Sum-over-histories origin of the composition laws of relativistic quantum mechanics and quantum cosmology

Jonathan J. Halliwell*
Center for Theoretical Physics, Laboratory for Nuclear Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
Blackett Laboratory, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BZ, United Kingdom†

Miguel E. Ortiz‡
Center for Theoretical Physics, Laboratory for Nuclear Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 2 November 1992)

This paper is concerned with the question of the existence of composition laws in the sum-over-histories approach to relativistic quantum mechanics and quantum cosmology, and its connection with the existence of a canonical formulation. In nonrelativistic quantum mechanics, the propagator is represented by a sum over histories in which the paths move forward in time. The composition law of the propagator then follows from the fact that the paths intersect an intermediate surface of constant time once and only once, and a *partition* of the paths according to their crossing position may be affected. In relativistic quantum mechanics, by contrast, the propagators (or Green functions) may be represented by sums over histories in which the paths move backward and forward in time. They therefore intersect surfaces of constant time more than once, and the relativistic composition law, involving a normal derivative term, is not readily recovered. The principal technical aim of this paper is to show that the relativistic composition law may, in fact, be derived directly from a sum over histories by partitioning the paths according to their *first* crossing position of an intermediate surface. We review the various Green functions of the Klein-Gordon equation, and derive their composition laws. We obtain path-integral representations for all Green functions except the causal one. We use the proper time representation, in which the path integral has the form of a nonrelativistic sum over histories but is integrated over time. The question of deriving the composition laws therefore reduces to the question of factoring the propagators of nonrelativistic quantum mechanics across an arbitrary surface in configuration space. This may be achieved using a known result called the path decomposition expansion (PDX). We give a proof of the PDX using a spacetime lattice definition of the Euclidean propagator. We use the PDX to derive the composition laws of relativistic quantum mechanics from the sum over histories. We also derive canonical representations of all of the Green functions of relativistic quantum mechanics, i.e., express them in the form $\langle x''|x'\rangle$, where the $\{|x\rangle\}$ are a complete set of configuration-space eigenstates. These representations make it clear why the Hadamard Green function $G^{(1)}$ does not obey a standard composition law. They also give a hint as to why the causal Green function does not appear to possess a sum-over-histories representation. We discuss the broader implications of our methods and results for quantum cosmology, and parametrized theories generally. We show that there is a close parallel between the existence of a composition law and the existence of a canonical formulation, in that both are dependent on the presence of a timelike Killing vector. We also show why certain naive composition laws that have been proposed in the past for quantum cosmology are incorrect. Our results suggest that the propagation amplitude between three-metrics in quantum cosmology, as constructed from the sum over histories, does not obey a composition law.

PACS number(s): 03.65.Ca, 04.60.+n, 11.10.Qr

I. INTRODUCTION

Quantum theory, in both its development and applications involves two strikingly different sets of mathematical tools. On the one hand, there is the canonical approach, involving operators, states, Hilbert spaces and Hamiltonians. On the other, there is the path integral, involving sums over sets of histories. For most purposes, the distinction between these two methods is largely regarded as a matter of mathematical rigor or calculational convenience. There may, however, be a more fundamental distinction: one method could be more general than the other. If this is the case, then it is of particular interest to explore the connections between the two formulations, and discover the conditions under which a route from one method to the other can or cannot be found.

A particular context in which the possible distinction

*Electronic address: j__halliwell@vax1.physics.imperial.ac.uk
†Present address.
‡Electronic address: ortiz@mitlns.mit.edu

between these two quantization methods will be important is quantum cosmology. There, the canonical formulation suffers from a serious obstruction known as the "problem of time" [1,2]. This is the problem that general relativity does not obviously supply the preferred time parameter so central to the formulation and interpretation of quantum theory. By contrast, in sum-over-histories formulations of quantum theory, the central notion is that of a quantum-mechanical history. The notion of time does not obviously enter in an essential way. Sum-over-histories formulations of quantum cosmology have therefore been promoted as promising candidates for a quantum theory of spacetime, because the problem of time is not as immediate or central, and may even be sidestepped completely [3]. In particular, as suggested by Hartle, a sum-over-histories formulation could exist even though a canonical formulation may not [3]. The broad aim of this paper is to explore this suggestion.

An object that one would expect to play an important role in sum-over-histories formulations of quantum cosmology is the "propagation amplitude" between three-metrics. Formally, it is given by a functional integral expression of the form

$$G(h_{ij}'',h_{ij}')=\int \mathcal{D}g_{\mu\nu}\exp(iS[g_{\mu\nu}]) \ .$$

Here, $S[g_{\mu\nu}]$ is the gravitational action, and the sum is over a class of four-metrics matching the prescribed three-metrics h_{ij}'', h_{ij}' on final and initial surfaces. The level of the present discussion is rather formal, so we will not go into the details of how such an expression is constructed (see [4] for details), nor shall we address the important question of its interpretation. It is, however, important for present purposes to assume that a definition of the sum over histories exists that is not dependent on the canonical formalism.

The above expression is closely analogous in its construction to the sum-over-histories representations of the propagators (or Green functions) of relativistic quantum mechanics, $\mathcal{G}(x''|x')$, where x denotes a spacetime coordinate. We shall make heavy use of this analogy in this paper.

In relativistic quantum mechanics, there exist both sum-over-histories and canonical formulations of the one-particle quantum theory. In the canonical formulation, one may introduce a complete set of configuration space states, $\{|x\rangle\}$. The propagators may then be shown to possess canonical representations; i.e., they may be expressed in the form

$$\mathcal{G}(x''|x')=\langle x''|x'\rangle \ ,$$

where the right-hand side denotes a genuine Hilbert space inner product. By insertion of a resolution of the identity, it may then be shown that the propagator satisfies a composition law, typically of the form

$$\langle x''|x'\rangle=\int d\sigma^{\mu}\langle x''|x\rangle\partial_{\mu}\langle x|x'\rangle \ ,$$

where $d\sigma^{\mu}$ denotes a normal surface element. The details of this type of construction will be given in later sections. For the moment, the point to stress is that the existence of a composition law is generally closely tied to the existence of canonical representations of $\mathcal{G}(x''|x')$.

Now in a sum-over-histories formulation of quantum cosmology, the path-integral representation of $G(h_{ij}'',h_{ij}')$ is taken to be the starting point. Relations such as the composition law, characteristic of canonical formulations, cannot be assumed but hold only if they can be *derived* directly from the sum over histories alone, without recourse to a canonical formulation. In particular, since the existence of a composition law seems to be a general feature of the canonical formalism, it is very reasonable to suppose that the existence of a composition law for a $G(h_{ij}'',h_{ij}')$ generated by the sum over histories is a *necessary condition* for the existence of an equivalent canonical formulation.

The object of this paper is to determine how a derivation of the composition law from the sum over histories may be carried out. We may then ask how this derivation might fail, i.e., whether the necessary condition for the recovery of a canonical formulation of quantum cosmology from a sum-over-histories formulation is satisfied.

Of course, a full quantum theory of cosmology, even if it existed, would be exceedingly complicated. Like many authors declaring interest in quantum cosmology, therefore, we will focus on the technically simpler case of the relativistic particle. As stated above, relativistic quantum mechanics possesses many of the essential features of quantum cosmology. Remarks on quantum cosmology of a more general and speculative nature will be saved until the end. We shall show how the composition laws of relativistic propagators may be derived directly from their sum-over-histories representations. To the best of our knowledge, this derivation has not been given previously. It is therefore of interest in the limited context of relativistic quantum mechanics, as well as being a model for the more difficult problem of quantum cosmology outlined above.

A. The problem

In nonrelativistic quantum mechanics the propagator $\langle \mathbf{x}'',t''|\mathbf{x}',t'\rangle$ plays a useful and important role. It is defined to be the object which satisfies the Schrödinger equation with respect to each argument,

$$\left[i\frac{\partial}{\partial t''}-\hat{H}''\right]\langle \mathbf{x}'',t''|\mathbf{x}',t'\rangle=0 \qquad (1.1)$$

(and similarly for the initial point), subject to the initial condition

$$\langle \mathbf{x}'',t'|\mathbf{x}',t'\rangle=\delta^{(n)}(\mathbf{x}''-\mathbf{x}') \ . \qquad (1.2)$$

It determines the solution to the Schrödinger equation at time t'', given initial data at time t':

$$\Psi(\mathbf{x}'',t'')=\int d^n\mathbf{x}'\langle \mathbf{x}'',t''|\mathbf{x}',t'\rangle\Psi(\mathbf{x}',t') \ . \qquad (1.3)$$

From this follows the composition law (semigroup property)

$$\langle \mathbf{x}'',t''|\mathbf{x}',t'\rangle=\int d^n\mathbf{x}\langle \mathbf{x}'',t''|\mathbf{x},t\rangle\langle \mathbf{x},t|\mathbf{x}',t'\rangle \ . \qquad (1.4)$$

In relativistic quantum mechanics, the most closely analogous object is at first sight the causal propagator $G(x''|x')$. It is defined to be the object satisfying the Klein-Gordon equation with respect to each argument,

$$(\Box_{x''}+m^2)G(x''|x')=0 \tag{1.5}$$

(and similarly for the initial point), and obeying the boundary conditions

$$G(\mathbf{x}'',x^{0''}|\mathbf{x}',x^{0'})|_{x^{0''}=x^{0'}}=0 \ ,$$

$$\frac{\partial}{\partial x^{0''}}G(\mathbf{x}'',x^{0''}|\mathbf{x}',x^{0'})|_{x^{0''}=x^{0'}}=-\delta^{(3)}(\mathbf{x}''-\mathbf{x}') \ .$$

It vanishes outside the light cone. It determines the solution at a spacetime point x'', given initial data on the spacelike surface Σ:

$$\phi(x'')=-\int_{\Sigma} d\sigma^{\mu}G(x''|x')\overleftrightarrow{\partial}_{\mu}\phi(x') \ , \tag{1.6}$$

where

$$\overleftrightarrow{\partial}_{\mu}=\overrightarrow{\partial}_{\mu}-\overleftarrow{\partial}_{\mu} \ , \tag{1.7}$$

and $d\sigma^{\mu}$ is normal to the surface Σ in the future timelike direction. From (1.6) follows the composition law

$$G(x''|x')=-\int_{\Sigma} d\sigma^{\mu}G(x''|x)\overleftrightarrow{\partial}_{\mu}G(x|x') \ . \tag{1.8}$$

There are of course a number of other Green functions associated with the Klein-Gordon equation, and many of them also obey composition laws similar to (1.8), involving the derivative operator (1.7) characteristic of relativistic field theories. For example, the Feynman Green function obeys a slightly modified version of (1.8).

Because of the presence of the derivative operator (1.7) in (1.8), the relativistic and nonrelativistic composition laws assume a somewhat different form. The difference is readily understood. The wave functions of nonrelativistic quantum mechanics obey a parabolic equation, and so are uniquely determined by the value of the wave function on some initial surface. In contrast, the wave functions in the relativistic case obey a hyperbolic equation, and so are uniquely determined by the value of the wave function and its normal derivative on some initial surface, hence the derivative term in (1.8).

A convenient way of representing the propagator in nonrelativistic quantum mechanics is in terms of a sum over histories. Formally, one writes

$$\langle \mathbf{x}'',t''|\mathbf{x}',t'\rangle=\sum_{p(\mathbf{x}',t'\to\mathbf{x}'',t'')}\exp[iS(\mathbf{x}',t'\to\mathbf{x}'',t'')] \ . \tag{1.9}$$

Here, $p(\mathbf{x}',t'\to\mathbf{x}'',t'')$ denotes the set of paths beginning at $\mathbf{x}'$ at time t' and ending at $\mathbf{x}''$ at t'', and $S(\mathbf{x}',t'\to\mathbf{x}'',t'')$ denotes the action of each individual such path. The propagator of nonrelativistic quantum mechanics is obtained by restricting to paths $\mathbf{x}(t)$ that are single-valued functions of t, that is, they *move forward in time.* There are many ways of defining a formal object such as (1.9). A common method worth keeping in mind is the time-slicing definition, in which the time interval is divided into N equal parts of size ϵ, $N\epsilon=(t''-t')$, and one writes

$$\langle \mathbf{x}'',t''|\mathbf{x}',t'\rangle = \lim_{N\to\infty}\prod_{k=1}^{N}\int\frac{d^n\mathbf{x}_k}{(2\pi i\epsilon)^{n/2}}\exp[iS(\mathbf{x}_{k+1},t_{k+1}|\mathbf{x}_k,t_k)] \ .$$

Here, $\mathbf{x}_0=\mathbf{x}'$, $t_0=t'$, $\mathbf{x}_{N+1}=\mathbf{x}''$, $t_{N+1}=t''$ and $S(\mathbf{x}_{k+1},t_{k+1}|\mathbf{x}_k,t_k)$ is the action of the classical path connecting $(\mathbf{x}_k,t_k)$ to $(\mathbf{x}_{k+1},t_{k+1})$. More rigorous definitions also exist, such as that in which (the Euclidean version of) (1.9) is defined as the continuum limit of a sum over paths on a discrete spacetime lattice. Indeed, we will find it necessary to resort to such a rigorous definition below.

