot \mathbf{x})}}{(\Omega^2 + \mathbf{p}^2 + m^2)}, \qquad (103)$$

given by Eq. (94), to be replaced by $(2mi\omega + \boldsymbol{p}^2)$ to give the non-relativistic propagator:

$$
\begin{aligned}
\bar{G}_{\mathrm{NR}} &= \int \frac{d\omega\, d^n\boldsymbol{p}}{(2\pi)^D} \frac{2m e^{i(\omega t + \boldsymbol{p}\cdot\boldsymbol{x})}}{(2mi\omega + \boldsymbol{p}^2)} \\
&= \int \frac{d\omega\, d^n\boldsymbol{p}}{(2\pi)^D} \frac{(2m) e^{i(\omega t + \boldsymbol{p}\cdot\boldsymbol{x})}}{\boldsymbol{p}^2 + 2mi(\Omega - im)}.
\end{aligned}
\tag{104}
$$

Let us see how this comes about when we restrict paths to go only forward along x^0.

Each of the $2\cos p_j\epsilon = e^{ip_j\epsilon} + e^{-ip_j\epsilon}$ in the denominator of Eq. (89) is contributed by paths going forward along the jth direction (contributing $e^{ip_j\epsilon}$) and paths going backward along the jth direction (contributing $e^{-ip_j\epsilon}$). So, when we restrict the paths to moving only forward along the x^0 axis and repeat the analysis, along the 0 direction we only pick up a $e^{ip_0\epsilon_0}$ factor. This modifies the denominator $\mathcal{D}$ of Eq. (89) to the expression

$$
\mathcal{D} = 1 - 2e^{-\mu(\epsilon_1)\epsilon_1} \sum_1^n \cos p_j\epsilon_1 - e^{\mu(\epsilon_0)\epsilon_0}\, e^{ip_0\epsilon_0}.
\tag{105}
$$

We have taken the lattice spacing to be ϵ_0 along the time direction and ϵ_1 for all the space directions. This is essential because the transition from a Klein–Gordon equation to a Schroedinger equation involves a transition from wave equation to a diffusion equation; the propagation in the Euclidean lattice will mimic a diffusion only if $t \propto x^2$, requiring $\epsilon_0 \propto \epsilon_1^2$, when we take the continuum limit. (If you do not do this and assume the same lattice spacing along both direction you will *not* reproduce the form of the propagator in Eq. (104).) Straightforward computation now reduces Eq. (105) to the form

$$
\mathcal{D} = (A p_0 + \boldsymbol{p}^2 + B)C
\tag{106}
$$

where

$$
C = \epsilon_1^2\, e^{-\mu_1\epsilon_1},
\tag{107}
$$

$$
B = \frac{1}{\epsilon_1^2} e^{\epsilon_1\mu_1} \left(1 - 2n e^{-\epsilon_1\mu_1} - e^{-\epsilon_0\mu_0}\right),
\tag{108}
$$

$$
A = \frac{1}{\epsilon_1^2} e^{\epsilon_1\mu_1} \left(-i\epsilon_0\, e^{-\epsilon_0\mu_0}\right).
\tag{109}
$$

Ignoring the overall constant C – which merely defines the overall measure like $M(\epsilon)$ in the previous analysis – and comparing $\mathcal{D}$ with the denominator in Eq. (104), we find that the following conditions need to be satisfied:

$$
A = 2m_0 i, \qquad B = 2m_0^2.
\tag{110}
$$

Some more algebra now shows that this can indeed be achieved with the choices

$$
\epsilon_0 = \frac{2m}{(2n-1)} \epsilon_1^2 \propto \epsilon_1^2
\tag{111}
$$

and

$$
\mu_0 = -\frac{1}{\epsilon_0} \ln(2n-1), \qquad \mu_1 = 2m^2\epsilon_1.
\tag{112}
$$

Equation (111) shows that $\epsilon_0 \propto \epsilon^2$ has to be expected in a diffusion process; Eq. (112) shows the scaling of μ_1 and μ_0 for this result to hold.

This feature can also be made more transparent along the following lines. While the real space expressions for $G_{\mathrm{NR}}(x)$ (given by Eq. (16)) and $G_R(x)$ (given by Eq. (81)) look very different, their spatial Fourier transforms are very similar:

$$
\begin{aligned}
G_R(t, \boldsymbol{p}) &\equiv \int d^3\boldsymbol{x}\; G(x_2; x_1) e^{-i\boldsymbol{p}\cdot\boldsymbol{x}} \\
&= \begin{cases}
e^{-i\omega_{\boldsymbol{p}} t} & \text{(non-relativistic)} \\[2mm]
\dfrac{1}{2\omega_{\boldsymbol{p}}} e^{-i\omega_{\boldsymbol{p}}|t|} & \text{(relativistic)}
\end{cases}
\end{aligned}
\tag{113}
$$

where $\omega_{\boldsymbol{p}} = \boldsymbol{p}^2/2m$ in the non-relativistic case, while $\omega_{\boldsymbol{p}} = (\boldsymbol{p}^2 + m^2)^{1/2}$ in the relativistic case. Using the Fourier transform of $G_+(x)$ in Eq. (15), it is easy to relate $G_R(x)$ and $G_+(x)$. We find that $G_R(x_2, x_1) = G_+(x_2; x_1)$ when $t_2 > t_1$ and $G_R(x_2; x_1) = G_+^*(x_2; x_1)$ when $t_2 < t_1$ where $G_-(x_2; x_1) \equiv G_+^*(x_2; x_1) = G_+(x_1; x_2)$ is the complex conjugate of $G_+(x_2; x_1)$. That is,[26]

$$
\begin{aligned}
G_R(x_2; x_1) &= \theta(t) G_+(x_2; x_1) + \theta(-t) G_+^*(x_2; x_1) \\
&= \theta(t) G_+(x_2; x_1) + \theta(-t) G_-(x_2; x_1).
\end{aligned}
\tag{114}
$$

Since we know that G_+ uses only paths which go forward in time it is clear that G_R propagates particles with energy $\omega_{\boldsymbol{p}}$ forward in time and propagates particles with energy $-\omega_{\boldsymbol{p}}$ backward in time. This feature arises from summing over paths which go back and forth in the time direction. So, $G_R(x_2, x_1)$ is actually two propagators rolled into one; we will come back to this aspect in Sect. 6.

It is obvious that, while the relativistic propagator G_R in Eq. (94) arises very naturally through the lattice regularization approach, we have to make several artificial choices *based on our hindsight* for obtaining a non-relativistic propagator by lattice regularization. Once again there is no natural limiting process within the lattice regularization which allows us to obtain the non-relativistic propagator from the relativistic one.

---

[26] $\theta(t)$ is Lorentz invariant only when x_2 and x_1 are separated by a time-like interval, in which case the $\theta(t)$ picks out one of the two terms. When x_2 and x_1 are separated by a space-like interval, $\theta(t)$ is not Lorentz invariant and hence the expression could pick either of the two terms depending on the Lorentz frame. But since we know – from any of the explicit expressions like Eq. (116) – that $G_R(x_2; x_1)$ is Lorentz invariant, it follows that $G_+(x_2; x_1) = G_-(x_2; x_1)$ when x_2 and x_1 are separated by a space-like interval. This is indeed true and can be explicitly verified. (See Appendix B.)

5.2 Jacobi action and its path integral

A convenient expression for G_R in the coordinate space is obtained using the Schwinger proper-time representation.[27] We write $(|\boldsymbol{p}|^2 + m^2)^{-1}$ as an integral over λ of $\exp[-\lambda(|\boldsymbol{p}|^2 + m^2)]$ and do the $\boldsymbol{p}$ integration to obtain

$$G_R = \int_0^\infty \frac{d\lambda}{(4\pi\lambda)^{D/2}} \exp\left(-\lambda m^2 - \frac{|\mathbf{x}|^2}{4\lambda}\right)$$
$$\Rightarrow \frac{1}{16\pi^2} \int_0^\infty \frac{d\lambda}{\lambda^2} \exp\left(-m^2\lambda - \frac{|\mathbf{x}|^2}{4\lambda}\right) \quad (115)$$

where the second expression is for $D = 4$. The analytic continuation from the Euclidean to the Lorentzian spacetime changes the sign of one of the coordinates in $|\mathbf{x}|^2$ to give $|\boldsymbol{x}|^2 - t^2 = -x^2$ and we set $\lambda = is$. This gives the final result:

$$G_R = -\frac{i}{16\pi^2} \int_0^\infty \frac{ds}{s^2} \exp\left(-im^2 s - \frac{i}{4s}x^2\right)$$
$$= \frac{m}{4\pi^2 i \sqrt{x^2}} K_1(im\sqrt{x^2}). \quad (116)$$

This proper-time representation of the G_R has an alternative interpretation. The integral expression in Eq. (116) can be expressed, after a rescaling of $s \to s/m$, as

$$G_R(x_2; x_1) \propto \int_0^\infty ds\, e^{-ims} \langle x_2, s | x_1, 0 \rangle$$
$$= C_m \int_0^\infty ds\, e^{-ims} \sum_{x(\tau)} e^{iA[x(\tau)]} \quad (117)$$

where C_m is an unimportant constant and

$$\langle x_2, s | x_1, 0 \rangle = \theta(s) i \left(\frac{m}{4\pi i s}\right)^2 \exp\left(-\frac{i}{4}\frac{mx^2}{s}\right) \quad (118)$$

can be thought of as a propagator for a (fictitious) particle moving in the four dimensional Lorentzian spacetime from x_1^i at $\tau = 0$ to x_2^i at $\tau = s$, where τ parameterizes the path in spacetime $x^i(\tau)$. The relevant action for this particle is a quadratic one, given by

$$A[x(\tau)] = -\frac{1}{4}m \int_0^s d\tau\, \dot{x}_a \dot{x}^a. \quad (119)$$

Classically, this action could *also* be thought of as representing the free relativistic particle (since it leads to the equation of motion $d^2 x^i / d\tau^2 = 0$). But – unlike the action in Eq. (75): (i) it is not reparametrization invariant and (ii) it does not have a geometrical interpretation. The path integral in Eq. (117) gives the amplitude $\langle x_2, s | x_1, 0 \rangle$ for the particle to propagate from x_1 to x_2 during the proper-time interval s. The Fourier transform of this amplitude with respect to s can be thought of as giving the amplitude for this propagation to

occur with the energy mc^2 in the rest frame. This suggests that $G_R(x_2, x_1)$ gives an amplitude for propagation *at a constant energy* rather than for a given *time interval*. Such a path integral can be defined in a more general context using what is known as the Jacobi action functional. We will now discuss this interpretation of the Feynman propagator.

The Jacobi action A_J can be thought of as the integral of $\boldsymbol{p} \cdot d\boldsymbol{x}$ where $\boldsymbol{p}$ is expressed as a function of energy E by solving the equation $H(\boldsymbol{p}) = E$. In our case, for a system with $H(\boldsymbol{p}) = H(|\boldsymbol{p}|)$, the $\dot{\boldsymbol{x}}$ and $\boldsymbol{p}$ will be in the same direction allowing us to write $\boldsymbol{p} \cdot d\boldsymbol{x} = \mathcal{P}(E) d\ell$ where ℓ is the arc-length of the path and $\mathcal{P}(E)$ is the magnitude of the momentum $|\boldsymbol{p}|$, expressed as a function of E. Since E is constant, the Jacobi action in our case reduces to

$$A_J = \mathcal{P}(E) \int d\ell = \mathcal{P}(E)\, \ell(\boldsymbol{x}_b, \boldsymbol{x}_a), \quad (120)$$

which has the geometrical meaning of the length of the path connecting the two events. This expression is manifestly reparameterization invariant with no reference to the time coordinate.