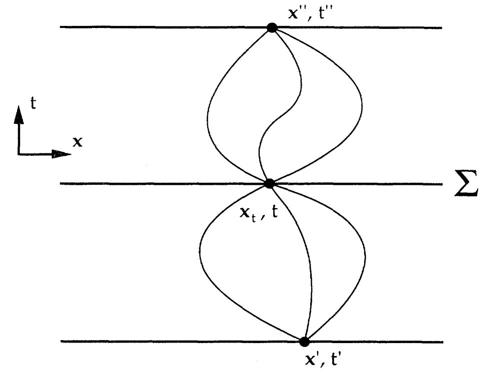

FIG. 1. Paths for the nonrelativistic propagator in the set $p(\mathbf{x}',t'\to\mathbf{x}_t,t\to\mathbf{x}'',t'')$.

Given the representation (1.9) of the propagator, it becomes pertinent to ask whether the composition law (1.4) may be derived directly from the sum-over-histories representation (1.9). This is indeed possible. The crucial notion permitting such a derivation is that of an exclusive *partition* of the histories into mutually exclusive alternatives. Consider the surface labeled by t, where $t'\le t\le t''$. Because the paths move forward in time, each path intersects this surface once and only once, at some point $\mathbf{x}_t$, say. The paths may therefore be exhaustively partitioned into mutually exclusive sets, according to the value of $\mathbf{x}$ at which they intersect the surface labeled by t (see Fig. 1). We write this as

$$p(\mathbf{x}',t'\to\mathbf{x}'',t'')=\bigcup_{\mathbf{x}_t} p(\mathbf{x}',t'\to\mathbf{x}_t,t\to\mathbf{x}'',t'') \ ,$$

$$p(\mathbf{x}',t'\to\mathbf{x}_t,t\to\mathbf{x}'',t'')\cap p(\mathbf{x}',t'\to\mathbf{y}_t,t\to\mathbf{x}'',t'')=\varnothing \ \text{ if } \mathbf{x}_t\neq\mathbf{y}_t \ .$$

Each path from $(\mathbf{x}',t')$ to $(\mathbf{x}'',t'')$ may then be uniquely expressed as the composition of a path from $(\mathbf{x}',t')$ to $(\mathbf{x}_t,t)$ with a path from $(\mathbf{x}_t,t)$ to $(\mathbf{x}'',t'')$ for some $\mathbf{x}_t$. Consider what this implies for the sum over histories. First of all, any sensible definition of the measure in the sum over histories should satisfy

$$\sum_{p(\mathbf{x}',t'\to\mathbf{x}'',t'')} = \sum_{\mathbf{x}_t} \sum_{p(\mathbf{x}',t'\to\mathbf{x}_t,t\to\mathbf{x}'',t'')} = \sum_{\mathbf{x}_t} \sum_{p(\mathbf{x}',t'\to\mathbf{x}_t,t)} \sum_{p(\mathbf{x}_t,t\to\mathbf{x}'',t'')} . \tag{1.10}$$

This is readily shown to be true of the time-slicing definition, for example. Second, the action should satisfy

$$S(\mathbf{x}',t'\to\mathbf{x}'',t'')=S(\mathbf{x}',t'\to\mathbf{x}_t,t)+S(\mathbf{x}_t,t\to\mathbf{x}'',t'') . \tag{1.11}$$

Combining (1.10) and (1.11), it is readily seen that one has

$$\begin{aligned}\langle \mathbf{x}'',t''|\mathbf{x}',t'\rangle &= \sum_{\mathbf{x}_t} \sum_{p(\mathbf{x}',t'\to\mathbf{x}_t,t)} \sum_{p(\mathbf{x}_t,t\to\mathbf{x}'',t'')} \exp[iS(\mathbf{x}',t'\to\mathbf{x}_t,t)+iS(\mathbf{x}_t,t\to\mathbf{x}'',t'')] \\ &= \sum_{\mathbf{x}_t} \langle \mathbf{x}'',t''|\mathbf{x}_t,t\rangle\langle \mathbf{x}_t,t|\mathbf{x}',t'\rangle .\end{aligned} \tag{1.12}$$

The composition law therefore follows directly from the partitioning of the sets of paths in the sum over histories.

Turn now to the relativistic particle. There also certain Green functions may be represented by sums over histories. Formally, one writes

$$\mathcal{G}(x''|x')=\sum_{p(x'\to x'')} \exp[iS(x'\to x'')] \tag{1.13}$$

(we will be precise later about which Green function $\mathcal{G}$ may be). In fact, a number of such representations are available, since the classical relativistic particle is a constrained system, and there is more than one way of constructing the path integral for constrained systems [5]. Here we shall be largely concerned with those constructions for which the set of paths summed over in (1.13) is all paths in spacetime. In particular, unlike the nonrelativistic case, the paths will generally move *forward and backward* in the time coordinate, x^0 (see Fig. 2).

It again becomes pertinent to ask whether a composition law of the form (1.8) may be derived from the sum-over-histories representation. However, because the paths move both backward and forward in time, they typically intersect an intermediate surface of constant x^0 many times, and the points at which they intersect the intermediate surface therefore *do not* effect a partition of the paths into exclusive sets. The argument for the nonrelativistic case, therefore, cannot be carried over directly to the relativistic case. Furthermore, even if this partition did work, it would then not be clear how the derivative term in the composition law might arise from the path representation (1.13). We are thus led to the question, is there a different way of partitioning the paths, that leads to a composition law of the form (1.8), and explains the appearance of the derivative term? This question is the topic of this paper.

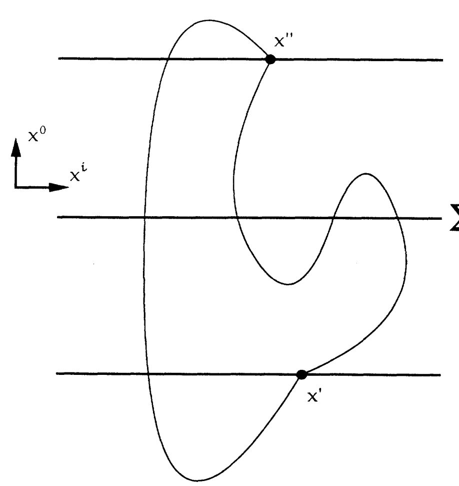

FIG. 2. Paths for the relativistic propagator.

In detail, we will study sum-over-histories expressions of the form (1.13) for relativistic Green functions. We will focus on the "proper time" sum over histories, in which the Green functions are represented by an expression of the form

$$\mathcal{G}(x''|x')=\int dT\, g(x'',T|x',0) , \tag{1.14}$$

Here $g(x'',T|x',0)$ is a Schrödinger operator satisfying (1.1), (1.2), and (1.4), with the Hamiltonian taken to be the Klein-Gordon operator in (1.5). g may therefore be represented by a sum over paths of the form (1.9). We will derive (1.14) below, but for the moment note that (1.14) will be a solution to (1.5) if T is taken to have an infinite range, and will satisfy (1.5) but with a δ function on the right-hand side if T is taken to have a half-infinite range.

B. Outline

We begin in Sec. II by reviewing the various Green functions associated with the Klein-Gordon equation and their properties. We determine which Green functions satisfy a composition law of the form (1.6). We briefly describe the sum-over-histories representation, and derive (1.14). An important question we address is that of which Green functions are obtained by the sum over paths (1.14). We also discuss the connection of sum-over-histories representations with canonical representations. By this we mean representations in which the propagators may be expressed in the form $\langle x''|x'\rangle$, where the $\{|x\rangle\}$ with a single time argument x^0 form a complete set of configuration-space eigenstates.

In the representation (1.14), the time coordinate x^0 is treated as a "spatial" coordinate, when g is thought of as an ordinary Schrödinger propagator like that of nonrelativistic quantum mechanics. Comparing (1.14) with the expression to be derived from it, (1.8), we therefore see that our problem of factoring the sum over histories (1.14) across a surface of constant x^0 is very closely related to that of factoring the sum over histories (1.9), not across a surface of constant parameter time t, as in (1.4), but across a surface on which one of the spatial coordinates is constant. It turns out that a solution to this problem exists, and the result goes by the name of the path decomposition expansion (PDX) [6]. The crucial observation that leads to this result is that although the paths may cross the factoring surface many times, they may nevertheless be partitioned into exclusive sets according to the parameter time and spatial location of their *first crossing* of the surface. We describe this result in Sec. III, and give a rigorous derivation of it.

In Sec. IV, we give our main result. This is to show how the composition law (1.8) follows from the sum over histories (1.14), using the PDX. We also explain why certain naive composition laws that have been proposed in the past are problematic.

Our principal result is admittedly simple, and has been derived largely by straightforward application of the PDX. However, it has broader significance in the context of the sum-over-histories approach to quantum theory. In particular, it is closely related to the question of the conditions under which a sum-over-histories formulation of quantum theory implies the existence of a Hilbert space formulation. In Sec. V, we therefore discuss the generalizations and broader implications of our result.

II. THE PROPAGATORS OF RELATIVISTIC QUANTUM MECHANICS

A. Green functions of the Klein-Gordon equation

We begin this section with a review of the various Green functions of the Klein-Gordon equation in Minkowski space relevant to our discussion. The section is intended to set out the conventions we shall use throughout this paper, and to list the relevant properties of the Green functions. A metric of signature $(+,-,-,-)$ is used throughout. Readers familiar with the intricacies of this subject may wish to move directly to Sec. II B.

The kernel $\mathcal{G}(x|y)$ of the operator $(\Box+m^2)$, satisfying

$$(\Box_x+m^2)\mathcal{G}(x|y)=-\delta^4(x-y)\ , \tag{2.1}$$

where x and y are four-vectors, may be shown by Fourier transformation to be given by the expression

$$\mathcal{G}(x|y)=\frac{1}{(2\pi)^4}\int d^4k\frac{e^{-ik\cdot(x-y)}}{k^2-m^2}\ . \tag{2.2}$$

$\mathcal{G}(x|y)$ is not uniquely defined in Minkowski space due to the presence of poles in the integrand. The k_0 integration

$$\int_{-\infty}^{\infty}dk_0\frac{e^{ik_0(x^0-y^0)}}{{k_0}^2-\mathbf{k}^2-m^2}$$

has poles on the real axis at $k_0=\pm(\mathbf{k}^2+m^2)^{1/2}$, and the various possible deformations of this contour determine the possible solutions to (2.1), each with different support properties. Below we shall list some possible contours and their corresponding Green functions. Closed contours yield solutions to the Klein-Gordon equation. We also discuss these below since they play an important role in relativistic quantum mechanics.

1. *Wightman functions: $G^+(x|y)$ and $G^-(x|y)$*

A closed anticlockwise contour around one or other of the poles yields the Wightman functions $\pm iG^{\pm}(x|y)$, which are solutions of the Klein-Gordon equation, and of its positive and negative square roots respectively:

$$\left[i\frac{\partial}{\partial x^0}\mp(m^2-\nabla_{\mathbf{x}}^2)^{1/2}\right]G^{\pm}(x|y)=0\ .$$

They are given by

$$G^{\pm}(x|y)=\frac{1}{(2\pi)^3}\int d^4k\ \theta(k^0)\delta(k^2-m^2)e^{\mp ik\cdot(x-y)}$$

or

$$G^{\pm}(x|y)=\frac{1}{(2\pi)^3}\int_{k_0=\pm\omega_{\mathbf{k}}}\frac{d^3\mathbf{k}}{2\omega_{\mathbf{k}}}e^{-ik\cdot(x-y)}\ ,$$

and are related by

$$G^+(x|y)=G^-(y|x)\ .$$

The two Wightman functions satisfy relativistic composition laws

$$G^{\pm}(x''|x')=\pm i\int_{\Sigma}d\sigma^{\mu}G^{\pm}(x''|x)\overleftrightarrow{\partial}_{\mu}G^{\pm}(x|x') \tag{2.3}$$

(where $d\sigma^{\mu}$ is normal to Σ and future pointing) and are orthogonal in the sense that

$$\int_{\Sigma}d\sigma^{\mu}G^{\pm}(x''|x)\overleftrightarrow{\partial}_{\mu}G^{\mp}(x|x')=0\ .$$

In field theory they are given by the expressions

$$G^+(x|y)=\langle 0|\phi(x)\phi(y)|0\rangle$$

and

$$G^-(x|y)=\langle 0|\phi(y)\phi(x)|0\rangle\ .$$

2. *Feynman propagator: $G_F(x|y)$*

A contour going under the left pole and above the right gives the Feynman propagator. This satisfies Eq. (2.1), and may be written as

$$iG_F(x|y)=\theta(x^0-y^0)G^+(x|y)+\theta(y^0-x^0)G^-(x|y)\ .$$

Alternatively,

$$G_F(x|y)=\frac{-i}{(2\pi)^4}\int_0^{\infty}dT\int d^4k\ e^{-i[k\cdot(x-y)-T(k^2-m^2+i\varepsilon)]}$$
$$=\frac{1}{(2\pi)^4}\int d^4k\frac{e^{-ik\cdot(x-y)}}{k^2-m^2+i\varepsilon}\ .$$