Since A_J describes an action principle for determining the path of a particle with energy E classically, the sum over $\exp(iA_J)$ could be interpreted as the amplitude for the particle to propagate from x_1^α to x_2^α *with energy* E. Since A_J is not quadratic in the velocities, even for a non-relativistic free particle (because $d\ell$ involves a square root), one has to again do a lattice regularization to compute the result, just as we did for a relativistic particle. The path integral defined using the Jacobi action then reduces to the sum over paths of the kind considered in Eq. (83) with m replaced by $\mathcal{P}(E)$. So the propagator for the Jacobi action will be given by the expression obtained earlier in Eq. (94) with m^2 replaced by $\mathcal{P}^2(E)$. That is, the Jacobi action propagator will be

$$\mathcal{G}(\boldsymbol{x}, E) = \sum \exp(-A_J) = \int \frac{d^D \boldsymbol{p}}{(2\pi)^D} \frac{e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}}{p^2 + \mathcal{P}^2(E)}. \quad (121)$$

In the case of a non-relativistic free particle with $\mathcal{P}^2(E) = 2mE$, this gives us the result

$$(2m)\,\mathcal{G}(\boldsymbol{x}, E) = \int \frac{d^D \boldsymbol{p}}{(2\pi)^D} \frac{e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}}{E + (p^2/2m)}, \quad (122)$$

which makes sense.[28]

---

[27] This representation, using a "fifth time" was introduced by Stueckelberg [35,36] and developed further extensively by Schwinger.

[28] The Fourier transform of a Lorentzian propagator $G(t, \boldsymbol{x})$ with respect to Lorentzian time t will be defined using an $\exp(-i\omega t)$ factor, while the Fourier transform of a Euclidean propagator $G_E(t_E, \boldsymbol{x})$ with respect to Euclidean time t_E will be usually defined using an $\exp(-i\omega t_E)$ factor. Sometimes it is more convenient to use the factor $\exp(-\omega t_E)$ (and restrict the time integration to the range $(0, \infty)$) and define the Laplace transform in the Euclidean sector. This is what we have done here; the usual Fourier transform can be obtained by the replacement $E \to -iE$.

But there is another way of determining $\mathcal{G}(\boldsymbol{x}, E)$. Since we already have the standard path integral defined for the non-relativistic particle, we can use it *to give meaning* to this sum over $\exp(-A_J)$. In the process, we would have obtained a procedure for *defining* the sum over paths for any non-quadratic action that is proportional to the length of the path. The idea is to write the sum over all paths in the conventional Lagrangian action principle (with amplitude $\exp(iA_{\boldsymbol{x}})$) as a sum over paths with energy E followed by a sum over all E. So we write, formally,

$$\sum_{0,\boldsymbol{x}_1}^{t,\boldsymbol{x}_2} \exp(iA_{\boldsymbol{x}}) = \sum_E \sum_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} e^{-iEt} \exp iA_J[E, \boldsymbol{x}(\tau)]$$
$$\propto \int_0^\infty dE\, e^{-iEt} \sum_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \exp(iA_J). \qquad (123)$$

In the last step we have treated the sum over E as an integral over $E > 0$ (since, for any Hamiltonian which is bounded from below, we can always achieve this by adding a suitable constant to the Hamiltonian) but there could be an extra proportionality constant which will depend on the measure used to define the sum over $\exp(iA_J)$. Inverting the Fourier transform, we get the Jacobi propagator:

$$\mathcal{G}(\boldsymbol{x}_2, \boldsymbol{x}_1; E) \equiv \sum_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \exp(iA_J) = C \int_0^\infty dt\, e^{iEt} \sum_{0,\boldsymbol{x}_1}^{t,\boldsymbol{x}_2}$$
$$\exp(iA) = C \int_0^\infty dt\, e^{iEt} G(x_2; x_1) \qquad (124)$$

where we have denoted the proportionality constant by C. This result shows that the sum over the Jacobi action A_J involving a *square root of velocities* can be re-expressed in terms of the standard path integral; if the latter can be evaluated for a given system, then the sum over Jacobi action can be defined by this procedure. For the case of a free particle we get

$$\sum_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \exp i\sqrt{2mE}\,\ell(\boldsymbol{x}_2, \boldsymbol{x}_1) = C \int_0^\infty dt\, e^{iEt} \sum_{0,\boldsymbol{x}_1}^{t,\boldsymbol{x}_2}$$
$$\times \exp \frac{im}{2} \int_0^t d\tau \left(g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta \right) \qquad (125)$$

where we have denoted the length of the path connecting x_1^α and x_2^α by $\ell(\boldsymbol{x}_2, \boldsymbol{x}_1)$, Since the action for the relativistic particle in Eq. (75) has the same structure as the Jacobi action for a non-relativistic free particle, the propagator, $G_R(x_2; x_1)$, can be obtained directly from Eq. (125). We first take the complex conjugate of Eq. (125) (in order to get the overall minus sign in the action in Eq. (75)) and generalize the result from space to spacetime, leading to

$$\sum_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \exp -i\sqrt{2mE}\,\ell(\boldsymbol{x}_2, \boldsymbol{x}_1) = C \int_0^\infty d\tau\, e^{-iE\tau} \sum_{0,\boldsymbol{x}_1}^{t,\boldsymbol{x}_2}$$
$$\times \exp -\frac{im}{2} \int_0^\tau d\lambda \left(g_{ab} \dot{x}^a \dot{x}^b \right). \qquad (126)$$

In order to get $-im\ell(\boldsymbol{x}_2, \boldsymbol{x}_1)$ on the left-hand side we take $E = m/2$ and put $\tau = 2s$ to get an $\exp(-ims)$ factor. The path integral over the quadratic action trivial and in $D = 4$, we get the expression in Eq. (118). Therefore the path integral propagator reduces to the expression in Eq. (117):

$$G(x_2; x_1) = -(2Cm)i \left(\frac{m}{16\pi^2} \right) \int_0^\infty \frac{ds}{s^2}$$
$$\times \exp \left(-ims - \frac{i}{4} \frac{mx^2}{s} \right)$$
$$= -\frac{i}{16\pi^2} \int_0^\infty \frac{d\mu}{\mu^2} \exp \left(-im^2\mu - \frac{i}{4} \frac{x^2}{\mu} \right) \qquad (127)$$

where we have rescaled the variable s to μ by $s \equiv m\mu$ and made the choice $C = 1/2m$ to match with conventional result in Eq. (116).

Once we introduce the idea of a fictitious particle propagating in spacetime, governed by a quadratic action in Eq. (119), we can also introduce a complete set of (space-time) position eigenkets $|x\rangle$ and momentum eigenkets $|p\rangle$. The Hamiltonian relevant for the action in Eq. (119) will be $H = -p^2$ (corresponding to the mass $m = 1/2$) and the matrix element of the proper-time evolution operator will be

$$\langle x_2, s | x_1, 0 \rangle \equiv \langle x_2 | e^{-isH} | x_1 \rangle = \langle x_2 | e^{isp^2} | x_1 \rangle. \qquad (128)$$

So, the relativistic propagator G_R, treated as a function of $\mu = m^2$ can be expressed as the integral

$$G_R(x_2, x_1; \mu) \equiv \int_0^\infty ds \langle x_2 | e^{-is(H+\mu)} | x_1 \rangle$$
$$= -i \langle x_2 | (\mu + H)^{-1} | x_1 \rangle. \qquad (129)$$

This result will be useful later on.

Our propagator can also be obtained by using the quadratic action Eq. (119) in the path integral and imposing the reparametrization invariance through a Lagrange multiplier. (This is also equivalent to imposing the condition $H = -p^2 = -m^2$ on the Hamiltonian.) The path integration over the Lagrange multiplier will reduce to integration over τ leading to the same final expression. I will quickly run through this procedure [37–39] to connect with our previous discussion. We begin by recalling that, for the relativistic Lagrangian, $L_R = -m \left[\eta_{mn} \dot{x}^m \dot{x}^n \right]^{1/2} = -m(\dot{x}^2)^{1/2}$, the momenta $p_m = \partial L / \partial \dot{x}^m$ satisfy the constraint $\mathcal{H} \equiv p_m p^m - m^2 = 0$. While constructing the Hamiltonian form of

the action, this constraint is incorporated through a Lagrange multiplier $N(\tau)$, leading to

$$A_R = \int_{r_1}^{r_2} d\tau \left(p_m \dot{x}^m + N\mathcal{H} \right). \tag{130}$$

This action, in turn, retains the memory of re-parameterization invariance of L_R because it remains invariant under the gauge transformation generated by $\mathcal{H}$ given by

$$\delta x = \epsilon(\tau)\{x, \mathcal{H}\}, \quad \delta p = \epsilon(\tau)\{p, \mathcal{H}\}, \quad \delta N = \dot{\epsilon}(\tau) \tag{131}$$

where $\epsilon(\tau)$ vanishes at the end points. The simplest gauge-fixing condition [3] is to take $\dot{N} = 0$, making N a constant. The Hamiltonian path integral will now require an integration over the parameter N, which will lead to the correct propagator G_R if the range of integration is restricted to $0 \leq N < \infty$. That is,

$$G_R(x_2 - x_1) = -i \int_0^\infty dN \int \mathcal{D}p\mathcal{D}x$$
$$\times \exp\left(i \int_{r_1}^{r_2} d\tau \, (p\dot{x} + N\mathcal{H}) \right). \tag{132}$$

This is yet another popular route to the Feynman propagator discussed in the literature.

To avoid possible misunderstanding, I stress the following fact. It is certainly possible to come up with schemes by which the path integral for a relativistic particle can be evaluated. We have already seen three such procedures which lead to the "correct" propagator, $G_F(x)$: (a) lattice regularization, (b) the Jacobi action method, and (c) the gauge-fixing approach. (The approach based on lattice regularization or the one based on the interpretation of Eq. (75) as a Jacobi action seems more transparent than the one in which gauge fixing is used, but this could be a matter of taste.) The key common feature is that all 'successful' approaches – which lead to the 'correct' $G_F(x)$ – allow for the paths to go backwards and forwards in Minkowski time coordinate t, which will not be allowed in the standard time-slicing approach to a path integral based on the Lagrangian $L_R = -m(1 - \dot{\mathbf{x}}^2)^{1/2}$. In fact, the class of paths summed over in each of the three approaches are formally very different. For example, it is certainly true that the Lagrangian $L_R(\dot{\mathbf{x}})$ arises from the 'gauge-invariant' Lagrangians, in a specific gauge. But the path integral involves the sum over totally different sets of paths ($x^a(\tau)$ versus $\mathbf{x}(t)$) in these two approaches; summing over $x^a(\tau)$ with time slicing in τ allows for paths $\mathbf{x}(t)$ which go backwards in the Minkowski time coordinate t.

The existence of these three (and possibly many other) procedures does not provide the answer to the simple question: *How does it come that the most natural procedure, based on paths $\mathbf{x}(t)$ and time slicing in t, which works so well in the case of a NR particles, fails for a relativistic particle?*

Given the action for a *non-relativistic* particle I can construct NRQM by path integral, just with time slicing, without knowing the Schroedinger equation or the Heisenberg operator algebra. But given the action for the *relativistic* particle, I cannot do it in a natural fashion and, in fact, the corresponding single-particle RQM does not exist! Of course, if you think of relativistic particles as excitations of an underlying field and quantize the field – rather than use the action principle for the particle – you will get $G_F(x)$ as well as the antiparticles. You can then cook up several ways to get it from path integrals; that is hardly satisfactory if you want to do everything upfront from the path integral.

The technical reason, as we will see in Sect. 6, has to do with the fact that $G_F(x)$ actually propagates *two* fields and *two* kinds of particles, not one. The procedures which actually "work," for defining the path integrals, have this feature built into them one way or another – usually by allowing paths to go backwards and forwards in Minkowski time coordinate t – so that they can lead to the "correct" propagator, $G_F(x)$. This is hardly a satisfactory situation because we already need to know the answer (and the existence of pairs of particles) from some other approach to define the suitable procedure for the path integral. I will say more about this in Sect. 8 in the text surrounding Eq. (194).

I conclude this section with a technical comment related to the time-slicing approach for determining the relativistic propagator which, as we saw earlier, does not work. In Eq. (117), the amplitude $\langle x_2, s | x_1, 0 \rangle$ has a natural path integral expression with time slicing in the proper time s. If we divide the proper-time interval into N slices and write the usual time-sliced expression for $\langle x_2, s | x_1, 0 \rangle$ in Eq. (118), we can write the relativistic propagator in Eq. (117) in the form

$$G_R = \int_0^\infty ds \int \prod_{n=1}^{N-1} i d^4 x_n \left(\frac{mN}{4\pi i s} \right)^2$$
$$\times \exp\left(-i \sum_{n=0}^{N-1} \frac{m(x_{n+1} - x_n)^2}{4s/N} - ims \right). \tag{133}$$

In the absence of the integration over s, the propagator $\langle x_2, s | x_1, 0 \rangle$ satisfies the non-relativistic composition law in Eq. (33). But once we introduce the integration over s, this composition law fails and – as we will see later in Eq. (158) – is replaced by a composition law involving the Klein–Gordon inner product. The crucial point is that the expression in Eq. (133), after integration of s, is not related to the exponential of the infinitesimal action. If we define the sum in the exponent as

$$R^2 \equiv \frac{N}{4} \sum_{n=0}^{N-1} (x_{n+1} - x_n)^2 \tag{134}$$

then the integration over s will lead to a weight for each path given by (with $s = m\bar{s}$):

$$
\begin{aligned}
W &= \int_0^\infty d\bar{s}\, \bar{s}^{2-2N} \exp\left(-\frac{iR^2}{\bar{s}} - im^2\bar{s}\right) \\
&= 2\left(\frac{-R^2}{m^2}\right)^{\nu/2} e^{-i\pi\nu/2} K_{-\nu}(2miR)
\end{aligned}
\tag{135}
$$

with $\nu = 3 - 2N$. Obviously, this expression has no simple relation with the exponential of relativistic action.

5.3 Non-relativistic limit of the Feynman propagator

We have obtained the propagator G_+ using the Hamiltonian path integral and G_R from the lattice regularization of the Lagrangian path integral. We already know that G_+ is related to G_{NR} through the limit in Eq. (19), which is not very surprising because we could derive both G_+ and G_{NR} at one go, in Eq. (14). But the derivation of G_R, using the lattice regularization, was quite different and there is no simple correspondence to G_{NR}. So the question arises as to whether one can get the non-relativistic propagator $G_{NR}(x_b, x_a)$ from the relativistic propagator $G_R(x_b, x_a)$ in the limit of $c \to \infty$.
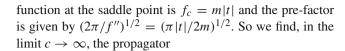
This is *not* possible in spite of occasional claims to the contrary made in the literature. This should, in fact, be obvious from Eq. (113). When you take $c \to \infty$ limit of $G_R(t, \boldsymbol{p})$, the pre-factor becomes $1/2m$, which is an inconsequential scaling. In the phase ω_p can be approximated as $m + \boldsymbol{p}^2/2m$. The factor $\exp(-imt)$ could have been interpreted as due to the rest energy mc^2 contributing to the phase. *But the $|t|$ never becomes t when we take this limit.*[29] So the factor $\exp(-m|t|)$ does not have a straightforward interpretation. Thus, while we can barely escape[30] in the case of $t > 0$, the expressions are quite different for $t < 0$.

To see this more explicitly, we have only have to evaluate G_R in the $c \to \infty$ limit using the saddle point approximation to the integral. Rescaling $\lambda \to \lambda/m$ we can express the Euclidean G_R in the form

$$
\begin{aligned}
G &= \frac{1}{(4\pi)^2} \int_0^\infty \frac{d\lambda}{\lambda^2} e^{-\lambda m^2} e^{-(1/4\lambda)(t^2+x^2)} \\
&\to \frac{m}{(4\pi)^2} \int_0^\infty \frac{d\lambda}{\lambda^2} e^{-m\lambda - (m/4\lambda)(t^2+x^2)}.
\end{aligned}
\tag{136}
$$

We need the saddle point of the function $f(\lambda) = m\lambda + (mt^2/4\lambda)$, which occurs at $\lambda = \lambda_c = |t|/2$. The value of the

---

[29] The Feynman propagator is an even function of t and it will remain an even function of t even when you take the limit $c \to \infty$. On the other hand, $G_{NR}(t, \boldsymbol{x})$ is not an even function of t since $G_{NR}(-t, \boldsymbol{x}) = G_{NR}^*(t, \boldsymbol{x})$. So you cannot get G_{NR} from G_R for all t.

[30] Equation (19) tells you that, for $t > 0$, we have $G_R = G_+$ and we already know the relation, given by Eq. (19), between G_+ and G_{NR}. So this is not a big deal. The real issue is in the comparison of G_R and G_{NR} for $t < 0$ where they do not match.

function at the saddle point is $f_c = m|t|$ and the pre-factor is given by $(2\pi/f'')^{1/2} = (\pi|t|/2m)^{1/2}$. So we find, in the limit $c \to \infty$, the propagator

$$
G = \frac{1}{2m}\left(\frac{m}{2\pi|t|}\right)^{3/2} e^{-m|t| - mx^2/2|t|}.
\tag{137}
$$

The overall scaling by $2m$ is of no consequence and arises from $2\omega_p$ in the limit $c \to \infty$. But you find that the result has $|t|$ rather than t in the expression. When you analytically continue to the Lorentzian sector, this effect will persist and you will get

$$
G = \frac{1}{2m}\left(\frac{m}{2\pi i|t|}\right)^{3/2} e^{-im|t| + imx^2/2|t|}.
\tag{138}
$$

One can understand the factor $\exp(-imc^2|t|)$ as signaling the rest energy mc^2 of the particle, which has to be taken away from the phase of the wave function, to reach the non-relativistic limit when $t > 0$. But one cannot make sense of this phase for $t < 0$; more generally we cannot interpret the occurrence of $|t|$ in NRQM. This issue is actually quite non-trivial and we will discuss it again in Sect. 6 from a different perspective.

The usual folklore that the Feynman propagator has the correct NRQM limit originates: (i) either from considering only $t > 0$ case, (ii) or from mixing up momentum space and real space descriptions. The momentum space argument goes along the following lines: In momentum space, the Feynman propagator is governed by a term in the denominator $(p^2 - m^2 - i\epsilon)$ where $p^a = (E, \boldsymbol{p})$. If we write $E \equiv m + \epsilon$ removing the rest energy, then $p^2 - m^2 = \epsilon^2 + 2m(\epsilon - \boldsymbol{p}^2/2m)$ and when we study processes involving non-relativistic energies, one can ignore the ϵ^2 term and use the approximate expression proportional to $(\epsilon - \boldsymbol{p}^2/2m)$ in the *momentum* space. This approximation completely changes the pole structure of propagator from two poles in the complex plane to one. To get the real space propagator from the momentum space propagator, you need to integrate over all ϵ *without* ignoring the ϵ^2 term. Making the approximation $\epsilon \ll 1$, obtaining an approximate momentum space propagator and *then* integrating over *all* ϵ to get the real space propagator is conceptually incorrect.

5.4 Feynman propagator as a matrix element of time evolution operator

The non-relativistic propagator G_{NR} can be expressed as the matrix element $G(x_b, x_a) = \langle x_b | e^{-itH} | x_a \rangle$ of the time evolution operator in a straightforward manner. In the case of G_+, we could again do this but the states $|x\rangle$ did not have the interpretation as eigenstates of the position operator; instead we had to define them using a Fourier transform. Let us now address the corresponding question for the relativistic propa-

gator $G_R(x_b, x_a)$ obtained above from lattice regularization, viz., whether it can be expressed in the form $G(x_b, x_a) = \langle x_b | e^{-itH} | x_a \rangle$ where $H = H(\boldsymbol{p}) = (\boldsymbol{p}^2 + m^2)^{1/2}$. We already know that this is not going to happen with the procedure we have adopted for defining the states $|x\rangle$; it only leads (at best) to G_+ and $G_R \neq G_+$. Obviously we have to cheat a little bit somewhere along the line if such a relation should hold. I will now describe how this can be achieved (with a bit of cheating) because the procedure highlights some key issues we have been discussing.

To do this, we will first consider the case when $t > 0$ and use the easily proved (operator) identity

$$2H \int_0^\infty d\mu \, \exp\left(-i\mu^2 H^2 - \frac{it^2}{4\mu^2}\right) = \left(\frac{\pi}{i}\right)^{1/2} e^{-iHt}, \tag{139}$$

which allows us to write

$$\begin{aligned}
\langle x_b | e^{-iHt} | x_a \rangle &= \left(\frac{i}{\pi}\right)^{1/2} \int_0^\infty d\mu \\
&\quad \times e^{(-it^2/4\mu^2)} \langle x_b | 2H(\boldsymbol{p}) e^{-i\mu^2 H^2(\boldsymbol{p})} | x_a \rangle \\
&= \left(\frac{i}{\pi}\right)^{1/2} \int_0^\infty d\mu \\
&\quad \times e^{(-it^2/4\mu^2)} e^{-i\mu^2 m^2} \langle x_b | 2H(\boldsymbol{p}) e^{-i\mu^2 \boldsymbol{p}^2} | x_a \rangle.
\end{aligned} \tag{140}$$

The matrix element can be evaluated by introducing a complete basis of momentum eigenkets $|\boldsymbol{p}\rangle$ with integration measure $d\Omega_p = d^n \boldsymbol{p}/(2\pi)^n (1/2\omega_p)$ for the momentum integration. This will give us, in three dimensions with $\boldsymbol{\ell} \equiv \boldsymbol{x}_b - \boldsymbol{x}_a$:

$$\begin{aligned}
\langle x_b | 2H(\boldsymbol{p}) e^{-i\mu^2 \boldsymbol{p}^2} | x_a \rangle &= \int \frac{d^3 \boldsymbol{p}}{(2\pi)^3} \frac{1}{2\omega_p} e^{i\boldsymbol{p}\cdot\boldsymbol{\ell}} [2\omega_p e^{-i\mu^2 \boldsymbol{p}^2}] \\
&= \left(\frac{\pi}{i\mu^2}\right)^{3/2} \frac{1}{8\pi^3} \exp\left(\frac{i\boldsymbol{\ell}^2}{4\mu^2}\right).
\end{aligned} \tag{141}$$

Note that the $2\omega_p$ arising from $2H$ in the left-hand side of Eq. (139) cancels nicely with the $(1/2\omega_p)$ in the measure of integration in the momentum space, giving a simple result. Substituting Eq. (141) into Eq. (140) we get the final result, with $x^2 = x^a x_a = t^2 - \boldsymbol{\ell}^2$,

$$\begin{aligned}
\langle x_b | e^{-iHt} | x_a \rangle &= \left(\frac{i}{\pi}\right)^{1/2} \left(\frac{\pi}{i}\right)^{3/2} \frac{1}{8\pi^3} \int_0^\infty \frac{ds}{2s^2} \\
&\quad \times \exp\left(-\frac{ix^2}{4s} - im^2 s\right), \tag{142} \\
&= \frac{1}{i} \frac{1}{16\pi^2} \int_0^\infty \frac{ds}{s^2} \exp{-i\left(\frac{x^2}{4s} + m^2 s\right)}. \tag{143}
\end{aligned}$$

This is, of course, the standard expression for the Feynman propagator and we have obtained it earlier in Eq. (116); it was

equal to the matrix element in the left-hand side. So where did we cheat?