It may be checked that G_F obeys a relativistic composition law

$$G_F(x''|x')=-\int_{\Sigma}d\sigma\ G_F(x''|x)\overleftrightarrow{\partial}_n G_F(x|x')\ , \tag{2.4}$$

where Σ is an arbitrary spacelike three-surface, and $\partial_n=n^{\mu}\partial_{\mu}$ with n^{μ} now the normal to Σ in the direction of propagation. In free scalar field theory, the Feynman propagator is of course given by

$$iG_F(x|y)=\langle 0|T(\phi(x)\phi(y))|0\rangle \ .$$

3. *Causal Green function:* $G(x|y)$

A closed clockwise contour around both poles gives what is generally known as the commutator or causal Green function, which is written simply as $G(x|y)$. It is a solution of the Klein-Gordon equation. $G(x|y)$ has the representations

$$G(x|y)=\frac{-i}{(2\pi)^3}\int d^4k\ \varepsilon(k^0)\delta(k^2-m^2)e^{-ik\cdot(x-y)} \tag{2.5}$$

or

$$G(x|y)=\frac{-1}{(2\pi)^3}\int \frac{d^3\mathbf{k}}{\omega_\mathbf{k}}\sin[\omega_\mathbf{k}(x^0-y^0)]e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} \ .$$

Since

$$G(\mathbf{x},x^0|\mathbf{y},y^0)|_{x^0=y^0}=0 \ ,$$

$$\left.\frac{\partial}{\partial x^0}G(\mathbf{x},x^0|\mathbf{y},y^0)\right|_{x^0=y^0}=-\delta^3(\mathbf{x}-\mathbf{y}) \ ,$$

and G is Lorentz invariant, it has support only within the light cone of $x-y$. G also obeys the relativistic composition law

$$G(x''|x')=-\int_\Sigma d\sigma^\mu G(x''|x)\overleftrightarrow{\partial}_\mu G(x|x') \ ,$$

and, as mentioned in the Introduction, it propagates solutions $\phi(x)$ of the Klein-Gordon equation via

$$\phi(y)=-\int_\Sigma d\sigma^\mu G(y|x)\overleftrightarrow{\partial}_\mu \phi(x) \ .$$

In field theory, G is given by the commutator

$$iG(x|y)=\langle 0|[\phi(x),\phi(y)]|0\rangle=[\phi(x),\phi(y)] \ .$$

Finally, note that

$$iG(x|y)=G^+(x|y)-G^-(x|y) \ .$$

4. *Hadamard Green function:* $G^{(1)}(x|y)$

A closed figure of eight contour around the two poles gives the Hadamard or Schwinger Green function $iG^{(1)}(x|y)$, which is a solution of the Klein-Gordon equation. It may be written as

$$\begin{aligned}G^{(1)}(x|y)&=\frac{1}{(2\pi)^3}\int d^4k\ \delta(k^2-m^2)e^{-ik\cdot(x-y)}\\ &=\frac{1}{(2\pi)^4}\int_{-\infty}^{\infty}dT\int d^4k\ e^{-i[k\cdot(x-y)-T(k^2-m^2)]}\end{aligned} \tag{2.6}$$

or

$$G^{(1)}(x|y)=\frac{1}{(2\pi)^3}\int \frac{d^3\mathbf{k}}{\omega_\mathbf{k}}\cos[\omega_\mathbf{k}(x^0-y^0)]e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} \ .$$

Perhaps the most important property of $G^{(1)}(x|y)$ is that it does not satisfy the standard relativistic composition law. In fact

$$\begin{aligned}G^{(1)}(x''|x')&=-\int_\Sigma d\sigma^\mu G(x''|x)\overleftrightarrow{\partial}_\mu G^{(1)}(x|x')\\ &=-\int_\Sigma d\sigma^\mu G^{(1)}(x''|x)\overleftrightarrow{\partial}_\mu G(x|x')\end{aligned} \tag{2.7}$$

and

$$G(x''|x')=\int_\Sigma d\sigma^\mu G^{(1)}(x''|x)\overleftrightarrow{\partial}_\mu G^{(1)}(x|x') \ .$$

In field theory, $G^{(1)}$ is given by the anticommutator

$$G^{(1)}(x|y)=\langle 0|\{\phi(x),\phi(y)\}|0\rangle \ .$$

It is related to the Wightman functions via

$$G^{(1)}(x|y)=G^+(x|y)+G^-(x|y) \ .$$

5. *Newton-Wigner propagator:* $G_{\mathrm{NW}}(\mathbf{x},x^0|\mathbf{y},y^0)$

The Newton-Wigner propagator is a solution of the Klein-Gordon equation, and indeed of its first-order positive square root. It is not given by the integral (2.2) for any contour. We nevertheless include it since it plays an important role in the quantum mechanics of the relativistic particle. G_{NW} is defined by

$$G_{\mathrm{NW}}(\mathbf{x},x^0|\mathbf{y},y^0)=\frac{1}{(2\pi)^3}\int_{k_0=\omega_\mathbf{k}} d^3\mathbf{k}\ e^{-ik\cdot(x-y)} \ . \tag{2.8}$$

The support property

$$G_{\mathrm{NW}}(\mathbf{x},x^0|\mathbf{y},y^0)|_{x^0=y^0}=\delta^3(\mathbf{x}-\mathbf{y})$$

shows that G_{NW} is analogous to the quantum-mechanical propagator (1.2). It propagates solutions of the first-order Schrödinger equation with Hamiltonian $H=(\mathbf{k}^2-m^2)^{1/2}$. G_{NW} also obeys the usual quantum-mechanical composition law:

$$\begin{aligned}G_{\mathrm{NW}}(\mathbf{x}'',x^{0''}|\mathbf{x},x^0)=&\int d^3\mathbf{x}' G_{\mathrm{NW}}(\mathbf{x}'',x^{0''}|\mathbf{x}',x^{0'})\\ &\times G_{\mathrm{NW}}(\mathbf{x}',x^{0'}|\mathbf{x},x^0) \ .\end{aligned} \tag{2.9}$$

Finally, note that G_{NW} is not Lorentz invariant.

An analogous operator, which we shall call the negative-frequency Newton-Wigner propagator, may also be defined. It is given by

$$\tilde{G}_{\mathrm{NW}}(\mathbf{x},x^0|\mathbf{y},y^0)=\frac{1}{(2\pi)^3}\int_{k_0=-\omega_\mathbf{k}} d^3\mathbf{k}\ e^{-ik\cdot(x-y)}$$

and has the same support properties as $G_{\mathrm{NW}}(\mathbf{x},x^0|\mathbf{y},y^0)$. It solves the negative-frequency square root of the Klein-Gordon equation and propagates its solutions.

B. Sum-over-histories formulation of relativistic quantum mechanics

We are interested in Green functions which may be represented by sums over histories of the form (1.13). We will take sum-over-histories expressions of the form (1.13) as our starting point and determine which Green functions they give rise to. The expression (1.13) is rather formal as it stands, and various aspects of it need to be specified more precisely before it is properly and uniquely defined. These include the action, class of paths, gauge-fixing conditions, and the domains of integration of certain variables. The particular Green function obtained will depend on how these particular features are specified. We note, however, that there is no guarantee that *all* known Green functions may be obtained in this way, and indeed, we are not able to obtain the causal propagator, G.

The action for a relativistic particle is usually written as

$$S=-m\int_{\tau'}^{\tau''}d\tau\left[\frac{\partial x^{\mu}}{\partial\tau}\frac{\partial x^{\nu}}{\partial\tau}\eta_{\mu\nu}\right]^{1/2}, \qquad (2.10)$$

the length of the world line of the particle in Minkowski space. τ parametrizes the world line, and S is invariant under reparametrizations $\tau\rightarrow f(\tau)$. Since (2.10) is highly nonlinear, its quantization presents certain difficulties which have hitherto prevented its direct use in a sum over histories. These difficulties may be bypassed by the introduction of an auxiliary variable N, which can be thought of as a metric on the particle world line. The action may then be rewritten as

$$S=-\int_{\tau'}^{\tau''}d\tau\left[\frac{\dot{x}^2}{4N}+m^2N\right],$$

where a dot denotes a derivative with respect to τ. Passing to a Hamiltonian form, the action becomes

$$S=\int_{\tau'}^{\tau''}d\tau[p_{\mu}\dot{x}^{\mu}+NH], \qquad (2.11)$$

where N is now a Lagrange multiplier which enforces the constraint $H=(p^2-m^2)=0$. The Hamiltonian form (2.11) of the action is still invariant under reparametrizations. Infinitesimally, these are generated by the constraint H:

$$\delta x=\varepsilon(\tau)\{x,H\},\quad \delta p=\varepsilon(\tau)\{p,H\},\quad \delta N=\dot{\varepsilon}(\tau)$$

for some arbitrary parameter $\varepsilon(\tau)$. Since H is quadratic in momentum, the action is only invariant up to a surface term [7]

$$\delta S=\left[\varepsilon(\tau)\left[p\frac{\partial H}{\partial p}-H\right]\right]_{\tau'}^{\tau''},$$

which constrains the reparametrizations at the end points to obey

$$\varepsilon(\tau'')=0=\varepsilon(\tau').$$

We shall discuss briefly the use of a sum over histories to evaluate the amplitude for a transition from x' to x'', which we shall write as $\mathcal{G}(x''|x')$. The sum is over paths beginning at x' at parameter time τ', and ending at x'' at parameter time τ''. Trajectories may in principle move forward and backward in the physical time x^0, although it is also possible to define an amplitude constructed from paths that move only forward in x^0, as we shall discuss below.

It is necessary to fix the reparametrization invariance, and this may be done in a number of ways. We shall give a brief description of the two most commonly used prescriptions: the so-called proper time gauge $\dot{N}=0$, and the canonical gauge $x^0=\tau$. The proper time gauge is a good prescription in the Gribov sense [5]. The canonical gauge has the feature that it restricts the class of paths in configuration space to move forwards in the time coordinate x^0. These two gauge-fixing conditions lead to quite different results. There is, however, no conflict with the standard result that the path integral is independent of the choice of gauge fixing [8]. That result applies only to families of gauge-fixing conditions which may be smoothly deformed into each other, which is not true of the two gauges described above.

1. Proper time gauge $\dot{N}=0$

The proper time gauge has been extensively discussed in the literature [5,7,9,10], and we shall therefore only state some well-known results.

The condition $\dot{N}=0$ is implemented by adding a gauge-fixing term $\Pi\dot{N}$ to the Lagrangian. The Batalin-Fradkin-Vilkovisky (BFV) prescription also requires the addition of a ghost term (details may be found in [8,10]). The path integration over the ghosts factorizes, and the gauge-fixing condition, realized by the integration over the Lagrange multiplier Π, reduces the functional integration over N to a single integration, leaving

$$\mathcal{G}(x''|x')=\int dN(\tau''-\tau')\times\int\mathcal{D}p\mathcal{D}x\exp\left[i\int_{\tau'}^{\tau''}d\tau[p\dot{x}-NH]\right]. \qquad (2.12)$$

Redefining $T=N(\tau''-\tau')$, this may be rewritten as

$$\mathcal{G}(x''|x')=\int dT\, g(x'',T|x',0),$$

where $g(x'',T|x',0)$ is an ordinary quantum-mechanical transition amplitude with Hamiltonian $H=p^2-m^2$. The amplitude is given explicitly by

$$\mathcal{G}(x''|x')=\frac{1}{(2\pi)^4}\int dT\int d^4p\, e^{i[p\cdot(x''-x')-T(p^2-m^2)]}.$$

All that remains is to specify the range of T integration. If T is integrated over an infinite range, then the Hadamard Green function is obtained:

$$\mathcal{G}(x''|x')=G^{(1)}(x''|x').$$

On the other hand, if the range of integration is limited to $T\in[0,\infty)$, then, introducing a regulator to make the T integration converge, the Feynman Green function is obtained:

$$\mathcal{G}(x''|x')=iG_F(x''|x').$$

From this the sum-over-histories representations of $G^{\pm}$ are readily obtained. G^+ is obtained by taking $T>0$ and $x^{0''}>x^{0'}$, or $T<0$ and $x^{0''}<x^{0'}$, with the reverse yielding G^-. In all of these cases, the class of paths is taken to be all paths in spacetime connecting the initial and final points. Note that the causal propagator $G(x''|x')$ is not obtained by these means. We will return to this point later.

2. Canonical gauge $x^0=\tau$

It is also of interest to consider a sum over histories in which the paths are restricted to move forwards in the physical time x^0. On this class of paths, $x^0=\tau$ may be shown to be a valid gauge choice, provided that one sets up the parameter time interval so that $\tau'=x^{0'}$ and $\tau''=x^{0''}$. It may be implemented in the action by the addition of a gauge-fixing term $\Pi(x^0-\tau)$. An evaluation of the path integral, using an infinite range for N, leads to the amplitude

$$\mathcal{G}(\mathbf{x}'',x^{0''}|\mathbf{x}',x^{0'})=G_{\rm NW}(\mathbf{x}'',x^{0''}|\mathbf{x}',x^{0'})+\tilde{G}_{\rm NW}(\mathbf{x}'',x^{0''}|\mathbf{x}',x^{0'}) . \tag{2.13}$$

This includes contributions from both positive- and negative-frequency sectors of the relativistic particle, in the sense that trajectories with both positive and negative p_0 are summed over. If the integrations over N are restricted to positive N [equivalently, a factor $\theta(N)$ is included on every time slice], then only the positive-frequency sector is included. In this case the amplitude is given by the Newton-Wigner propagator

$$\mathcal{G}(x''|x')=G_{\rm NW}(x''|x') .$$

The choice of a canonical gauge leads in both cases to an amplitude which is not Lorentz invariant, a consequence of the preferred status acquired by the coordinate x^0. A comprehensive discussion of this gauge may be found in [11].