The identity in Eq. (139) is actually valid when the right-hand side has $\exp(-iH|t|)$. Note that, in our final expression given by Eq. (143), the right-hand side is an even function of t. So the left-hand side should also be an even function of t. This is ensured only because the result we have proved continues to be valid for $t < 0$ as well, if we replace $\exp(-iHt)$ by $\exp(-iH|t|)$. In other words, the evolution operator we have sandwiched between the eigenkets is *not* $\exp(-iHt)$ but

$$U(t) = e^{-iH|t|} = \theta(t)e^{-iHt} + \theta(-t)e^{iHt}. \tag{144}$$

So we are not computing the matrix element of the evolution operator e^{-iHt} as per the standard rule but evaluating a matrix element of the operator $U(t) = e^{-iH|t|}$. This modification of the evolution operator, in which propagation forward in time is dictated by H and the propagation backward in time is dictated by $-H$, makes all the difference in the world.

But the real surprise is the following: We have now shown that the propagator G_R can be expressed as $\langle y | \exp(-iH|t|) | x \rangle$ where $|x\rangle$ and $|y\rangle$ are *non-localized states*! It is not obvious that, merely by using $U(t) = e^{-iH|t|}$ rather than e^{-iHt}, we can still express the correct propagator without solving the problem of localized particle state. This is the real surprise I want to highlight against the background of the discussion in this section.

5.5 Aside: composition law for propagators

In NRQM, the propagator $G_{NR}(x_b, x_a)$ actually propagates the wave function from the event $\mathcal{A}$ to the event $\mathcal{B}$. Such an interpretation relies crucially on the propagator satisfying the composition law in Eq. (33). This composition law, in turn, is a *trivial* consequence of two facts: (i) G_{NR} can be expressed as the matrix element $\langle x_b | \exp[-iH(t_b - t_a)] | x_a \rangle$ and (ii) the set $|x\rangle$ forms a complete set of an orthonormal basis. So multiplying two propagators $G_{NR}(x_b, x_c)$ and $G_{NR}(x_c, x_a)$ and integrating over the variable occurring in $|x_c\rangle$ reduces the composition law to an identity:

$$\begin{aligned}
\int dx_c \, &\langle x_b | e^{-iH(t_b - t_c)} | x_c \rangle \langle x_c | e^{-iH(t_c - t_a)} | x_a \rangle \\
&= \langle x_b | e^{-iH(t_b - t_a)} | x_a \rangle. \tag{145}
\end{aligned}$$

Obviously, this will not hold for the relativistic propagator because the condition (ii) is violated.

It is, however, straightforward to derive the corresponding composition law *with integration over spacetime* rather than just space, for the relativistic propagator. From the integral representation of the propagator in Eq. (129), we immediately

see that[31] with $\mu = m^2$:

$$\int d^D x\, G(\mu; x_2, x) G(\mu; x, x_1) = -\langle x_2 | (\mu + H)^{-2} | x_1 \rangle$$

$$= i \frac{\partial}{\partial \mu} G(\mu; x_2, x_1).$$

(146)

But the integration now is over, say, $d^D x$ at the intermediate event, rather than over $d^n \boldsymbol{x}$; so the physical meaning of this composition law is unclear; you certainly cannot use it to propagate a wave function. (It does not help to restrict the integration over spatial coordinates in Eq. (146); see Appendix A.) It is also obvious from the derivation of Eq. (146) that it is the integration over ds in Eq. (129) which makes the relativistic case very different from the non-relativistic one.

Incidentally, this composition law can be iterated N times to give the result

$$\int d^D x_1 \cdots d^D x_N\, G(\mu; x_b, x_N) \cdots G(\mu; x_1, x_a)$$

$$= (i)^N \frac{\partial}{\partial \mu^N} G(\mu; x_b, x_a).$$

(147)

This result suggests a curious way of reconstructing $G(\mu; x_b, x_a)$. We first note that the Euclidean version of Eq. (146) (in which the i factor on the right-hand side is replaced by -1) can be rewritten, after integrating over μ in the range $m^2 < \mu < \infty$, in the form

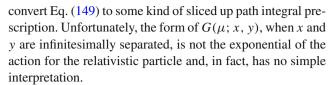$$\int_{m^2}^{\infty} d\mu\, d^D x_1\, G(\mu; x_b, x_1) G(\mu; x_1, x_a)$$

$$\equiv \int d\mathcal{M}_1\, G(\mu; x_b, x_1)$$

$$\times\, G(\mu; x_1, x_a) = G(m^2; x_b, x_a)$$

(148)

where we have treated the propagator as a function of the variable μ and defined the measure of integration as $d\mathcal{M} \equiv d\mu\, d^D x$ and used the fact that the Euclidean propagator vanishes when $m^2 \to \infty$. This equation can be iterated an infinite number of times by keeping two events in $G(\mu; x_j, x_{j-1})$ infinitesimally close to each other. Iterating N times will give the result

$$\int d\mathcal{M}_1 \cdots d\mathcal{M}_N\, G(\mu_N; x_b, x_N) \cdots$$

$$G(\mu_1; x_1, x_a) = G(m^2; x_b, x_a).$$

(149)

This is very similar in structure to the non-relativistic composition law in Eq. (33). Therefore, one can, in principle,

convert Eq. (149) to some kind of sliced up path integral prescription. Unfortunately, the form of $G(\mu; x, y)$, when x and y are infinitesimally separated, is not the exponential of the action for the relativistic particle and, in fact, has no simple interpretation.

The expression for the relativistic propagator in terms of the Jacobi action offers some further insight into the composition law and demystifies it. Even in NRQM, the *energy propagator* $\mathcal{G}(\boldsymbol{x}_2, \boldsymbol{x}_1; E)$, obtained from the path integral sum over the Jacobi action, does not obey the composition law in Eq. (33); instead it satisfies an analog of the composition law in Eq. (146). This is, again, obvious from the structure of $\mathcal{G}(\boldsymbol{x}_2, \boldsymbol{x}_1; E)$, defined in Eq. (124). Expressing the propagator $G(x_2, x_1)$ in Eq. (124) as the matrix element of the time evolution operator of NRQM, we get

$$\mathcal{G}(\boldsymbol{x}_2, \boldsymbol{x}_1; E) = \int_0^{\infty} dt\, e^{it(E + i\epsilon)} \langle \boldsymbol{x}_2 | e^{-it\hat{H}} | \boldsymbol{x}_1 \rangle$$

$$= i \langle \boldsymbol{x}_2 | (E - \hat{H} + i\epsilon)^{-1} | \boldsymbol{x}_1 \rangle$$

(150)

where we have introduced an $i\epsilon$ factor, with an infinitesimal ϵ, to ensure convergence. From Eq. (150), it immediately follows that

$$\int d^D \boldsymbol{y}\, \mathcal{G}(\boldsymbol{x}_2, \boldsymbol{y}; E) \mathcal{G}(\boldsymbol{y}, \boldsymbol{x}_1; E) = -i \left[\frac{\partial \mathcal{G}(\boldsymbol{x}_2, \boldsymbol{x}_1; E)}{\partial E} \right],$$

(151)

which has the same form as the result in Eq. (146), for pretty much the same algebraic reasons. So this composition law in Eq. (146) has *nothing to do* with relativity; it arises because the G_R can be interpreted as arising from a Jacobi action. (One can also write down an iterated relation, identical in form to Eq. (149) in this case as well; unfortunately its physical meaning is not clear.)

The composition law in Eq. (146) induces corresponding composition laws in the Fourier transform of the propagators. Consider first $G_{\boldsymbol{p}}(t_b - t_a)$, which is the spatial Fourier transform of the propagator in Eq. (113). This function satisfies, in the Euclidean sector, the composition law

$$\int_{-\infty}^{\infty} dt\, G_p(t_2 - t) G_p(t - t_1) = -\frac{\partial}{\partial \mu} G_p(t_2 - t_1).$$

(152)

It is straightforward to verify that the integrals on both the left-hand side and right-hand side can be expressed in the form

$$I_{\text{RHS}} = -\frac{G}{2\omega_p} \frac{\partial \ln G}{\partial \omega_p} = \frac{G}{2\omega_p} \left\{ (t_2 - t_1) + \frac{1}{\omega_p} \right\} = I_{\text{LHS}}.$$

(153)

It would be interesting to ask whether one can recover the non-relativistic composition law in Eq. (33) from this result – which looks quite different – in the appropriate limit. This *cannot* be done with the expressions in Eq. (113) but if we change the propagator for the non-relativistic

---

[31] You can obtain the same result by multiplying the two propagators and using their explicit form in Fourier space. For a discussion of the composition laws for the propagators, see Refs. [15,28] and the references therein.

case by multiplying it by a $\theta(t)$ [that is, we take the non-relativistic propagator in the Fourier space to be $G_{\text{NR}}(t, \boldsymbol{p}) = \theta(t) \exp(-i\omega_p t)$] then one can obtain the non-relativistic limit correctly. This is based on the fact that in the non-relativistic limit we have the approximate form

$$-\left.\frac{\partial \ln G}{\partial \mu}\right|_{\text{NR}} = +\frac{1}{2\omega}\left\{(t_2 - t_1) + \frac{1}{\omega}\right\} \approx \frac{t_2 - t_1}{(2m)}. \quad (154)$$

Then, as long as $t_1 < t < t_2$, the composition law in Eq. (152) reduces to the composition law of NRQM in Eq. (113). (Some of the details of these computations are given in Appendix A.)

One can also consider the Fourier transform of the Euclidean propagator with respect to time obtaining $G_E(\boldsymbol{x})$. A simple calculation shows that

$$G_E(\boldsymbol{x}) \equiv \int_{-\infty}^{\infty} G \, e^{iEt} \mathrm{d}t = \int \frac{d^n \boldsymbol{p}}{(2\pi)^3} \frac{e^{i \boldsymbol{p} \cdot \boldsymbol{x}}}{E^2 + \boldsymbol{p}^2 + m^2}. \quad (155)$$

This function satisfies the composition law

$$\int d^n \boldsymbol{x} \, G_E(\boldsymbol{x}_2, \boldsymbol{x}) \, G_E(\boldsymbol{x}, \boldsymbol{x}_1) = -\frac{\partial}{\partial \mu} G_E(x_2, x_1), \quad (156)$$

which is easy to verify.

Finally, let us consider the composition law which *does* lead to the propagator in terms of two other propagators in the relativistic case. Since the scalar product for the relativistic Klein–Gordon equation is defined as

$$(\phi_1, \phi_2) \equiv i \int d\sigma^a [\phi_1^* \partial_a \phi_2 - \phi_2 \partial_a \phi_1^*]$$
$$= i \int d\sigma^a \, \phi_1^* \overleftrightarrow{\partial_a} \phi_2, \quad (157)$$

it is straightforward to show that the propagator, treated as a function of x, satisfies the composition law:

$$(G^*(x_2, x), G(x, x_1)) = G(x_2, x_1). \quad (158)$$

This result holds *only* as long as $x_2^0 > x^0 > x_1^0$. On the other hand, if $x^0 > x_2^0 > x_1^0$, say, the integral on the left-hand side vanishes. One simple way to prove this result is to Fourier transform Eq. (158) with respect to spatial coordinates and write the corresponding condition involving $G_R(t, \boldsymbol{p})$ and $\partial_t G_R(t, \boldsymbol{p})$. We next note from Eq. (113) that $G_R(t, \boldsymbol{p})$ and its time derivative can be expressed as

$$G = \frac{1}{2\omega}\left[\theta(t)e^{-i\omega t} + \theta(-t)e^{+i\omega t}\right];$$
$$\partial_t G = +\frac{i}{2}\left[-\theta(t)e^{-i\omega t} + \theta(-t)e^{i\omega t}\right], \quad (159)$$

leading to $\partial_t G = -i\omega G[\text{Sg}(t)]$ where $\text{Sg}(t) = t/|t|$ is the sign function. With this result, it is easy to show that the combination occurring on the left-hand side of Eq. (158) in Fourier space is proportional to $G_R(t_2 - t, \boldsymbol{p})G_R(t - t_1, \boldsymbol{p})[\text{Sg}(t_2 - t) + \text{Sg}(t - t_1)]$, which is non-zero only in

the interval $t_1 < t < t_2$. In this interval the relation Eq. (158) is identically satisfied.[32] So, even though the Feynman propagator can propagate backwards in time, it does not work in the composition laws.

6 A seamless route from QFT to NRQM

In Sect. 3 we found that one is led to a notion of a field operator $A(x)$ fairly naturally from the propagator both in RQM and in NRQM. This was done by introducing the "creation" and "annihilation" operators in the Fourier space, $A_{\boldsymbol{p}}$ and $A_{\boldsymbol{p}}^\dagger$, and defining $A(x)$ by Eq. (47). This approach, therefore, holds promise for a seamless transition from QFT to NRQM.

There was, however, one serious difficulty. We found that, in QFT, the commutator $[A(x_2), A^\dagger(x_1)] = \langle x_2|x_1\rangle$ (where the state $|x\rangle$ is defined by Eq. (27)) does not reduce to a Dirac delta function on a space-like hypersurface. This is a reflection of the non-localizability of the particle position. So if you build observables from A and $A^\dagger$, they will not commute for events separated by a space-like interval. A sensible way of incorporating causality into quantum theory will be to arrange matters such that commutator between observables vanish for space-like separated events. So we cannot treat $A(x)$ as the basic building block in the theory and need to do a little bit more work.

To tackle this issue, we will introduce another field $B(x)$ whose commutator will lead to $G_-(x_2, x_1) = G_+^*(x_2, x_1)$ just as the commutator in Eq. (51) lead to $G_+(x_2, x_1)$. This is achieved through the definition

$$B(x) \equiv \int d\,\Omega_{\boldsymbol{p}} B_{\boldsymbol{p}} e^{-ipx} \quad (160)$$

with the assumption that $B(x)$ commutes with $A(x)$. It is straightforward to verify that $[B(x_2), B^\dagger(x_1)] \equiv G_-(x_2; x_1)$. Let us now define the combination $\phi(x) = A(x) + B^\dagger(x)$. This field ϕ will also satisfy the Klein–Gordon equation since A and B do. But ϕ has better behavior as regards causality. It is straightforward to show that

$$[\phi(x_2), \phi^\dagger(x_1)] = [A(x_2) + B^\dagger(x_2), A^\dagger(x_1) + B(x_1)]$$
$$= [A(x_2), A^\dagger(x_1)] - [B(x_1), B^\dagger(x_2)]$$
$$= G_+(x_2; x_1) - G_+(x_1; x_2). \quad (161)$$

This commutator vanishes at space-like separation because $G_+(x_2; x_1) = G_+(x_1; x_2)$ in that case. (See Appendix B; for a nice discussion of the role of causality in QFT, see [41].)

---

[32] If you do the integrals in Eq. (158) in the real space, this result arises, after a bit of tedious algebra, because a factor in the numerator cancels a pole in the denominator in a rather subtle manner. It is also related to the orthogonality of $G_+(x, y)$ and $G_-(x, y) \equiv G_+^*(x, y)$. I do not know of any simple way to "guess" this result.

So we find that to maintain relativistic causality we need to work with two fields A and B and define the physical field as $\phi(x) = A(x) + B^\dagger(x)$. We also have the relations giving the propagator directly in terms of ϕ and $\phi^\dagger$:

$$\langle 0|\phi(x_2)\phi^\dagger(x_1)|0\rangle = \langle 0|A(x_2)A^\dagger(x_1)|0\rangle$$
$$= \int d\,\Omega_{\boldsymbol{p}}\, e^{-ipx} = G_+(x_2; x_1),$$
$$\langle 0|\phi^\dagger(x_1)\phi(x_2)|0\rangle = \langle 0|B(x_1)B^\dagger(x_2)|0\rangle = \int d\,\Omega_{\boldsymbol{p}}\, e^{+ipx}$$
$$= G_-(x_2; x_1). \tag{162}$$

These relations, in turn, allow us to express our relativistic propagator entirely in terms of ϕ through the relation

$$G(x_2, x_1) = \theta(t_2 - t_1)\langle 0|A(x_2)A^\dagger(x_1)|0\rangle$$
$$+ \theta(t_1 - t_2)\langle 0|B(x_1)B^\dagger(x_2)|0\rangle$$
$$= \langle 0|T(\phi(x_2)\phi^\dagger(x_1))|0\rangle. \tag{163}$$

To summarize, we first introduced a primitive field $A(x)$ based on the relationship between $|x\rangle$ and $|\boldsymbol{p}\rangle$. This field satisfies the Klein–Gordon equation but not our notion of causality. Looking at the structure of the commutator of A field, we introduced another field $B(x)$, which also satisfies the Klein–Gordon equation and, finally, a physical field $\phi(x)$ which obeyed the Klein–Gordon equation and the notion of causality. Obviously, the notion of causality introduced here will disappear in the non-relativistic limit and the two primitive fields A and B will – so to speak – be liberated. They will have appropriate non-relativistic limits which will allow us to construct NRQM in a proper manner.

To see how this comes about, we first introduce two "non-relativistic" fields $a(x)$ and $b(x)$ in place of $A(x)$ and $B(x)$ by

$$A(x) \equiv \frac{e^{-imt}}{\sqrt{2m}}\, a(x), \qquad B(x) \equiv \frac{e^{-imt}}{\sqrt{2m}}\, b(x). \tag{164}$$
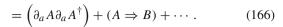
This rescaling does two things: (i) It separates out a rapidly oscillating phase $\exp(-imc^2t)$ from the fields; this phase arises from the relativistic rest energy of the particle. (ii) It factors out $(1/\sqrt{2m})$, which is a vestige of the relativistic momentum measure $(1/2\omega_p)$, which goes over to $(1/2m)$ in the non-relativistic limit. Thus we have eliminated two key relativistic factors (one due to rest energy, mc^2, and the other due to the change of measure in momentum integration) from the fields A and B to define a and b.

We next express the Lagrangian $L = \partial_a\phi\,\partial^a\phi^\dagger - m^2\phi\phi^\dagger$ for the physical field ϕ in terms of a and b fields. The kinetic energy part is

$$\partial_a\phi\,\partial^a\phi^\dagger = \left(\partial_a A + \partial_a B^\dagger\right)\left(\partial^a A^\dagger + \partial^a B\right), \tag{165}$$
$$= \left(\partial_a A\partial^a A^\dagger\right)$$
$$+ \left(\partial_a B^\dagger\partial^a B\right) + \partial_a A\partial^a B + \partial_a A^\dagger\partial^a B^\dagger$$

$$= \left(\partial_a A\partial_a A^\dagger\right) + (A \Rightarrow B) + \cdots. \tag{166}$$

Here and in what follows, the $\cdots$ represent terms with factors $\exp(\pm 2imt)$, which can be ignored since they rapidly oscillate and average out to zero in the non-relativistic limit.[33] In terms of the "non-relativistic" fields, the first term is given by

$$2m\partial_a A\,\partial^a A^\dagger = \left[(-ima + \dot{a})\left(ima^\dagger + \dot{a}^\dagger\right) - \partial_\mu a\,\partial^\mu a^\dagger\right] \tag{167}$$

and corresponding terms for b. Similarly,

$$m^2\phi^\dagger\phi = m^2 AA^\dagger + (A \Rightarrow B) + \cdots$$
$$= m^2 aa^\dagger + (a \Rightarrow b) + \cdots. \tag{168}$$

Using these results we can express the Lagrangian in terms of $a(x)$ and $b(x)$ as

$$2L = a^\dagger\left(i\partial_t - H\right)a + b^\dagger\left(i\partial_t - H\right)b + \text{h.c} + \cdots \tag{169}$$

where $H = -(1/2m)\nabla^2$ is the non-relativistic Hamiltonian for free particle – obtained by writing $\partial_\mu a\,\partial^\mu a^\dagger = \partial_\mu(a\,\partial^\mu a^\dagger) - a\,\partial_\mu\partial^\mu a^\dagger$ in Eq. (167) and ignoring the total divergence – and the dots indicate terms which can be ignored in the non-relativistic limit. These are terms of the kind

$$Q = |\dot{a}|^2 + |\dot{b}|^2 + e^{-2imt}(\ ) + e^{2imt}(\ ). \tag{170}$$

The terms $|\dot{a}|^2$ and $|\dot{b}|^2$ are ignorable because the leading time variation, viz. the e^{-imt} factor, has been pulled out on defining the non-relativistic fields a and b; therefore, it is justifiable to retain only up to first time derivative while working with a and b. We can also ignore terms multiplied by factors $\exp(\pm 2imt)$ since they rapidly average out to zero in the non-relativistic limit.

From the structure of Eq. (169) we see that, in the non-relativistic limit, our system is described by *two* fields a and b, which actually represent the particle and antiparticle of the original system. Both of them satisfy the non-relativistic Schroedinger equation in operator form. *So antiparticles do not go away when you take the non-relativistic limit if you do it correctly.*

We worked with the primitive fields A and B (which actually correspond to the particle and antiparticle, respectively) in order to show that the non-relativistic limit leads to a pair of fields a and b both obeying the Schroedinger equation. It is, however, possible to work entirely with the physical field ϕ and obtain the appropriate limit. To do this, we start with the definition of ϕ, viz.,

$$\phi = A + B^\dagger = \left(ae^{-imt} + b^\dagger e^{imt}\right)\frac{1}{\sqrt{2m}}. \tag{171}$$

---

[33] One can do this more formally in an RG-type analysis by integrating out the high frequency modes and defining a low-energy effective Lagrangian. But since the modes are decoupled in the free field theory, this is equivalent to just dropping the rapidly oscillating terms.

The canonical momentum associated with ϕ is

$$\Pi = \dot\phi \approx i\sqrt{\frac{m}{2}}\left(-ae^{-imt} + b^\dagger e^{imt}\right) \tag{172}$$

where we have ignored the time derivatives of a and b in comparison with the time derivatives coming from $\exp(\pm imt)$ factor. This allows us to write

$$a = e^{imt}\left(\sqrt{\frac{m}{2}}\,\phi + \frac{i}{\sqrt{2m}}\,\Pi\right);$$
$$b^\dagger = e^{-imt}\left(\sqrt{\frac{m}{2}}\,\phi - \frac{i}{\sqrt{2m}}\,\Pi\right). \tag{173}$$

This procedure works even for a real scalar field for which the antiparticle is identical to the particle. So, even real scalar fields have a natural non-relativistic limit, contrary to what is sometimes claimed in the literature.