C. Canonical formulation of relativistic quantum mechanics

We have listed the various Green functions, their composition laws, and their path-integral representations, where they exist. In this subsection we discuss the connection of these considerations with the canonical quantization of the relativistic particle. In particular, we ask whether the various Green functions have canonical representations of the form $\langle x''|x'\rangle$, where the $\{|x'\rangle\}$ are complete sets of configuration-space eigenstates for any particular value of x^0, and are constructed by taking suitable superpositions of physical states (i.e., ones satisfying the constraint). We will find that essentially all of the Green functions may be so represented. Which Green function is obtained depends on the choice of inner product in the space of physical states, and on which states are included in the superposition (positive frequency, negative frequency, or both). These considerations will shed some light on various features of the composition laws.

Dirac quantization of the relativistic particle leads to a space of states which may be expressed in terms of a complete set of momentum eigenstates,

$$\hat{p}_\mu|p\rangle=p_\mu|p\rangle ,$$

subject to the additional constraint

$$(p^2-m^2)|p\rangle=0 .$$

The solutions to the constraints may be labeled by the eigenstates of the three-momentum $\mathbf{p}$, and we denote them $|\mathbf{p}\rangle$. The states $|\mathbf{p}\rangle$ are not complete, since there remains an ambiguity in the action of

$$\hat{p}_0|\mathbf{p}\rangle=\pm(\mathbf{p}^2+m^2)^{1/2}|\mathbf{p}\rangle .$$

For free particles, the positive- and negative-frequency states decouple. Canonical representations are therefore possible involving the positive- and negative-frequency sectors separately, or both together. We consider each in turn. Our aim is to find canonical representations in which each of the Green functions may be represented in the form $\langle x''|x'\rangle$.

1. Positive-frequency sector

In the positive-frequency sector, $p_0>0$, the $|\mathbf{p}\rangle$ such that

$$\hat{p}_0|\mathbf{p}\rangle=(\mathbf{p}^2+m^2)^{1/2}|\mathbf{p}\rangle$$

form a complete basis. The appropriate choice of inner product is

$$\langle\mathbf{p}|\mathbf{p}'\rangle=2\omega_{\rm p}\delta(\mathbf{p}-\mathbf{p}')$$

and the completeness relation

$$1=\int\frac{d^3\mathbf{p}}{2\omega_{\rm p}}|\mathbf{p}\rangle\langle\mathbf{p}|$$

follows, where $\omega_{\rm p}=(\mathbf{p}^2+m^2)^{1/2}$. Two choices of configuration-space representations of this Hilbert space are possible: the Newton-Wigner representation, and the relativistic representation.

Newton-Wigner representation. From the basis $|\mathbf{p}\rangle$, we may change to the Newton-Wigner basis defined by the states

$$|\mathbf{x},x^0\rangle=\frac{1}{(2\pi)^{3/2}}\int_{p_0=\omega_{\rm p}}\frac{d^3\mathbf{p}}{(2\omega_{\rm p})^{1/2}}e^{ip\cdot x}|\mathbf{p}\rangle .$$

They are not Lorentz invariant, but they are orthogonal at equal times and satisfy the completeness relation

$$1=\int d^3\mathbf{x}|\mathbf{x},x^0\rangle\langle\mathbf{x},x^0| . \tag{2.14}$$

Any Newton-Wigner wave function $\Psi(\mathbf{x},x^0)=\langle\mathbf{x},x^0|\Psi\rangle$ satisfies the positive square root of the Klein-Gordon equation:

$$i\frac{\partial}{\partial x^0}\Psi=(m^2-\nabla^2)^{1/2}\Psi ,$$

which reflects the fact that we are only considering the positive-frequency excitations. The inner product on wave functions is the usual one

$$\langle\Phi|\Psi\rangle=\int d^3\mathbf{x}\,\Phi^\dagger(\mathbf{x},x^0)\Psi(\mathbf{x},x^0) ,$$

and the propagator $\langle\mathbf{x}'',x^{0''}|\mathbf{x}',x^{0'}\rangle$ is precisely the Newton-Wigner propagator (2.8). Its composition law (2.9) follows immediately from (2.14).

Relativistic representation. It is possible to define a Lorentz-invariant configuration-space representation, using the basis states

$$|x\rangle=\frac{1}{(2\pi)^{3/2}}\int_{p_0=\omega_{\rm p}}\frac{d^3\mathbf{p}}{2\omega_{\rm p}}e^{ip\cdot x}|\mathbf{p}\rangle ,$$

where the states $|x\rangle$ with x^0 fixed form a basis on the space of physical states. They are not orthogonal, since at equal times x^0 one has

$$\langle x|x'\rangle = \frac{1}{(2\pi)^3}\int_{p_0=\omega_{\mathbf{p}}}\frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}}e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} . \tag{2.15}$$

They satisfy the relativistic completeness relation

$$1 = i\int_\Sigma d\sigma^\mu |x\rangle \overleftrightarrow{\partial}_\mu \langle x| , \tag{2.16}$$

where Σ is an arbitrary spacelike three-surface. The corresponding wave functions $\psi(x)=\langle x|\psi\rangle$ solve the positive square root of the Klein-Gordon equation and their positive-definite inner product is given by the usual relativistic expression

$$\langle\phi|\psi\rangle = i\int_\Sigma d\sigma^\mu \phi^\dagger(x)\overleftrightarrow{\partial}_\mu\psi(x) . \tag{2.17}$$

The propagator $\langle x'|x\rangle$ given in Eq. (2.15) is equal to $G^+(x'|x)$.

Similarly, by restricting attention to the negative-frequency sector, it is readily shown that $\langle x'|x\rangle$ is equal to $G^-(x'|x)$. Canonical representations of $G^\pm$ are therefore readily obtained. A canonical representation of the Feynman Green function comes from those for G^+ and G^-, although it is not immediately apparent how to construct a more direct one than this. These propagators all obey suitably modified versions of the relativistic composition law (1.8), as readily follows from the completeness relation (2.16).

2. *Positive- and negative-frequency sectors*

The discussion above provides a canonical description of both the Feynman and Newton-Wigner propagators which arose in the path-integral formulation of Sec. II B, with $N>0$. However, if the range of integration of the lapse function N is not restricted to a half-infinite range for the proper time gauge, we saw that the path integral leads to the propagator $\mathcal{G}(x''|x') = G^{(1)}(x''|x')$ where $G^{(1)}$ is the Hadamard Green function (2.6). Since restricting N to be positive (or negative) in the sum over histories appears to correspond to the positive- or negative-frequency sectors in the canonical representations, it is very plausible that a canonical representation of $G^{(1)}$ will involve both sectors simultaneously. This is indeed the case, as we now show.

In momentum space, there are two orthogonal copies of the space of states $|\mathbf{p}\rangle$. We label these two copies $|\mathbf{p},\pm\rangle$ where

$$\hat{p}_0|\mathbf{p},\pm\rangle = \pm(\mathbf{p}^2+m^2)^{1/2}|\mathbf{p},\pm\rangle .$$

The space of states is now a sum of the two copies, on which we choose the completeness relation

$$1 = \int\frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}}[\,|\mathbf{p},+\rangle\langle\mathbf{p}+| + |\mathbf{p},-\rangle\langle\mathbf{p},-|\,] . \tag{2.18}$$

The corresponding inner product is positive definite for all states:

$$\langle p,i|p',j\rangle = 2\omega_{\mathbf{p}}\delta^3(\mathbf{p}-\mathbf{p}')\delta_{ij} , \tag{2.19}$$

where $i,j=\pm$ [7].

Newton-Wigner representation. We define Newton-Wigner states as

$$|\mathbf{x},x^0\rangle = \left[\int_{p_0=\omega_{\mathbf{p}}}\frac{d^3\mathbf{p}}{(2\omega_{\mathbf{p}})^{1/2}}e^{ip\cdot x}|\mathbf{p},+\rangle + \int_{p_0=-\omega_{\mathbf{p}}}\frac{d^3\mathbf{p}}{(2\omega_{\mathbf{p}})^{1/2}}e^{ip\cdot x}|\mathbf{p},-\rangle\right] ,$$

where $p=(\omega_{\mathbf{p}},\mathbf{p})$. This definition is compatible with (2.18) and (2.19) provided that the usual completeness relation (2.14) is amended. Defining

$$|\mathbf{x},x^0,+\rangle = \int_{p_0=\omega_{\mathbf{p}}}\frac{d^3\mathbf{p}}{(2\omega_{\mathbf{p}})^{1/2}}e^{ip\cdot x}|\mathbf{p},+\rangle ,$$

$$|\mathbf{x},x^0,-\rangle = \int_{p_0=-\omega_{\mathbf{p}}}\frac{d^3\mathbf{p}}{(2\omega_{\mathbf{p}})^{1/2}}e^{ip\cdot x}|\mathbf{p},-\rangle ,$$

as the positive- and negative-frequency parts of $|\mathbf{x},x^0\rangle$, (2.14) is replaced by

$$1 = \int d^3\mathbf{x}[\,|\mathbf{x},x^0,+\rangle\langle\mathbf{x},x^0,+| + |\mathbf{x},x^0,-\rangle\langle\mathbf{x},x^0,-|\,] .$$

The propagator for Newton-Wigner states is then given by

$$\langle\mathbf{x},x^0|\mathbf{x}',x^{0'}\rangle = G_{\mathrm{NW}}(\mathbf{x},x^0|\mathbf{x}',x^{0'}) + \tilde{G}_{\mathrm{NW}}(\mathbf{x},x^0|\mathbf{x}',x^{0'}) .$$

This is precisely the amplitude derived in Sec. II B, Eq. (2.13).

Note that now the wave function $\Psi(\mathbf{x},x^0)=\langle\mathbf{x},x^0|\Psi\rangle$ solves only the second-order Klein-Gordon equation. The inner product on wave functions $\Psi(\mathbf{x},x^0)$ is

$$\langle\Phi|\Psi\rangle = \int d^3\mathbf{x}[\Phi_+^\dagger(\mathbf{x},x^0)\Psi_+(\mathbf{x},x^0) + \Phi_-^\dagger(\mathbf{x},x^0)\Psi_-(\mathbf{x},x^0)] .$$

Relativistic representation. Lorentz-invariant states in this canonical representation involving positive- and negative-frequency states may be defined by

$$|x\rangle = \int_{p_0=\omega_{\mathbf{p}}}\frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}}e^{ip\cdot x}|\mathbf{p},+\rangle + \int_{p_0=-\omega_{\mathbf{p}}}\frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}}e^{ip\cdot x}|\mathbf{p},-\rangle . \tag{2.20}$$

The usual treatment of the relativistic particle involves these states along with the usual relativistic completeness relation (2.16), which is equivalent to the usual relativistic inner product (2.17) on wave functions [12]. However, it is well known that (2.17) is not positive definite on the class of functions with both positive- and negative-frequency parts. Hence, (2.16) and (2.20) are not compatible with the positive-definite inner product (2.19). In fact, working backwards, they imply that

$$\langle\mathbf{p},\pm|\mathbf{p}',\pm\rangle = \pm 2\omega_{\mathbf{p}}\delta(\mathbf{p}-\mathbf{p}') , \tag{2.21}$$

and

$$1 = \int\frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}}[\,|\mathbf{p},+\rangle\langle\mathbf{p},+| - |\mathbf{p},-\rangle\langle\mathbf{p},-|\,] .$$

If we are to keep (2.16) and (2.20), therefore, we must use the indefinite inner product (2.21) in place of (2.19). In this way, we do in fact obtain the canonical representation for the causal Green function, for it is readily shown that one has,

$$\langle x|x'\rangle = iG(x|x') \ .$$

Its composition law follows from inserting the resolution of the identity (2.16).

A different canonical representation may be obtained by keeping (2.20) and the positive-definite inner product (2.19), but modifying the completeness relation (2.16). Define

$$|x,+\rangle = \int_{p_0=\omega_{\mathbf{p}}} \frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}} e^{ip\cdot x}|\mathbf{p},+\rangle \ ,$$

$$|x,-\rangle = \int_{p_0=-\omega_{\mathbf{p}}} \frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}} e^{ip\cdot x}|\mathbf{p},-\rangle \ ,$$

and replace (2.16) by

$$1 = i\int_\Sigma d\sigma^\mu [\,|x,+\rangle \overleftrightarrow{\partial}_\mu \langle x,+| - |x,-\rangle \overleftrightarrow{\partial}_\mu \langle x,-|\,] \ , \quad (2.22)$$

which is compatible with (2.18) and (2.19). The appropriate relativistic propagator is

$$\langle x|x'\rangle = G^{(1)}(x|x') = G^+(x|x') + G^-(x|x') \ ,$$

this giving a canonical representation of the Hadamard-Green function. The wave functions $\psi(x) = \langle x|\psi\rangle$ satisfy the Klein-Gordon equation and obey

$$\psi(x') = i\int_\Sigma d\sigma^\mu [G^+(x'|x)\overleftrightarrow{\partial}_\mu \psi_+(x) - G^-(x'|x)\overleftrightarrow{\partial}_\mu \psi_-(x)] \ ,$$

where $\psi_\pm$ are the positive- and negative-frequency parts of ψ. Since

$$\int_\Sigma d\sigma^\mu G^+ \overleftrightarrow{\partial}_\mu \psi_- = \int_\Sigma d\sigma^\mu G^- \overleftrightarrow{\partial}_\mu \psi_+ = 0 \ ,$$

it follows that

$$\psi(x') = -\int_\Sigma d\sigma^\mu G(x',x)\overleftrightarrow{\partial}_\mu \psi(x) \ ,$$

where $iG = G^+ - G^-$ is the causal Green function. This is precisely the evolution equation we expect for ψ, a solution to the Klein-Gordon equation with both positive- and negative-frequency parts. The unusual form (2.22) of the completeness relation explains how it is that the Green function $G^{(1)}$, which does not propagate solutions of the Klein-Gordon equation, is nevertheless compatible with causal evolution of a wave function $\psi(x)$. The inner product on $\psi(x)$ is now not (2.17) but rather

$$\langle \phi|\psi\rangle = i\int_\Sigma d\sigma^\mu [\phi_+^\dagger(x)\overleftrightarrow{\partial}_\mu \psi_+(x) - \phi_-^\dagger(x)\overleftrightarrow{\partial}_\mu \psi_-(x)] \ , \quad (2.23)$$

which is by construction positive definite.

We note that a significant and seemingly anomalous property of $G^{(1)}$ is that, unlike all the other Green functions, it does not obey a composition law of the usual form, but instead obeys (2.7) [13]. Our study of canonical representations now makes it clear why this is. The composition laws of G_F, $G^\pm$, and G readily follow from their canonical representations $\langle x|x'\rangle$ by simply inserting the resolution of the identity, Eq. (2.16). Recall, however, that the canonical representation of $G^{(1)}$ involves dropping (2.16) in favor of (2.22), from which follows the result

$$G^{(1)}(x''|x') = i\int_\Sigma d\sigma^\mu [G^+(x''|x)\overleftrightarrow{\partial}_\mu G^+(x|x') - G^-(x''|x)\overleftrightarrow{\partial}_\mu G^-(x|x')] \ , \quad (2.24)$$

which is readily shown to be equivalent to (2.7). The important point, therefore, is that the unusual form of the composition law for $G^{(1)}$ is explained by the nonstandard resolution of the identity in its canonical representation, which is in turn necessitated by the assumed positive-definite inner product on both positive- and negative-frequency states.

We have therefore derived canonical representations of all the Green functions. Our results, together with the composition laws and path-integral representations are summarized in Table I.

Finally, we make the following comments on the connection between the sum over histories and canonical formulations of relativistic quantum mechanics. A sum-over-histories representation of a given propagator may be derived from its canonical representation by a standard procedure, which involves inserting resolutions of the identity into the canonical expression $\langle x|x'\rangle$ (except for the causal Green function; see below). It is then reasonable to ask how one might proceed in the opposite direction, i.e., given a propagator, as supplied by the sum over histories, how does one derive the Hilbert space inner product from which the canonical representation is constructed? The answer to this question lies in the observation that the inner products given above for the relativistic representations all have the general form

$$\langle \phi|\psi\rangle = -\int_{\Sigma,\Sigma'} d\sigma^\mu d\sigma^{\mu'} \phi^\dagger(x)\overleftrightarrow{\partial}_\mu \mathcal{G}(x|x')\overleftrightarrow{\partial}_{\mu'}\psi(x') \ .$$

So, for example, by taking $\mathcal{G}$ to be G^+, one obtains the inner product (2.17). This observation is a natural starting point for the possible derivation of a canonical formulation from a sum over histories, as we shall discuss further in Sec. V.

D. Summary of Sec. II

In words, our results may be summarized as follows.

(a) The Green functions $G^\pm$ and G_F obey standard composition laws [Eqs. (2.3) and (2.4)]. They may be obtained by sums over histories over either positive or negative proper time. Their canonical representations may be obtained by restriction to the positive- or negative-frequency sector, with a positive-definite inner product and with the usual resolution of the identity.

(b) The causal Green function G obeys the standard composition law. It does not obviously have a sum-over-histories representation. Its canonical representation involves both the positive- and negative-frequency sectors, with an indefinite inner product and the usual resolution of the identity.

(c) The Hadamard Green function $G^{(1)}$ does not obey the standard composition law. It may be obtained by a

TABLE I. The various Green functions and their roles in nonrelativistic quantum mechanics. $\mathcal{G}\circ\mathcal{G}$ and $\mathcal{G}\times\mathcal{G}$ denote relativistic and nonrelativistic composition laws respectively. Unless otherwise stated, sums over histories are over arbitrary paths in spacetime from x to y.

Green function	Composition law	Sum over histories	Canonical representation $\langle x\|y\rangle$
$G^+(x,y)$	$G^+=G^+\circ G^+$; $G^+\circ G^-=0$	$N>0$ $x^0>y^0$ $N<0$ $x^0<y^0$	$p_0>0$; $\langle p\|p'\rangle>0$, $\mathbb{1}$=usual
$G^-(x,y)$	$G^-=G^-\circ G^-$; $G^+\circ G^-=0$	$N>0$ $x^0<y^0$ $N<0$ $x^0>y^0$	$p_0<0$; $\langle p\|p'\rangle>0$, $\mathbb{1}$=usual
$G_F(x,y)$	$G_F=G_F\circ G_F$	$N>0$ $x^0\gtrless y^0$	See G^+ and G^-
$G^{(1)}(x,y)$	$G=G^{(1)}\circ G^{(1)}$ $G^{(1)}=G^{(1)}\circ G$ $G^{(1)}=G^+\circ G^+-G^-\circ G^-$	$-\infty<N<\infty$	$p_0\gtrless 0$ $\langle p\|p'\rangle>0$ $\mathbb{1}$=unusual
$G(x,y)$	$G=G\circ G$	?	$p_0\gtrless 0$ $\langle p\|p'\rangle$ indefinite $\mathbb{1}$=usual
$G_{\rm NW}(x,y)$	$G_{\rm NW}=G_{\rm NW}\times G_{\rm NW}$	(i) Paths moving forward in x^0, $N>0$ (ii) All paths not crossing final surface, $N>0$ (see Sec. IV A)	$p_0>0$ $\langle p\|p'\rangle>0$ $\mathbb{1}$=usual, nonrelativistic

sum over histories over both positive and negative proper time. Its canonical representation involves both positive- and negative-frequency sectors, with a positive-definite inner product and a nonstandard resolution of the identity. The latter explains the absence of the usual composition law.

(d) The Newton-Wigner propagator obeys the composition law of the nonrelativistic type. It may be obtained by a sum over histories of the form (2.12) in which the paths move forwards in the physical time x^0. It has a canonical representation in the positive-frequency sector with a positive-definite inner product, with the usual quantum-mechanical resolution of the identity. We will find below that an alternative, rather novel representation in the proper time gauge is also available.

It is striking that unlike all the other Green functions, the causal Green function is represented canonically with an indefinite inner product. We conjecture that this is the reason why it does not have an obvious sum-over-histories representation in configuration space of the form (1.13). Briefly, a phase-space path-integral representation may be constructed by inserting resolutions of the identity into the canonical expression $\langle x|x'\rangle$, and the configuration-space path integral is obtained by integrating out the momenta. For the causal Green function, however, the indefinite inner product leads to the appearance of factors of $\varepsilon(p_0)$ in the phase-space path integral [cf. Eq. (2.5)]. This prevents the momenta from being integrated out in the usual way, and a configuration-space sum over histories of the form (1.13) is not obviously obtained.

The relativistic particle is frequently studied as a toy model for quantum gravity, and this is indeed part of the motivation for the study described in this paper. In such investigations, it is often stated that the problem with the Klein-Gordon equation is that the standard inner product is indefinite, and thus it is necessary to discard half of the solutions [1]. We would like to point out, however, that it is not necessary to view the problem in this way. As we have seen, there does in fact exist a positive-definite inner product on the set of all solutions to the Klein-Gordon equation, namely (2.23). It is therefore not necessary to discard any of the solutions if one uses this inner product. Of course, the real problem with the Klein-Gordon equation is that it is not possible to sort out the solutions into positive and negative frequency, except in the simplest of situations. This problem is present whatever view one takes.

III. THE PATH DECOMPOSITION EXPANSION

Our ultimate task is to derive the various relativistic composition laws from the sum over histories (2.12). The sum over histories for the relativistic particle readily reduces to the proper time representation (1.14). The derivation of the desired composition law is therefore intimately related to that of factoring a sum over histories of the nonrelativistic form (1.9) across an arbitrary sur-

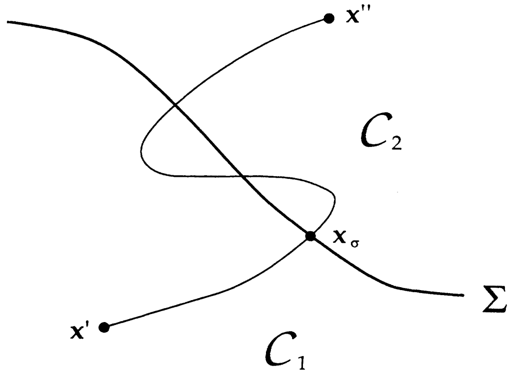

FIG. 3. The surface Σ divides the configuration space $\mathcal{C}$ into two components: $\mathcal{C}_1$ and $\mathcal{C}_2$. A path typically crosses Σ many times; the point of first crossing is at $\mathbf{x}_\sigma$.

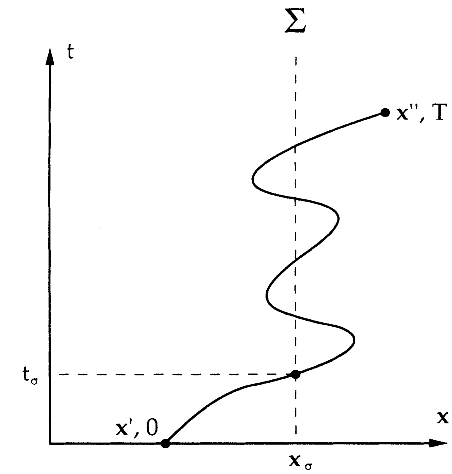

FIG. 4. The path crosses the surface Σ for the first time at $\mathbf{x}=\mathbf{x}_\sigma$ and $t=t_\sigma$ and is in the set $p(\mathbf{x}',0\rightarrow\mathbf{x}_\sigma,t_\sigma\rightarrow\mathbf{x}'',T)$.

face in configuration space. As noted in the Introduction, the solution to this problem already exists, and goes by the name of the path decomposition expansion (PDX). In this section, we will describe this result, and give a rigorous derivation of it.

A. The PDX as a partitioning of paths

Consider nonrelativistic quantum mechanics in a configuration space $\mathcal{C}$ (here taken to be $\mathbb{R}^n$), described by a propagator $g(\mathbf{x}'',T|\mathbf{x}',0)$. The propagator may be expressed as a sum over histories, which we write

$$g(\mathbf{x}'',T|\mathbf{x}',0)=\int \mathcal{D}\mathbf{x}(t)\exp\left[i\int_0^T dt\left[\frac{1}{2}M\dot{\mathbf{x}}^2-V(\mathbf{x})\right]\right]. \tag{3.1}$$

The sum is taken over all paths in configuration space, $\mathbf{x}(t)$, satisfying the boundary conditions $\mathbf{x}(0)=\mathbf{x}'$ and $\mathbf{x}(T)=\mathbf{x}''$. Denote this set of paths by $p(\mathbf{x}',0\rightarrow\mathbf{x}'',T)$.

Let Σ be a surface between $\mathbf{x}''$ and $\mathbf{x}'$. It therefore divides $\mathcal{C}$ into two parts, $\mathcal{C}_1$ and $\mathcal{C}_2$, say, with $\mathbf{x}'\in\mathcal{C}_1$ and $\mathbf{x}''\in\mathcal{C}_2$. Σ may be closed or infinite. We would like to factor the sum over histories across the surface Σ.

Consider the set of paths $p(\mathbf{x}',0\rightarrow\mathbf{x}'',T)$. Every path crosses Σ at least once, but will generally cross it many times (see Fig. 3). Unlike surfaces of constant time in spacetime, therefore, the position of crossing does not label each path in a unique and unambiguous manner. However, each path is uniquely labeled by the time and location of its *first* crossing of Σ. This means that there exists a partition of the paths according to their time t and location $\mathbf{x}_\sigma$ of first crossing (see Fig. 4). We write

$$p(\mathbf{x}',0\rightarrow\mathbf{x}'',T)=\bigcup_{\mathbf{x}_\sigma\in\Sigma}\ \bigcup_{t\in[0,T]} p(\mathbf{x}',0\rightarrow\mathbf{x}_\sigma,t\rightarrow\mathbf{x}'',T)$$

and

$$p(\mathbf{x}',0\rightarrow\mathbf{x}_\sigma,t\rightarrow\mathbf{x}'',T)\cap p(\mathbf{x}',0\rightarrow\mathbf{y}_\sigma,s\rightarrow\mathbf{x}'',T)=\varnothing \quad \text{if } \mathbf{x}_\sigma\neq y_\sigma,\ \ t\neq s\ .$$

Each path in each part $p(\mathbf{x}',0\rightarrow\mathbf{x}_\sigma,t\rightarrow\mathbf{x}'',T)$ of the partition may then be split into two pieces: (i) a restricted path lying entirely in $\mathcal{C}_1$, beginning at $\mathbf{x}'$ at time 0 and ending on Σ at $\mathbf{x}_\sigma$ at its first-crossing time t; (ii) an unrestricted path exploring $\mathcal{C}_1$ and $\mathcal{C}_2$, beginning on Σ at $\mathbf{x}_\sigma$ at time t and ending at $\mathbf{x}''$ at time T.