The most important feature which has come about in the non-relativistic limit is the transition from second time derivatives to first time derivatives in the equation obeyed by the operators. That is, the relevant operator changes from $(\partial_t^2 - \nabla^2 + m^2)$ to $(i\partial_t + (1/2m)\nabla^2)$ or equivalently $(\nabla^2 - m^2 - \partial_t^2)$ goes over to $\nabla^2 + (2mi\,\partial_t)$. So the net effect is the replacement

$$(\partial_t^2 + m^2) \Longrightarrow (-2mi\,\partial_t). \tag{174}$$

Almost all the key differences between QFT and NRQM are directly or indirectly connected with this change. In view of its importance, it is worth going over the algebraic features which led to this reduction.

Since the spatial dependence is governed by the *same* operator ∇^2 *both* in the relativistic and non-relativistic field equations, we can work in the Fourier space – with modes labeled by the magnitude of a wave vector k – in both cases. In the relativistic case the Fourier mode will satisfy a harmonic oscillator equation with frequency $\Omega_k^2 \equiv k^2 + m^2$. All we need to do is to look at appropriate features of harmonic oscillators to understand what is going on. So consider a dynamical degree of freedom $f(t)$, which satisfies the harmonic oscillator equation of the form

$$\left(\frac{d^2}{dt^2} + \Omega_k^2\right)f = \left(\frac{d^2}{dt^2} + k^2 + m^2\right)f. \tag{175}$$

To study NRQM, we want to look at the limit $k^2 \ll m^2$ when the frequency of oscillation of f will be dominated by a factor like $\exp(\pm imt)$. It makes sense to pull this factor out of f and redefine another dynamical variable F by the relation $f = e^{-imt}F$. It is now straightforward to show that the Lagrangian that leads to Eq. (175) can be re-expressed in terms of F as

$$L = f^\dagger\left(\frac{d^2}{dt^2} + \Omega_k^2\right)f = F^\dagger\left(\frac{k^2}{2m} - i\partial_t\right)F + \frac{F^\dagger \ddot F}{2m}. \tag{176}$$

The first term on the right-hand side involves only the first time derivative. The second term contains $\ddot F$, which can be ignored compared to $\dot F$ in the limit we are interested in. This is how the reduction of time derivatives occurs when we proceed from QFT to NRQM. The culprit is the rest energy, which introduces rapid time oscillations through the factor $\exp(-imt)$.

The idea of the primitive fields a and b and the physical field ϕ can also be understood without worrying about the spatial dependence, and working with Fourier modes which behave like oscillators. To do this, let us consider a dynamical variable $q(t)$ described by a Lagrangian

$$L = \dot q^\dagger \dot q - \Omega^2 q^\dagger q + \text{h.c.} \tag{177}$$

We now introduce two primitive fields $a(t)$ and $b(t)$, such that $q = a + b^\dagger$, and re-express L in terms of a and b. You will find that the Lagrangian separates into two parts as $L = L_1 + L_2$ where

$$L_1 = |\dot a|^2 - \Omega^2|a|^2 + |\dot b|^2 - \Omega^2|b|^2 \tag{178}$$

and

$$L_2 = -a(\ddot b - \Omega^2 b) - a^\dagger(\ddot b - \Omega^2 b)^\dagger. \tag{179}$$

The second part of the Lagrangian L_2 actually leads to the identical field equations as L_1. For example, if you vary a in L_2 you get $\ddot b = -\Omega^2 b$, which is identical to the field equation you get from the second pair of terms in L_1. Therefore we can ignore L_2 and think of the dynamics as being dictated by L_1 itself. The L_1 describes two independent oscillators a, b with frequency Ω. By an analysis similar to the one done before, we can reduce this system to one which involves only the first time derivative. This is exactly analogous to what we have done earlier in the case of the field.

One feature which emerges out of this analysis is the sharp distinction between (i) any direct approach to quantum theory of relativistic particle and (ii) relativistic particles emerging as excitations of a quantized field. Conceptually these constructions are completely different. To describe a relativistic particle, we can start with an eigenstate $|p\rangle$ of its three-momentum (with its energy ω_p determined by $\omega_p = +(p^2 + m^2)^{1/2}$. One can build further states like, for example, $|x\rangle$ and other useful operators like, for example, $A(x)$ etc. and build a theory in a suitable Hilbert space. But such a field $A(x)$ will not obey a sensible notion of causality. To remedy this situation *we have to double up the number of particles by associating with each particle another particle* with (an unfortunate) nomenclature antiparticle. This is roughly what the introduction of the field $B(x)$ does. Then the combination $\phi(x) = A(x) + B^\dagger(x)$ obeys a natural notion

of micro-causality. So the answer to the question "why do antiparticles exist" is simply "to ensure causality in a Lorentz invariant theory". It has nothing to do with square roots in Hamiltonians or some funny notion of negative energy states; there are no negative energy states in the one-particle sector of the Fock space. The two fields $A(x)$ and $B(x)$ have to be treated on an equal footing and both have a right to exist in NRQM. In short, a pair of fields in NRQM gets mapped to a single field in QFT.

7 Propagators as correlators

We have seen that when you take the non-relativistic limit properly, an operator remains an operator. All that happens is that the Lagrangian and the field equation describing the field operator $\hat{\phi}(x)$ are different from the ones describing the field operator $\hat{a}(x)$ and $\hat{b}(x)$. $\phi(x)$ satisfies a field equation which is of second order in time, while $a(x)$ and $b(x)$ satisfy field equations which are of first order in time. These non-relativistic fields are what are usually called – in a confusing and incorrect nomenclature – the "second quantized" version of Schroedinger wave functions. By using the language of field operators both in QFT and NRQM we can make a seamless transition from QFT to NRQM. (This is a familiar aspect of condensed matter physics but is not usually explored in detail in the context of non-relativistic limit of QFT; some earlier work is cited in [40].) The last issue which remains to be answered is the role of the propagators: How do we obtain the non-relativistic propagator in the appropriate limit since we are no longer talking about particle positions and trajectories *even* in the NRQM limit?

To answer this question, let us start by examining an action which is quadratic in the fields and can be expressed in the form
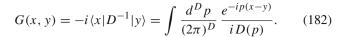
$$A = \int d^D x \ \Phi^* \hat{D}(i\partial_a)\Phi = \int d^D x \ \Phi^*(\hat{Q} + \mu)\Phi. \quad (180)$$

The first equation defines an operator $\hat{D}$ which is built from the time and space derivatives ∂_a; for convenience we have introduced a parameter μ and written this operator as $\hat{D} \equiv \hat{Q} + \mu$. (In the case of Klein–Gordon field, for example, μ could be identified with m^2.) We will now define the propagator for the field as the correlator averaged using e^{iA} through

$$G(x, y) \equiv \langle \Phi(x)\Phi^*(y) \rangle \equiv \frac{1}{Z} \int \mathcal{D}\Phi \ \mathcal{D}\Phi^* \ \Phi(x)\Phi^*(y) \ e^{iA};$$
$$Z \equiv \int \mathcal{D}\Phi \ \mathcal{D}\Phi^* e^{iA}. \quad (181)$$

Since the action is quadratic, it is straightforward to evaluate this correlator, which is the matrix element of D^{-1} in Fourier space. We get

$$G(x, y) = -i \langle x|D^{-1}|y \rangle = \int \frac{d^D p}{(2\pi)^D} \frac{e^{-ip(x-y)}}{i D(p)}. \quad (182)$$

As an example, consider the standard Klein–Gordon field. In this case, we have $\hat{D} = (\Box + \mu - i\epsilon) = (-p^2 + \mu - i\epsilon)$, so that $-iD^{-1} = +i(p^2 - \mu + i\epsilon)^{-1}$. This will lead to the standard Feynman propagator G_R.

One can give a nicer interpretation to *any* such propagator by using the integral representation for D^{-1} and writing

$$G(x, y) = \int_0^\infty ds \ \langle x|e^{-is\hat{D}}|y \rangle$$
$$= \int_0^\infty ds \int e^{-ip \cdot (x-y)} \ e^{-isD(p)} \frac{d^D p}{(2\pi)^D}. \quad (183)$$

The second expression is obtained by introducing a complete set of momentum eigenstates and using $\langle x|p \rangle = e^{-ipx}$ etc. The matrix element $\langle x|e^{-isD}|y \rangle$ can be thought of as a quantum mechanical propagator for a particle to go from y to x under the action of a Hamiltonian D in "time" interval s. The structure of this expression immediately leads to the composition law for the propagator. Since

$$\int dx \ \langle x_2|(iD)^{-1}|x \rangle \langle x|(iD)^{-1}|x_1 \rangle = \langle x_2|(iD)^{-2}|x_1 \rangle$$
$$= i\frac{\partial}{\partial \mu} \langle x_2|(iD)^{-1}|x_1 \rangle, \quad (184)$$

we obtain the result

$$\int d^D x \ G(x_2, x) \ G(x, x_1) = i\frac{\partial}{\partial \mu} G(x_2, x_1). \quad (185)$$

The discussion so far has been completely general. Let us now consider the question recovering NRQM from this approach. We start with the Fourier transform of the propagator with respect to spatial coordinates which can be expressed as

$$G_{\boldsymbol{k}}(t) \equiv \int d^D \boldsymbol{x} \ G(t, \boldsymbol{x}) \ e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} = \int_{-\infty}^\infty \frac{d\omega}{(2\pi)} \frac{e^{-i\omega t}}{i D(\omega, \boldsymbol{k})}. \quad (186)$$

Obviously, the form of the propagator depends on the pole structure of $D(\omega, \boldsymbol{k})$ in the complex plane. We saw in the last section that the essential difference between QFT and NRQM is in the reduction of second time derivatives to first time derivatives, indicated by Eq. (174). This, in turn, suggests that in the Fourier domain, a second order pole is replaced by a first order pole in ω. In fact, this is indeed the case. You will get the standard form of NRQM if the pole structure of $D(\omega, \boldsymbol{k})$ has the form

$$D(\omega, \boldsymbol{k}) - i\epsilon = [-\omega + F(\boldsymbol{k}) - i\epsilon] (2\Omega_{\boldsymbol{k}}). \quad (187)$$

Then a simple contour integration of the integral in Eq. (186) will give the momentum space propagator:

$$G_{\boldsymbol{k}}(t) = \frac{\theta(t)}{2\Omega(\boldsymbol{k})} \exp\left[-it\,F(\boldsymbol{k})\right], \qquad F \equiv H + \mu. \quad (188)$$

This will lead to standard NRQM if $2\Omega(\boldsymbol{k}) = $ constant. This propagator obeys the composition law in Fourier space given by

$$\int_{-\infty}^{\infty} dt\,(2\Omega_k)G_k(t_2,t)\,(2\Omega_k)G_k(t,t_1)$$
$$= i\frac{\partial}{\partial\mu}(2\Omega_k)G_k(t_2,t_1). \quad (189)$$

From the explicit form of the propagator in Eq. (188), we see that the right-hand side of Eq. (189) is given by

$$i\frac{\partial}{\partial\mu}(2\Omega_k)G_k(t_2,t_1) = (t_2 - t_1)\,G_k(t_2,t_1). \quad (190)$$

The left-hand side of Eq. (189) will also reduce to this expression because of the theta functions in time and we will recover the standard result in NRQM. Thus it is clear that NRQM is recovered when $D(\omega,\boldsymbol{k})$ has a single pole in the lower half plane.