This suggests that there exists a composition of (3.1) across Σ, consisting of a restricted propagator in $\mathcal{C}_1$ from $(\mathbf{x}',0)$ to $(\mathbf{x}_\sigma,t)$, composed with a standard unrestricted propagator in $\mathcal{C}$ from $(\mathbf{x}_\sigma,t)$ to $(\mathbf{x}'',T)$, with summations over both $\mathbf{x}_\sigma$ and t. There is indeed such a composition law. It is the path decomposition expansion [6,14]:

$$g(\mathbf{x}'',T|\mathbf{x}',0)=\int_0^T dt\int_\Sigma d\sigma\, g(\mathbf{x}'',T|\mathbf{x}_\sigma,t)\frac{i}{2M}\mathbf{n}\cdot\nabla g^{(r)}(\mathbf{x},t|\mathbf{x}',0)\bigg|_{\mathbf{x}=\mathbf{x}_\sigma}. \tag{3.2}$$

Here, $d\sigma$ is the integration over the surface Σ. The quantity $g^{(r)}$ is the restricted propagator in $\mathcal{C}_1$, and satisfies the boundary condition that it vanish on Σ. Its normal derivative $\mathbf{n}\cdot\nabla g^{(r)}$, however, does not vanish on Σ. We give a precise definition of $g^{(r)}$ below. Also note that $\mathbf{n}$ is defined to be the normal to Σ pointing *away* from the region of restricted propagation, in this case $\mathcal{C}_1$. The reason for the appearance of the normal derivative term will become fully apparent in the rigorous derivation given below. For the moment we comment that it is related to the fact that we are interested in restricted propagation to a final point which actually lies on the boundary.

The path decomposition expansion is central to this paper, and we will be making heavy use of it in what follows.

We now record some useful closely related results. First of all, it is also possible to partition the paths according to their *last*-crossing times. This would lead to the composition law

$$g(\mathbf{x}'',T|\mathbf{x}',0)=-\int_0^T dt\int_\Sigma d\sigma\frac{i}{2M}\mathbf{n}\cdot\nabla g^{(r)}(\mathbf{x}'',T|\mathbf{x},t)\bigg|_{\mathbf{x}=\mathbf{x}_\sigma} g(\mathbf{x}_\sigma,t|\mathbf{x}',0) \ , \tag{3.3}$$

where t is the last-crossing time. The overall minus sign arises because the restricted propagator is now in the region $\mathcal{C}_2$, and the normal $\mathbf{n}$ (whose definition is unchanged) now points *into* the region of restricted propagation.

Second, it is of interest to consider the case in which the surface Σ does not lie between the initial and final points, $\mathbf{x}',\mathbf{x}''\in\mathcal{C}_1$, say. Then it is no longer true that every path crosses Σ. In this case, one first partitions the paths into paths that never cross Σ and paths that always cross. The paths that always cross may then be further partitioned as above. The sum over paths which never cross simply yields a restricted propagator in the region $\mathcal{C}_1$ that vanishes on Σ. One thus obtains

$$g(\mathbf{x}'',T|\mathbf{x}',0)=g^{(r)}(\mathbf{x}'',T|\mathbf{x}',0)+\int_0^T dt\int_\Sigma d\sigma\, g(\mathbf{x}'',T|\mathbf{x}_\sigma,t)\frac{i}{2M}\mathbf{n}\cdot\nabla g^{(r)}(\mathbf{x},t|\mathbf{x}',0)\bigg|_{\mathbf{x}=\mathbf{x}_\sigma} \ , \tag{3.4}$$

where t is the first-crossing time, and $g^{(r)}$ is the restricted propagator in $\mathcal{C}_1$. $\mathbf{n}$ is again the normal pointing away from $\mathcal{C}_1$. Similarly, in the case that the paths are partitioned according to their final-crossing times, one obtains

$$g(\mathbf{x}'',T|\mathbf{x}',0)=g^{(r)}(\mathbf{x}'',T|\mathbf{x}',0)+\int_0^T dt\int_\Sigma d\sigma\frac{i}{2M}\mathbf{n}\cdot\nabla g^{(r)}(\mathbf{x}'',T|\mathbf{x},t)\bigg|_{\mathbf{x}=\mathbf{x}_\sigma} g(\mathbf{x}_\sigma,t|\mathbf{x}',0) \ . \tag{3.5}$$

Here t is the final-crossing time, $g^{(r)}$ is again the restricted propagator in $\mathcal{C}_1$, and $\mathbf{n}$ is again the normal pointing away from $\mathcal{C}_1$. Note that there is no minus sign in the second term in Eq. (3.5), in contrast with Eq. (3.3). This is because in both (3.4) and (3.5), the region of restricted propagation is $\mathcal{C}_1$ in each case, whereas in (3.2) and (3.3), it is $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively. These subtle differences will turn out to be significant in Sec. IV.

B. A lattice derivation of the PDX

The sum over histories (3.1) must be regarded as no more than a formal expression. Certain formal properties can sometimes be deduced from (3.1) as it stands, but care is generally necessary. In particular, the path decomposition expansion cannot be derived directly from the sum over histories without recourse to a more precise mathematical definition. The purpose of this section, therefore, is to give a rigorous derivation of the PDX from a properly defined sum over histories.

Real time path integrals cannot be rigorously defined [15], so we first rotate to the imaginary time (Euclidean) version, by writing $t=-i\tau$ (note that the Euclidean time τ bears no relation to the parameter time τ of the previous section), yielding

$$g_E(\mathbf{x}'',\tau|\mathbf{x}',0) =\int\mathcal{D}\mathbf{x}(\tau')\exp\left[-\int_0^\tau d\tau'\left[\frac{1}{2}M\dot{\mathbf{x}}^2+V(\mathbf{x})\right]\right] . \tag{3.6}$$

Euclidean sums over histories may be rigorously defined as the continuum limit of a discrete sum over histories on a spacetime lattice. The discrete sum over histories is then viewed as a sum of probability measures on the space of paths on the lattice for some suitable stochastic process. To illustrate the key features of the derivation of the PDX, we will first consider the case of the free particle, $V(\mathbf{x})=0$, and define the sum over histories using one particularly simple stochastic process, namely the random walk.

Consider a spacetime lattice with temporal spacing $\Delta\tau$ and spatial spacing Δx. We follow the methods of Itzykson and Drouffe [16]. Let the n-dimensional spatial lattice be generated by n orthonormal vectors $\mathbf{e}_\mu$ with $\mathbf{e}_\mu\cdot\mathbf{e}_\nu=\Delta x^2\delta_{\mu\nu}$. Each site is located at $\mathbf{x}=x^\mu\mathbf{e}_\mu$, where the x^μ are integers.

We propose to regard $p(\mathbf{x}'',\tau|\mathbf{x}',0)$. This quantity is defined to be the probability density that in a random walk on the spacetime lattice, the system (a particle, say) will be found at $\mathbf{x}''$ at time τ given that it was at $\mathbf{x}'$ initially. On the lattice it is meaningful to talk about the probability of an individual history from $\mathbf{x}'$ at time zero to $\mathbf{x}''$ at time τ. The probability density $p(\mathbf{x}'',\tau|\mathbf{x}',0)$ is therefore given by the sum of the probabilities for the individual histories connecting the initial and final points. Formally we write,

$$p(\mathbf{x}'',\tau|\mathbf{x}',0)=\sum_{\text{histories}}p(\text{history}) \ . \tag{3.7}$$

It is in this sense that it corresponds to a sum over histories.

$p(\mathbf{x}'',\tau|\mathbf{x}',0)$ satisfies the relations

$$(\Delta x)^n p(\mathbf{x}'',0|\mathbf{x}',0)=\delta_{\mathbf{x}'',\mathbf{x}'} \ , \tag{3.8}$$

$$(\Delta x)^n\sum_{\mathbf{x}''}p(\mathbf{x}'',\tau|\mathbf{x}',0)=1 \ , \tag{3.9}$$

where $\delta_{\mathbf{x}'',\mathbf{x}'}$ denotes a product of Kronecker δ's. The factors of $(\Delta x)^n$ enter because p is a density. Equation (3.8) expresses the initial condition, and (3.9) says that the particle must be somewhere at time τ.

In a random walk, the probabilities of stepping from any one site to any one of the adjacent sites are all equal, and equal to $1/2n$ in n spatial dimensions. The probability of an entire history in (3.7) is then just $1/2n$ raised to the power of the number of steps in that history. Proceeding in this way, one may evaluate (3.7) and calculate the probability density p. However, we find it instead to be more convenient to calculate p using the recursion relation:

$$p(\mathbf{x},\tau+\Delta\tau|\mathbf{x}',0)-p(\mathbf{x},\tau|\mathbf{x}',0) = \frac{1}{2n}\sum_{\mu=1}^{n}[p(\mathbf{x}+\mathbf{e}_\mu,\tau|\mathbf{x}',0)+p(\mathbf{x}-\mathbf{e}_\mu,\tau|\mathbf{x}',0) - 2p(\mathbf{x},\tau|\mathbf{x}',0)] \ . \tag{3.10}$$

This relation follows from the fact that if the walker is on site $\mathbf{x}$ at time $\tau+\Delta\tau$, she must have been on one of the immediately adjacent sites at time τ. Equation (3.10) is a discrete version of the diffusion equation. It may be solved by Fourier transform, yielding the result

$$p(\mathbf{x}'',\tau|\mathbf{x}',0)=\int_{-\pi/\Delta x}^{\pi/\Delta x}\frac{d^n\mathbf{k}}{(2\pi)^n}e^{i\mathbf{k}\cdot(\mathbf{x}''-\mathbf{x}')} \times\left[\frac{1}{n}\sum_{\mu=1}^{n}\cos\Delta x\, k_\mu\right]^{\tau/\Delta\tau} .$$

Taking the continuum limit, $\Delta\tau,\Delta x\to 0$, and holding fixed the combination

$$\frac{(\Delta x)^2}{2n\Delta\tau}=\frac{1}{2M} , \tag{3.11}$$

one obtains

$$g_E(\mathbf{x}'',\tau|\mathbf{x}',0)=\left[\frac{M}{2\pi\tau}\right]^{n/2}\exp\left[-\frac{M(\mathbf{x}''-\mathbf{x}')^2}{2\tau}\right] ,$$

where we use g_E to denote the continuum limit of p. The diffusion limit of this stochastic process therefore yields the Euclidean propagator for the free nonrelativistic particle of mass m.

Armed with a more precise notion of a discrete sum over histories, we may now proceed to the derivation of the PDX. For simplicity, we first restrict attention to the case in which the intermediate surface Σ is flat [17]. We view $p(\mathbf{x}'',\tau|\mathbf{x}',0)$ as a sum of the probabilities for each path on the lattice from the initial to the final point. As described in Sec. III A, the paths may be partitioned according to their position $\mathbf{x}_\sigma$ and time τ_c of first crossing of an intermediate surface Σ. We therefore expect a composition law on the lattice expressing the statement "the probability of going from $\mathbf{x}'$ at time zero to $\mathbf{x}''$ at time τ is the sum over $\mathbf{x}_\sigma$ and τ_c of the probabilities of going from the initial point to final point crossing the surface Σ for the first time at time τ_c at the point $\mathbf{x}_\sigma$." The composition law is

$$p(\mathbf{x}'',\tau|\mathbf{x}',0) =(\Delta x)^n\sum_{\mathbf{x}_\sigma\in\Sigma}\sum_{\tau_c=0}^{\tau}\tilde{p}(\mathbf{x}'',\tau|\mathbf{x}_\sigma,\tau_c)q(\mathbf{x}_\sigma,\tau_c|\mathbf{x}',0) \ . \tag{3.12}$$

Here, $q(\mathbf{x}_\sigma,\tau_c|\mathbf{x}',0)$ is defined to be a lattice sum over paths which never cross Σ but end on it at position $\mathbf{x}_\sigma$ at time τ_c. After reaching the surface at the point $\mathbf{x}_\sigma$ at time τ_c, the paths must then actually step across it, by definition of the partition. The quantity $\tilde{p}(\mathbf{x}'',\tau|\mathbf{x}_\sigma,\tau_c)$ is therefore a lattice sum over all paths from the surface to the final point, but with the restriction that the very first step moves off the surface in the normal direction. It is therefore given by

$$\tilde{p}(\mathbf{x}'',\tau|\mathbf{x}_\sigma,\tau_c)=\frac{1}{2n}p(\mathbf{x}'',\tau|\mathbf{x}_\sigma+\Delta x\,\mathbf{n},\tau_c+\Delta\tau) , \tag{3.13}$$

since $1/2n$ is the probability of stepping off the surface, and $p(\mathbf{x}'',\tau|\mathbf{x}_\sigma+\Delta x\,\mathbf{n},\tau_c+\Delta\tau)$ is the probability of going from the point just off the surface to the final point. Strictly the sum over τ_c should not begin at zero, because on the lattice it takes a finite amount of time for the first path to reach the surface, but this time interval goes to zero in the continuum limit.

Because $q(\mathbf{x}'',\tau|\mathbf{x}',0)$ is a sum over paths that never cross Σ (but may touch it), it will satisfy the boundary condition

$$q(\mathbf{x}_\sigma+\Delta x\,\mathbf{n},\tau_c|\mathbf{x}',0)=0 \tag{3.14}$$

where $\mathbf{n}$ is the normal to the surface. That is, the probability of making one step beyond Σ is zero. Now write

$$q(\mathbf{x}_\sigma,\tau_c|\mathbf{x}',0) =q(\mathbf{x}_\sigma+\Delta x\,\mathbf{n},\tau_c|\mathbf{x}',0) -\Delta x\left[\frac{q(\mathbf{x}_\sigma+\Delta x\,\mathbf{n},\tau_c|\mathbf{x}',0)-q(\mathbf{x}_\sigma,\tau_c|\mathbf{x}',0)}{\Delta x}\right] . \tag{3.15}$$

The boundary condition (3.14) implies that the first term vanishes. The part of the second term in square brackets converges to the normal derivative of q in the continuum limit. Inserting this in (3.12), one obtains, with some rearrangement,

$$p(\mathbf{x}'',\tau|\mathbf{x}',0)=\sum_{\mathbf{x}_\sigma\in\Sigma}(\Delta x)^{n-1}\sum_{\tau_c=0}^{\tau}\Delta\tau\, p(\mathbf{x}'',\tau|\mathbf{x}_\sigma+\Delta x\,\mathbf{n},\tau_c+\Delta\tau)\frac{(\Delta x)^2}{2n\Delta\tau}\left[\frac{q(\mathbf{x}_\sigma+\Delta x\,\mathbf{n},\tau_c|\mathbf{x}',0)-q(\mathbf{x}_\sigma,\tau_c|\mathbf{x}',0)}{\Delta x}\right] . \tag{3.16}$$

Now, using the continuum limits

$$\sum_{\mathbf{x}_\sigma\in\Sigma}(\Delta x)^{n-1}\to\int_\Sigma d\sigma, \quad \sum_{\tau_c=0}^{\tau}\Delta\tau\to\int_0^\tau d\tau_c \tag{3.17}$$

and using (3.11), we derive

$$g_E(\mathbf{x}'',\tau|\mathbf{x}',0)=\int_0^\tau d\tau_c \int_\Sigma d\sigma\, g_E(\mathbf{x}'',\tau|\mathbf{x}_\sigma,\tau_c)\frac{1}{2M}\mathbf{n}\cdot\nabla g_E^{(r)}(\mathbf{x},\tau_c|\mathbf{x}',0)\bigg|_{\mathbf{x}=\mathbf{x}_\sigma} . \tag{3.18}$$

This is the Euclidean version of the path decomposition expansion. The desired result (3.2) is then readily obtained by continuing back to real time. The closely related results (3.3)–(3.5) are derived in a similar manner.

It is perhaps worth noting that this result cannot be derived from formal manipulation of the continuum sum over histories (3.1). Each part of the composition law (3.12) is well defined and nonzero on the lattice, but not every part has a continuum analogue. In particular $q(\mathbf{x}_\sigma,\tau_c|\mathbf{x},0)$, where $\mathbf{x}_\sigma$ is on Σ, formally goes to zero in the continuum limit. The desired result arises because the various parts of (3.12) fortuitously conspire to give a result which is well defined in the continuum limit, even though the separate parts may not be.

Now consider the case of a nonzero potential, $V(\mathbf{x})\neq 0$. We will argue that the inclusion of a potential does not affect the key points of the derivation of the path decomposition expansion. The random walk process described above supplies a *measure* on the set of paths on the lattice. (In fact it is an important result that it also defines a measure in the continuum limit, but we prefer to work on the lattice.) Using this measure, one can compute the average value of various functions of the histories of the system. In particular, it is a standard result that the amplitude (3.6) may be defined as the average value of $\exp[-\int d\tau\, V(\mathbf{x}(\tau))]$ in this measure [18].

A different way of doing essentially the same calculation is more convenient for our purposes. The amplitude (3.6) may be calculated directly by constructing a measure on the set of paths different to that given above, which includes the effect of the potential. A weight w(history) may be defined for each history, and the density $w(\mathbf{x}'',\tau|\mathbf{x}',0)$ is again

$$w(\mathbf{x}'',\tau|\mathbf{x}',0)=\sum_{\text{histories}} w(\text{history}) .$$

Loosely speaking, w is defined by weighting the probability p of going from one lattice point to the next by $\exp[-\Delta\tau\, V(\mathbf{x})]$. w is of course no longer a probability density, and does not define a stochastic process. It obeys the recursion relation

$$w(\mathbf{x},\tau+\Delta\tau|\mathbf{x}',0)-w(\mathbf{x},\tau|\mathbf{x}',0)=\frac{1}{2n}\sum_{\mu=1}^{n}[w(\mathbf{x}+\mathbf{e}_\mu,\tau|\mathbf{x}',0)+w(\mathbf{x}-\mathbf{e}_\mu,\tau|\mathbf{x}',0)-2w(\mathbf{x},\tau|\mathbf{x}',0)]+\Delta\tau\, V(\mathbf{x})w(\mathbf{x},\tau|\mathbf{x}',0) , \tag{3.19}$$

which differs from (3.10) in that the "walker" may now stay at site $\mathbf{x}$ with a weight $\Delta\tau\, V(\mathbf{x})$. This recursion relation yields the Euclidean Schrödinger equation with potential $V(\mathbf{x})$ in the continuum limit, as expected.

The issue is now to determine whether the derivation (3.12)–(3.18) goes through for w as it did for p. It is relatively easy to see that it will. The quantities analogous to $\tilde{p}$ and q are defined in the obvious way, and all the steps go through as before. The important point is that (3.12) and (3.13) are not modified, since the weight for stepping off the surface is still $1/2n$, as may be seen from the recursion relation (3.19).

An equivalent approach is to rescale the weights w so that they describe a stochastic process, and can be regarded as probability densities [19]. The random walk is then characterized by a nonzero drift, that is by unequal probabilities of stepping in different directions due to the asymmetry of the potential. In the continuum limit, the rescaled w satisfies a Fokker-Planck equation. A composition law involving an object analogous to q may then be derived, which is a rescaled version of the path decomposition expansion. We will not pursue this here [20].

C. An important simplification

The restricted propagator appearing in (3.2)–(3.5) is somewhat inconvenient and for our purposes it is useful to reexpress it in terms of the usual propagator [21]. This is certainly possible if the potential $V(\mathbf{x})$ in (3.1) possesses a translational symmetry in a direction that we shall refer to as x^0, and Σ is a surface of constant x^0. To this end, consider $g^{(r)}(\mathbf{x},t|\mathbf{x}',0)$ in (3.2), where both $\mathbf{x}$ and $\mathbf{x}'$ are in $\mathcal{C}_1$, the region of restricted propagation. By the imposed symmetry of the random walk, it is possible to rewrite the restricted propagator as

$$g^{(r)}(\mathbf{x}'',t|\mathbf{x}',0)=g(\mathbf{x}'',t|\mathbf{x}',0)-g(\mathbf{x}_\sigma+(\mathbf{x}_\sigma-\mathbf{x}''),t|\mathbf{x}',0) , \tag{3.20}$$

where $\mathbf{x}_\sigma$ is the point on Σ closest to $\mathbf{x}''$. This is of course just the familiar method of images. That this is equivalent to a restricted sum over paths may be seen as follows. The full propagator is given by a sum over all paths from initial to final point. The sum over all paths g may be written as a sum over paths that never cross the surface, $g^{(r)}$, plus a sum over paths that do cross the surface at least once, $g^{(a)}$ [cf. Eq. (3.4)]. The paths that cross have a last-crossing position. Because of the symmetry, the segment of the path after the last crossing may be reflected about the surface without changing the value of the sum over paths (see Fig. 5). $g^{(a)}$ is therefore equal to the sum over all paths from the initial point to the reflection about the surface of the final point. Hence,

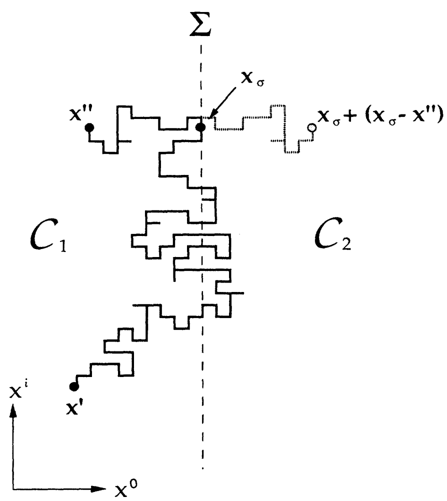

FIG. 5. A path crossing the surface Σ and ending at $\mathbf{x}''$ is canceled by a path crossing the surface and ending at $\mathbf{x}_\sigma+(\mathbf{x}_\sigma-\mathbf{x}'')$, provided that $V(\mathbf{x})$ is independent of x^0.

with a little rearrangement, one obtains (3.20).

Given (3.20), the normal derivative of the restricted propagator on Σ, which is the quantity that appears in (3.2), is just

$$\mathbf{n}\cdot\nabla g^{(r)}(\mathbf{x},t|\mathbf{x}',0)|_{\mathbf{x}=\mathbf{x}_\sigma}=2\mathbf{n}\cdot\nabla g(\mathbf{x},t|\mathbf{x}',0)|_{\mathbf{x}=\mathbf{x}_\sigma} . \tag{3.21}$$

We conclude that in the special case of a symmetric potential and a flat surface, (3.2) becomes

$$g(\mathbf{x}'',T|\mathbf{x}',0)=\int_0^T dt\int_\Sigma d\sigma\, g(\mathbf{x}'',T|\mathbf{x}_\sigma,t)\times\frac{i}{M}\mathbf{n}\cdot\nabla g(\mathbf{x},t|\mathbf{x}',0)|_{\mathbf{x}=\mathbf{x}_\sigma} \tag{3.22}$$

and likewise for (3.3). We will use this result in all subsequent applications of the PDX. The analysis so far is for flat surfaces Σ. Equation (3.21) will also follow for curved surfaces (in flat configuration spaces with constant potential), because the analysis leading to it is essentially local.

IV. DERIVATION OF RELATIVISTIC COMPOSITION LAWS

We now show how the path decomposition expansion is used to derive the relativistic composition laws for certain Green functions.

A. Composition laws for G_F, G^+, and G^-

Consider the Feynman Green function. As discussed earlier, its sum-over-histories representation readily reduces to

$$iG_F(x''|x')=\int_0^\infty dT\, g(x'',T|x',0) . \tag{4.1}$$

Here $g(x'',T|x',0)$ is a propagator of the nonrelativistic type, and is given by a sum over histories of the form (3.1), but with $2M=1$, $V(x)=m^2$ (which means that the results of Sec. III C apply for any flat surface), and with $\dot{\mathbf{x}}^2$ replaced by $\dot{x}^\mu\dot{x}^\nu\eta_{\mu\nu}$, where $\eta_{\mu\nu}$ is the Minkowski metric, with signature $(+---)$. It therefore obeys the Schrödinger equation

$$\left[-i\frac{\partial}{\partial T}+\Box_{x''}+m^2\right]g(x'',T|x',0)=0 ,$$

subject to the initial condition

$$g(x'',0|x',0)=\delta^{(4)}(x''-x') .$$

From this, it readily follows that $G_F(x''|x')$ satisfies Eq. (2.1). An explicit expression for $g(x'',T|x',0)$ is readily obtained:

$$g(x'',T|x',0)=\frac{1}{(2\pi iT)^2}\exp\left[-i\frac{(x''-x')^2}{4T}-im^2T\right] .$$

These basics out of the way, we may now derive the composition law. Consider first the case in which the initial and final points are on opposite sides of the surface Σ. Apply the path decomposition expansion (3.2) to (4.1). One obtains

$$G_F(x''|x')=-i\int_0^\infty dT\int_0^T dt\int_\Sigma d\sigma\, g(x'',T|x,t)2i\vec{\partial}_n g(x,t|x',0) . \tag{4.2}$$

Here, $\vec{\partial}_n$ denotes the normal derivative pointing away from x' and operating to the right, and we have used (3.21) to express the derivative of the restricted propagator in terms of the unrestricted propagator. Also, we use a simple x to denote the coordinates in the surface Σ. Now, in the integrals over time, one may perform the change of coordinates $v=T-t$, and $u=t$. Equation (4.2) then becomes

$$G_F(x''|x')=2\int_0^\infty dv\int_0^\infty du\int_\Sigma d\sigma\, g(x'',v|x,0)\vec{\partial}_n g(x,u|x',0) . \tag{4.3}$$

Comparing with (4.1), it is then readily seen that

$$G_F(x''|x')=-2\int_\Sigma d\sigma\, G_F(x''|x)\vec{\partial}_n G_F(x|x') . \tag{4.4}$$

Although this is a correct property of the Feynman-Green function, it is not quite the expected result. Furthermore, it does not manifestly exhibit the usual property of independence of the location of the factoring surface. To this end, we repeat the above with (3.3) instead of (3.2), obtaining,

$$G_F(x''|x')=2\int_\Sigma d\sigma\, G_F(x''|x)\overleftrightarrow{\partial}_n G_F(x|x') . \tag{4.5}$$

Finally, averaging (4.4) and (4.5) leads to the desired result

$$G_F(x''|x') = -\int_\Sigma d\sigma\, G_F(x''|x)\overleftrightarrow{\partial}_n G_F(x|x') \ . \qquad (4.6)$$

Define $iG_F(x''|x')$ to be $G^+(x''|x')$ when x'' is in the future cone of x', and to be $G^-(x''|x')$ when x'' is in the past cone of x'. Then it readily follows that G^+ and G^- each satisfy suitably modified versions of (4.4) and (4.5) and hence their composition laws (2.3).