We can also construct the propagator directly from Eq. (183) along the following lines. Introducing a complete set of momentum eigenkets $|p\rangle$ in the matrix element, this expression can be reduced to

$$G(x) = \int_0^\infty ds \int \frac{d^n\boldsymbol{p}}{(2\pi)^n}\,e^{i\,\boldsymbol{p}\cdot\boldsymbol{x}}$$
$$\times \int_{-\infty}^\infty \frac{d\omega}{2\pi}\,e^{-i\omega t}\,e^{-is(2\Omega)(-\omega+F-i\epsilon)}, \quad (191)$$
$$= \int \frac{d^n\boldsymbol{p}}{(2\pi)^n}\int_0^\infty ds\,\delta\left((2\Omega_p)\,s - t\right)\,e^{-isF(\boldsymbol{p})+i\,\boldsymbol{p}\cdot\boldsymbol{x}}, \quad (192)$$
$$= \theta(t)\int \frac{d^n\boldsymbol{p}}{(2\pi)^n}\frac{1}{2\Omega_{\boldsymbol{p}}}\,e^{-itF(\boldsymbol{p})+i\,\boldsymbol{p}\cdot\boldsymbol{x}}. \quad (193)$$

Notice that, when there is only one pole for $D(\omega,\boldsymbol{k})$, making it a linear function of ω, the ω integration in the first line leads to a Dirac delta function in time. This allows us to identify the "internal time" s with the physical time t, leading to the final result. The final expression also has a direct interpretation in terms of the Hamiltonian form of the action principle. Thus the definition of propagators as correlators work consistently both in QFT and in NRQM. The key difference between the two is in the pole structure of the operator $\hat{D}$, which, in turn, is related to the conversion of second time derivatives to first time derivatives as explained in the previous section.

8 Discussion

This has been a rather long journey and – for the sake of clarity – let me briefly describe the path we have followed and the landmarks on the way. (The reader is invited to revisit the summary of the results given in Sect. 1.2 at this stage, for more details.) I will then conclude by highlighting two important results we have obtained.

8.1 Brief overview

One main conclusion – which we have reached from several different perspectives – is that, to make a seamless transition from QFT to NRQM, you need to describe NRQM in a language which is closer to that of QFT and not the other way around. This conclusion by itself may not be surprising but it was necessary to demonstrate it from different perspectives, which was one of the main objectives achieved in the paper.

The NRQM limit can be obtained for a free particle by working with relativistic particle *and* antiparticle field operators $A^\dagger(x)$ and $B^\dagger(x)$. (The antiparticles do *not* "go away" in the NRQM limit.) These operators are, in turn, defined in terms of operators which create fixed three-momentum states from the no-particle state. The three-momentum continues to be a "good" operator in QFT while the three-position is not. I have commented on this aspect extensively, contrasting the non-relativistic and relativistic cases, where the Hamiltonian takes the forms $H(\boldsymbol{p}) = \boldsymbol{p}^2/2m$ or $H(\boldsymbol{p}) = \boldsymbol{p}^2 + m^2$, respectively.

A closely related question is whether the non-relativistic wave function can be recovered through some limiting procedure from a relativistic field operator. I addressed this by focusing on the propagator, an object that is well-defined in both NRQM and QFT. The technical issue, which makes all the difference between the two cases, is the fact that the measure of integration in momentum space has to be different in the two cases, which – in turn – arises from the requirement of Lorentz invariance. This difference features throughout the discussion, and it makes it impossible to perform a Fourier transform in the relativistic case that will yield Lorentz covariant coordinate wave functions representing spatially localized particles.

After discussing these aspects, I turned to the issue of obtaining NRQM from QFT using the path integral formalism. Once again, the simplest route is to try and define the respective propagators from the path integrals. You then find that the Lagrangian path integral cannot be defined through time slicing in the relativistic case for any sensible choice of measure. The Hamiltonian approach does work in both cases but does *not* lead to the correct Feynman propagator.

The best route seems to be the one based on a Euclidean lattice regularization scheme, which *does* lead to the Feynman propagator. In this approach we sum over the paths,
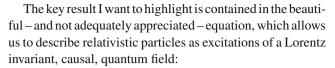
parametrized by proper time, including implicitly those that proceed both forward and backward in *coordinate* time. Exploiting the mathematical similarity of this method to the approach based on the Jacobi action principle, one can again understand the origin of the difficulties in obtaining a single-particle wave function. The Jacobi action approach tells us that, in the non-relativistic case, we need to construct a propagator for fixed energy and then sum over all energies while, in the relativistic case, we need to sum over paths for a fixed proper time followed by an integration over the proper time. It is this last integration (over energy in NRQM and over proper time in RQM) that ruins the composition property of the propagator in either situation.

To conclude this summary, I will comment briefly on two issues which are indirectly related to the discussion in this paper. The first comment has to do with the philosophical interpretation of the wave function in NRQM, which is still strongly debated. But note that: (i) QFT is more fundamental than NRQM, and (ii) we do not have a sensible notion of the single-particle wave function in RQM. Therefore, the debate over the ontological versus epistemological status of the wave function *within the context of NRQM* – in which it is often attempted – seems irrelevant and misplaced. At least one should expand the debate to full QFT (say in the Schroedinger functional formalism) for it to be meaningful; but then we will face several new serious, nontrivial, issues which might take precedence and change the nature of the debate.

The second comment is more technical. We saw that the consistent description of the NRQM limit of QFT requires us to work with a *pair* of fields, corresponding to a particle and its antiparticle. In the case of charged particles, these two will carry equal and opposite charges and hence there is a natural notion of charge conjugation with an associated operator in QFT. From our discussion it is clear that this is a purely relativistic feature and one does not have natural notion of charge conjugation operation in the NRQM, within the single-particle sector. There are attempts in the literature to introduce the notion of charge conjugation in NRQM but these attempts lack the naturalness with which one can introduce this notion in QFT.

8.2 Two intriguing results

The investigations of the path integral leads to some remarkable results, definitely worthy of further study. The first one is the expression for the Feynman propagator, $G_R(x_2, x_1) = \langle x_2 | e^{H|t|} | x_1 \rangle$, with the appearance of the absolute value of the time difference in the evolution operator. The second one is an intriguing relation between the path integral and the existence of antiparticles. I will now discuss these two results, starting from the second one.

The key result I want to highlight is contained in the beautiful – and not adequately appreciated – equation, which allows us to describe relativistic particles as excitations of a Lorentz invariant, causal, quantum field:

$$\sum_{\text{paths}} \exp\left(-\frac{im}{\hbar} \int_1^2 dt \sqrt{1 - \boldsymbol{v}^2}\right)$$
$$= \theta(t_2 - t_1)\langle 0|A(x_2)A^\dagger(x_1)|0\rangle$$
$$+ \theta(t_1 - t_2)\langle 0|B(x_1)B^\dagger(x_2)|0\rangle. \tag{194}$$

The equality of the left-hand side with the relativistic propagator $G_R(x_2, x_1)$ was demonstrated by lattice regularization in the Euclidean sector in Sect. 5; the equality of the right-hand side with the relativistic propagator $G_R(x_2, x_1)$ is provided by Eq. (163).

The remarkable fact about Eq. (194) *is that nobody understands it!*. That is to say, no one has found a simple, physical argument suggesting why the left- and right-hand sides of Eq. (194) should be equal without doing fairly elaborate calculations. This means that we do not quite understand the conceptual basis of QFT – and the structural implications of combining the principles of quantum theory and special relativity – in spite of its remarkable success as a working tool.

To see why 'explaining' Eq. (194) is hard, consider the two sides separately. On the left-hand side we have the action for a *single* relativistic particle summed over all paths in spacetime connecting two events. So the left-hand side combines the principles of quantum theory and special relativity in the most straightforward manner. The right-hand side, on the other hand, describes *two* kinds of particles propagating between the two events in spacetime. If $t_2 > t_1$, then the A-type particle propagates forward in time, while, if $t_2 < t_1$, the B-type particle again propagates forward in time. (There is no propagation of particles backward in time which textbooks are fond of invoking.) It is a mystery how the path integral for a *single* relativistic particle gets an equivalent description in terms of *two* kinds of particles – both propagating forward in time with the choice of particles determined by the time ordering. It would be nice if a prescription for the sum over paths can be devised which nicely separates the contributions from A and B type particles on the right-hand side. (I have some ideas on how to do this, but – as you could have guessed – none of them works properly.)[34]

The second result I want to highlight is the one we found in Sect. 5.4. We found that the relativistic propagator can be expressed in the form

$$G_R(x_2, x_1) = \langle \boldsymbol{x}_2 | e^{-iH|t|} | \boldsymbol{x}_1 \rangle. \tag{195}$$

---

[34] Usually textbooks combine the two terms in the right-hand side of Eq. (194) into a single time ordered product of the field $\phi = A + B^\dagger$, which does not help in resolving the mystery.

We are working throughout with $H = \sqrt{\boldsymbol{p}^2 + m^2}$, which is a positive definite operator. But for $t = -|t| < 0$, we have $\exp(-iH|t|) = \exp[-i(-H)t]$ and thus the minus sign in t can be transferred to H giving the illusion of a negative energy Hamiltonian. Since the operator $U(t) \equiv e^{-iH|t|}$ separates into two distinct evolution operators for $t > 0$ and < 0, it is obvious that two types of propagations are again incorporated in Eq. (195) just as in the case of Eq. (194). This is understandable *but the real surprise has to do with the quantum states between which the time evolution occurs* in Eq. (195). We have repeatedly seen that there are no localized particle states in QFT and hence we necessarily have to interpret $|\boldsymbol{x}_1\rangle$ and $|\boldsymbol{x}_2\rangle$ as some kind of smeared particle position states. Then Eq. (195) tells you that the relativistic propagator is obtained by the *standard* time evolution operator used with either H or $-H$ between such *smeared* states. Clearly, some subtle interplay is going on between the non-localizability of particle states and the existence of two kinds of propagation. As in the case of Eq. (194) I do not know of any simple way of explaining Eq. (195) vis a vis the occurrence of smeared states.

Finally, let me comment on some broader implications of these and other results highlighted in the paper. I believe we can learn lessons regarding combining General Relativity (GR) with QM from carefully exploring the new features which arise when we combine Special Relativity (SR) with QM, which is the motivation for these comments.

It often happens in physics that certain well-defined notions become approximate or, sometimes, even lose their utility when we proceed from an approximate description of Nature to a more exact description. It is possible that the spatial location of an event is one of such concepts. In classical physics, both relativistic and non-relativistic, the notion of a spatial location $\boldsymbol{x}$ is operationally identified either with the position of the particle $\boldsymbol{x}(t)$ at some time t or through the intersection of the world lines of two particles. Both these notions assume the existence of particles with arbitrarily small dimensions.

In the conventional formulation of NRQM this idea is retained except for elevating $\boldsymbol{x}(t)$ to a Heisenberg operator $\hat{\boldsymbol{x}}(t)$ while retaining the purely parametric (non-operator) status for time t. In NRQM you can still work with sharply localized one-particle states $|t, \boldsymbol{x}\rangle$, which are eigenstates of the operator $\hat{\boldsymbol{x}}(t)$, as long as you do not care about the momentum of the particle. But, as we have seen, the introduction of special relativity into QM makes this notion ill-defined. We no longer have localized particle states in RQM, which, of course, is well known in the literature. But if you do not have localized particle states, can you still use the notion of spatial coordinates as though they are well-defined? The usual belief is that one can. For example, combining the uncertainty principle of QM with the mass–energy equivalence of SR, we immediately reach the conclusion that the notion of

a single-particle position becomes ill-defined, for a particle of mass m, at length scales below $\lambda_C \equiv \hbar/mc$. So by considering hypothetical particles of arbitrarily high mass you can define spatial location with arbitrarily high accuracy.

This idea, of course, breaks down when you approach the Planck length, L_P. It is well known that one cannot (see, for e.g., [42,43] and the references therein) operationally define spatial locations with an accuracy better than a few Planck lengths, say. This in turn brings about an *extra* non-localization in the states $|\boldsymbol{x}\rangle$. In the absence of gravity, $\langle \boldsymbol{y}|\boldsymbol{x}\rangle$ differs from a Dirac delta function and has significant support over a region of the size $|\boldsymbol{x} - \boldsymbol{y}|^2 \approx \lambda_C^2$. When we introduce the Planck length into the consideration, we probably need to modify the form of $\langle \boldsymbol{y}|\boldsymbol{x}\rangle$ so that it has support in a region, say, $|\boldsymbol{x} - \boldsymbol{y}|^2 \approx \lambda_C^2 + L_P^2$ or something like that.[35]

To incorporate any such modification at a fundamental scale, we may have to abandon the notion of precise spatial location $\boldsymbol{x}$. Instead, one may want to consider creation and annihilation operators for spatial locations themselves; the action of these operators on a pre-geometric quantum state should produce the standard geometrical notion of a space-like hypersurface as a collection of spatial coordinates along with other geometrical notions. This is a coordinate-based notion of the more abstract idea that a creation operator $A^\dagger(^3\mathcal{G})$ creates a three-geometry $^3\mathcal{G}$ out of a pre-geometric state. Such an approach may be necessary to incorporate the breakdown of operational notion of spatial location at the Planck scale. The de-localization of the position by an amount λ_C, which arises when we combine SR with QM, suggests that some such structure is required to describe the spacetime when we combine GR with QM.

A Appendix A

The left-hand side of Eq. (152) has the form

---

[35] The standard smearing over the Compton wavelength arises with the Klein–Gordan equation of the form $(\Box + \lambda_C^{-2})\phi = 0$, while one can get the above modification if we use the equation $(\Box + \ell^{-2})\phi = 0$ where $\ell^2 = \lambda_C^2 + L_P^2$. One can also introduce such a zero-point-length into spacetime in a Lorentz invariant manner by modifying the path integral; see [44,45].

$$I_{\text{LHS}} = \int_{-\infty}^{\infty} dt \, \frac{1}{(2\omega_p)^2} e^{-\omega_p(|t_2-t|+|t-t_1|)}, \tag{196}$$

which, on integration, gives

$$4\omega_p^2 \, I_{\text{LHS}} = \left[(t_2 - t_1) + \frac{1}{\omega_p} \right] e^{-\omega_p(t_2-t_1)}. \tag{197}$$

Consider now the right-hand side of Eq. (152), which also evaluates to

$$I_{RHS} = -\frac{\partial}{\partial \mu} G_p(t_2 - t_1) = -\frac{1}{2\omega_p} \frac{\partial}{\partial \mu}$$

$$G = -\frac{G}{2\omega_p} \frac{\partial \ln G}{\partial \omega_p} = \frac{G}{2\omega_p} \left\{ (t_2 - t_1) + \frac{1}{\omega_p} \right\} = I_{\text{LHS}}. \tag{198}$$

Either side can also be expressed as

$$I_{\text{LHS}} = I_{RHS} = \frac{G(t_2 - t_1)}{2\omega_p} \left\{ -\frac{\partial \ln G}{\partial \omega_p} \right\}$$

$$= -\frac{1}{2\omega_p} \frac{\partial G}{\partial \omega_p} = -\frac{\partial G}{\partial \mu}. \tag{199}$$

The non-relativistic limit corresponds to $\omega \approx m$ and $\omega(t_2 - t_1) \approx m(t_2 - t_1) \gg 1$. Then we find that

$$-\frac{\partial \ln G}{\partial \mu}\bigg|_{\text{NR}} = +\frac{1}{2\omega} \left\{ (t_2 - t_1) + \frac{1}{\omega} \right\} \approx \frac{t_2 - t_1}{(2m)}. \tag{200}$$

The $(t_2 - t_1)$ dependency will arise in the left-hand side of the standard composition law because of the appropriate $\theta(t)$ functions.

If one restricts the integration in Eq. (146) to just spatial coordinates of the intermediate point, we get the result:

$$\int dx \, G(x_2, x) G(x, x_1) = \int \frac{d\boldsymbol{p}}{(2\pi)^3}$$

$$\times \frac{1}{(2\omega_p)^2} e^{-\omega_p(|t_2-t|+|t-t_1|)} e^{i\boldsymbol{p}\cdot(\boldsymbol{x}_2-\boldsymbol{x}_1)}. \tag{201}$$
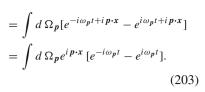
This expression, of course, is not Lorentz invariant and will depend on the intermediate time t except when $t_1 < t < t_2$. Further it is not related in any simple manner to $G(x_2, x_1)$. It is also clear from the form of the expression how it reduces to the correct non-relativistic composition law when the modulus signs are omitted in $|t_2 - t|$ and $|t - t_1|$ and we set $\omega_p = m$.

B Appendix B

This is most easily seen by manipulating this expression into the form

$$G_+(x_2; x_1) - G_-(x_2; x_1) = G_+(x_2; x_1) - G_+(x_1; x_2), \tag{202}$$

$$= \int d\,\Omega_{\boldsymbol{p}}[e^{-ipx} - e^{+ipx}]$$

$$= \int d\,\Omega_{\boldsymbol{p}}[e^{-i\omega_p t+i\boldsymbol{p}\cdot\boldsymbol{x}} - e^{i\omega_p t+i\boldsymbol{p}\cdot\boldsymbol{x}}]$$

$$= \int d\,\Omega_{\boldsymbol{p}} e^{i\boldsymbol{p}\cdot\boldsymbol{x}}[e^{-i\omega_p t} - e^{i\omega_p t}]. \tag{203}$$

To arrive at the third line, we have flipped the sign of $\boldsymbol{p}$ in the second term. Since the expression is Lorentz invariant, we can always evaluate it in the frame with $t_2 - t_1 = t = 0$, when the events are separated by a space-like interval. It vanishes, showing that $G_+(x_2; x_1) = G_-(x_2; x_1)$ when the events are separated by a space-like interval.

References

1. T. Padmanabhan, *Sleeping Beauties in Theoretical Physics* (Springer, Heidelberg, 2015)
2. T. Padmanabhan, *Quantum Field Theory: The Why, What and How* (Springer, Heidelberg, 2016)
3. C. Teitelboim, Phys. Rev. D **2**(5), 3159 (1982)
4. M. Pavšič, arXiv:1705.02774
5. A.S. Wightman, S.S. Schweber, Phys. Rev. **98**, 812–837 (1955)
6. E.B. Manoukian, Nuovo Cim. A **103**, 1495–1497 (1990)
7. P. Teller, *An Interpretative Introduction to Quantum Field Theory* (Princeton University Press, Princeton, 1995)
8. F. Buscemi, G. Compagno, Phys. Lett. A **334**, 357–362 (2005)
9. T. Padmanabhan, Unpublished work (2013)
10. L.P. Horwitz, *Relativistic Quantum Mechanics* (Springer, Berlin, 2015)
11. D. Alba, H. Crater, L. Lusanna, J. Math. Phys. **52**, 062301 (2011)
12. H. Crater, L. Lusanna, Int. J. Mod. Phys. A **29**, 1450163 (2014)
13. G.N. Fleming, J. Butterfield, in *From Physics to Philosophy, chapter 6*, ed. by J. Butterfield, C. Pagonis (Cambridge University Press, Cambridge, 1999), pp. 108–165
14. D. Malament, In defense of dogma: why there cannot be a relativistic quantum mechanics of (localizable) particles, in *Perspectives on Quantum Reality*, ed. by R. Clifton (Kluwer, Dordrecht, 1996), pp. 1–10
15. A. Anderson, Phys. Rev. D **49**, 4049 (1994)
16. I.H. Redmount, W.-M. Suen, Int. J. Mod. Phys. A **8**, 1629 (1993)
17. J.B. Hartel, K.V. Kuchar, Phys. Rev. D **34**, 2323 (1986)
18. J.J. Halliwell, J. Thorwart, Phys. Rev. D **64**, 124018 (2001)
19. T. Newton, E. Wigner, Rev. Mod. Phys. **21**, 400–406 (1949)
20. A.S. Wightman, Rev. Mod. Phys. **34**, 845–872 (1962)
21. A.J. Kálnay, Phys. Rev. D **1**, 1092–1104 (1969)
22. T.W. Ruijgrok, *On Localization in Relativistic Quantum Mechanics, Theoretical Physics Fin de Siècle, vol 539, Lecture Notes on Physics* (Springer, Heidelberg, 2000)
23. N. Barat, J.C. Kimball, Phys. Lett. A **308**, 110–115 (2003)
24. R.M. Mir-Kasimov, Phys. Part. Nucl. Lett. **3**, 280–289 (2006)
25. D.J. Cirilo-Lombardo, J. Math. Phys. **57**, 063503 (2016). https://doi.org/10.1063/1.4953368. arXiv:1610.03624 [hep-th]
26. L. O. Herrmann, *Localization in Relativistic Quantum Theories*, PhilSci Archive (2010). http://philsci-archive.pitt.edu/5427/
27. A.H. Monahan, M. McMillan, Phys. Rev. A **56**, 2563–2566 (1997)
28. J.J. Halliwell, M.E. Ortiz, Phys. Rev. D **48**, 748 (1993)
29. O. Rudolph, Phys. Rev. D **51**, 1818 (1995)
30. J.J. Halliwell, Phys. Rev. D **38**, 2468 (1988)
31. R. Ferraro, Phys. Rev. D **45**, 1198 (1992)
32. E. Seidewitz, Found. Phys. **37**, 572 (2007)
33. H. Ikemori, Phys. Rev. D **40**, 3512 (1989)
34. E.S. Fradkin, D.M. Gitman, Phys. Rev. D **44**, 3230 (1991)

35. E.C.G. Stueckelberg, Helv. Phys. Acta **14**, 588 (1941)
36. E.C.G. Stueckelberg, Helv. Phys. Acta **15**, 588 (1942)
37. C. Teitelboim, Phys. Rev. Lett **50**, 705 (1983)
38. C. Teitelboim, Phys. Rev. D **25**, 3159 (1982)
39. M. Henneaux, C. Teitelboim, Ann. Phys. (NY) **143**, 127 (1982)
40. Y. Feng, arXiv:1210.6418
41. U.H. Niederer, L.O. Raifeartaigh, Fortschr. Phys. **22**, 131–157 (1974)

42. T. Padmanabhan, Ann. Phys. **165**, 38 (1985)
43. T. Padmanabhan, Class. Quanum Gravity **4**, L107 (1987)
44. T. Padmanabhan, Phys. Rev. Lett. **78**, 1854 (1997). arXiv:hep-th-9608182
45. T. Padmanabhan, Phys. Rev. D **57**, 6206 (1998)