Now consider the case in which the initial and final points lie on the same side of the surface Σ. We therefore apply the path decomposition expansions (3.4) and (3.5). Direct application of either of these expressions to Eq. (4.1) does not lead to an obviously useful result, since it still involves a restricted propagator. However, equating (3.4) and (3.5), one obtains

$$\int_0^T dt \int_\Sigma d\sigma\, g(x'',T|x,t)\overrightarrow{\partial}_n g(x,t|x',0) = \int_0^T dt \int_\Sigma d\sigma\, g(x'',T|x,t)\overleftarrow{\partial}_n g(x,t|x',0) \ . \qquad (4.7)$$

Suppose that x'' is in the future cone of x', which in turn lies to the future of the surface Σ. Then performing the integration over T in (4.7) leads to the result

$$\int_\Sigma d\sigma^\mu G^+(x''|x)\overleftrightarrow{\partial}_\mu G^-(x|x') = 0 \ , \qquad (4.8)$$

demonstrating the expected orthogonality of G^+ and G^-.

It might appear that the above derivation of the composition law is valid for *any* choice of factoring surface. This impression would be false: the derivation holds only for spacelike surfaces. To see this, note that the integral representation of the Feynman Green function (4.1) is properly defined only in the Euclidean regime. The Euclidean version of (4.1) is obtained by rotating both the parameter time T and the physical time x^0. Write $T_E = iT$ and $x_E^0 = ix^0$. The first rotation is just a matter of distorting the integration contour in (4.1) and does not change the result of evaluating the integral. Indeed, (4.1) may be *defined* by an integral over real T_E. The second rotation actually changes the answer, so needs to be rotated back afterwards. Performing the rotations, one obtains

$$g_E(x'',T_E|x',0) = \frac{1}{(2\pi T_E)^2} \exp\left[-\frac{1}{4T_E}[(x_E^0)^2 + \mathbf{x}^2] - m^2 T \right] \qquad (4.9)$$

for the time-dependent propagator, where x denotes $x''-x'$ for both its time and space components. The Euclidean path decomposition expansion (3.18) is then clearly well defined for (4.9)—the integral over the surface Σ is clearly convergent. A composition law for the Euclidean Feynman propagator is therefore obtained across *any* surface. But suppose now we try to continue back the Euclidean PDX (3.18) to the Lorentzian spacetime. Leave T_E as it is, but continue back x^0. The integrand, previously exponentially decaying in all directions, becomes exponentially growing in the x^0 direction. This is not a problem if the surface Σ is spacelike, since x^0 is not integrated over. It is a problem if x^0 is integrated over, which it would be if Σ is timelike. It follows that the Euclidean composition law, valid for any surface, may be continued to a well-defined composition law for the Lorentzian propagator only if the surface Σ is spacelike in the Lorentzian regime.

At this stage it is perhaps useful to summarize how we have arrived at the results (4.6) and (4.8) from the sum over histories. First the sum over histories was written in the proper time representation (4.1). This is essentially a partition of the set of all paths from x' to x'', according to their total parameter time (which is effectively the same as their length). Then the paths were further partitioned according to the parameter time and position of their first (or last) crossing Σ. The path decomposition expansion then led to the desired result. In the final result (4.6), however, no reference is made to the parameter time involved in this sequence of partitions; only the first- (or last-)crossing position x is referred to. In the results (4.6) and (4.8), therefore, there is only one partition of the paths that is important, namely the partition according to the position x of first or last crossing. Differently put, suppose there existed a sum-over-histories representation of G_F referring only to the spacetime coordinates x^μ, and not requiring the explicit introduction of a parameter t. Then the composition law (4.6) could be derived by a single partitioning of the paths according to their first- or last-crossing position.

By way of a short digression, let us explore this idea further. Suppose one simply assumes that a sum-over-histories representation of $G_F(x''|x')$ is available, in which there is a sum over all paths in spacetime from x' to x''. As described above, one can therefore partition the paths according to their first-crossing position x of an intermediate surface Σ. It is therefore reasonable to *postulate* a relation of the form

$$G_F(x''|x') = \int_\Sigma d\sigma\, G_F(x''|x)\Delta(x|x') \ ,$$

where $\Delta(x|x')$ is defined by a restricted sum over paths beginning at x' which never cross Σ, but end on it at x. Comparing with Eq. (4.4), or by explicit calculation, one has

$$\Delta(x|x') = 2\partial_n G_F(x|x') \ . \qquad (4.10)$$

This gives a rather intriguing representation of $\partial_n G_F(x|x')$ in terms of a restricted sum over paths in spacetime.

It is also interesting to note that when $x^0 > x^{0'}$, by explicit calculation,

$$2i\partial_n G_F(x|x') = G_{NW}(x|x') \ ,$$

where $G_{\rm NW}$ is the Newton-Wigner propagator [22]. Via (4.10), this therefore gives a novel path-integral representation of the Newton-Wigner propagator. It is novel because $G_{\rm NW}$ is really a propagator of the Schrödinger type, and is therefore normally obtained by a sum over paths moving forward in time, as we saw in Sec. II B. In contrast, in the path-integral representation of $\Delta(x|x')$, the paths move backward and forward in time, although are restricted to lie on one side of the surface Σ in which x lies.

Note that using this representation of the Newton-Wigner propagator, its composition law (2.9) is easily derived. The sum over paths from x' to x'' ending on Σ and remaining below it, may be partitioned across an intermediate surface Σ' according to the point $x_{\sigma'}$ of first crossing of Σ'. That is

$$p(x'\to x'')=\bigcup_{x_{\sigma'}} p(x'\to x_{\sigma'}\to x'') ,$$

$$p(x'\to x_{\sigma'}\to x'')\cap p(x'\to y_{\sigma'}\to x'')=\varnothing \quad \text{if } x_{\sigma'}\neq y_{\sigma'} .$$

The sum over paths factorizes into a sum over paths from x' to $x_{\sigma'}$, ending on Σ' and remaining below it, and over paths from $x_{\sigma'}$ to x'', ending on Σ and remaining below it. This is precisely a composition of type (1.12), and leads directly to (2.9).

These observations may merit further investigation. They are, however, only incidental to the rest of this paper.

B. Other Green functions

By integrating T over an infinite range in (1.14) the Green function $G^{(1)}(x''|x')$ is obtained. Let us therefore repeat the steps (4.2) to (4.6) for this case. The integration over T and t in (4.2) is now

$$\int_{-\infty}^{\infty} dT \int_0^T dt=\int_0^{\infty} dT \int_0^T dt+\int_{-\infty}^0 dT \int_0^T dt . \qquad (4.11)$$

The first term in (4.11) leads to a composition of two Feynman Green functions, as before. The second term can be cast in a similar form by letting $T\to -T$ and $t\to -t$, which introduces an overall minus sign, and using the fact that $g(x,-t|x',0)=g^*(x,t|x',0)$. One thus obtains

$$\begin{aligned} G^{(1)}(x''|x')&=-i\int_\Sigma d\sigma[G_F(x''|x)\overleftrightarrow{\partial}_n G_F(x|x') \\ &\qquad -G_F^*(x''|x)\overleftrightarrow{\partial}_n G_F^*(x|x')] \\ &=i\int_\Sigma d\sigma^\mu[G^+(x''|x)\overleftrightarrow{\partial}_\mu G^+(x|x') \\ &\qquad -G^-(x''|x)\overleftrightarrow{\partial}_\mu G^-(x|x')] , \end{aligned} \qquad (4.12)$$

where $d\sigma^\mu$ and ∂_n are defined as in Sec. II. The result (4.12) may seem somewhat trivial, since it follows from (4.6) and the use of

$$\begin{aligned} G^{(1)}(x''|x')&=i[G_F(x''|x')-G_F^*(x''|x')] \\ &=G^+(x''|x')+G^-(x''|x') . \end{aligned}$$

However, the key point is that the composition law (2.24) for $G^{(1)}$ arises directly in the sum over histories. The splitting into positive- and negative-frequency parts, in the language of Sec. II, arises naturally from the identity (4.11).

Finally, consider the causal Green function. In terms of $G^\pm$ it is defined by

$$iG(x''|x')=G^+(x''|x')-G^-(x''|x') .$$

Then it straightforwardly follows that G obeys (1.8), since $G^\pm$ obey (2.3) and (4.8). However, there is no natural, quantum-mechanical derivation of the composition law for G directly from a sum over histories. This is because we do not have a direct path-integral representation of G, only an indirect one in terms of the path-integral representations of $G^\pm$ [which may be read off from (4.1)]. The question of finding a direct sum-over-histories representation of the causal propagator is, to the best of our knowledge, a question for which no entirely satisfactory answer exists at present. Indeed, as we conjectured in Sec. II, such a representation may not exist.

C. Why the naive composition law fails

In the context of quantum gravity, and parametrized theories generally, composition laws different in form to (1.8) have occasionally been proposed. In particular, a composition law of the form

$$\mathcal{G}(x''|x')=\int d^4x\, \mathcal{G}(x''|x)\mathcal{G}(x|x') \qquad (4.13)$$

has often been considered [23]. However, it is readily seen that there are difficulties associated with (4.13) [24]. The methods of this paper help to understand the reason why it cannot hold as it stands.

Let us first illustrate the problem with (4.13). Consider the proper time representation (1.14). It is a property of g that

$$g(x'',T''+T'|x',0)=\int d^4x\, g(x'',T''|x,0)g(x,T'|x',0) .$$

Integrating both sides over T'' and T', one obtains

$$-\tfrac{1}{2}\int du \int dv\, g(x'',u|x',0)=\int d^4x\, g(x''|x)g(x|x') , \qquad (4.14)$$

where we have introduced $u=T''+T'$, $v=T''-T'$. If T is taken to have an infinite range, then u and v have an infinite range, and the left-hand side of (4.14) is equal to $G^{(1)}(x''|x')$ multiplied by an infinite factor. If T is taken to have a half-infinite range, then things are yet more problematic. In that case v ranges from $-u$ to $+u$, and the left-hand side of (4.14) becomes

$$-\int_0^\infty du\, u\, g(x'',u|x',0) . \qquad (4.15)$$

This may converge, but it does not converge to the left-hand side of (4.13).

It should be clear from the discussion given in the introduction that (4.13) should not be expected to hold. The reason is, quite simply, that it does not correspond to a proper partitioning of the paths in the sum over histories (1.13). For in proposing an expression of the form (4.13), one is evidently contemplating a partitioning of the paths $(x'\to x'')$ in which the paths are labeled according to an intermediate spacetime point x through which they pass. That is, the set of all paths is regarded as the union over all x of paths passing through x:

$$p(x'\to x'')=\bigcup_x p(x'\to x\to x'') .$$

But this is not a proper partition because it is not exclusive:

$$p(x'\to x\to x'')\cap p(x'\to y\to x'')\neq\varnothing \quad \text{for } x\neq y \ . \tag{4.16}$$

It is not exclusive because passing through an intermediate point x does not prohibit the path from also passing through a different intermediate point y. The intermediate spacetime point x therefore does not supply the paths with a unique and unambiguous label.

Of course, the exhaustivity condition (4.16) is still in some sense true, but the failure of the exclusivity condition means that there is a vast amount of overcounting. It is this that leads to the divergent factor appearing in (4.15) in the case where T takes an infinite range.

The fact that (4.14) is equal to $G^{(1)}$ times an infinite factor is, however, suggestive. A similar feature was found in the Dirac quantization of the relativistic particle by Henneaux and Teitelboim [7]. They found that for functions $\psi(x)=\langle x|\psi\rangle$ solving the Klein-Gordon equation

$$\langle\phi|\psi\rangle=\int d^4x\, \phi^\dagger(x)\psi(x) \tag{4.17}$$

is a positive-definite inner product independent of x^0, and with all the necessary symmetry properties for an inner product on physical states. The only problem with (4.17) is that it is formally divergent. In fact, it is equal to the inner product (2.23) times a factor $\delta(0)$, which may be removed in a Lorentz-invariant way [7,25]. This inner product may therefore be of some value, despite the fact that it is not associated with a partition of the sum over histories. It is yet to be seen whether these features continue to hold in more complicated parametrized systems, such as quantum gravity.

It is also clear that a composition law in which the d^4x in (4.13) is replaced by a d^3x cannot be correct. This would at first sight be more in keeping with conventional quantum mechanics, since one of the four x^μ's is time, and the composition law (1.4) is at a fixed moment of time. However, it corresponds to contemplating a partition in which the paths are labeled according to the position x^i at which they cross a surface $x^0=$const. This fails because, as discussed in the Introduction, it is not a proper partition. The paths typically cross such a surface many times, and the crossing location does not label the paths in a unique and unambiguous way.

We have seen in this paper that there is a partition that does work, and does lead to the desired composition law. It is to partition the paths according to their position of *first* crossing of an intermediate surface.

V. DISCUSSION

The principal technical aim of this paper was to show that the composition laws of relativistic quantum mechanics may be derived directly from a sum over histories by partitioning the paths according to their first-crossing position of an intermediate surface. We also derived canonical representations of the propagators. These representations showed why the Hadamard Green function $G^{(1)}$, which is the propagator picked out by the sum over histories, does not obey a standard composition law. They also indicate why it is not obviously possible to construct a sum-over-histories representation of the causal Green function.

The notion of a sum over histories is extremely general. Indeed, as discussed in the Introduction, it has been suggested that sum-over-histories formulations of quantum theory are more general than canonical formulations. Central to such generalized formulations of quantum mechanics is the notion of a partition to paths. This simple but powerful notion replaces and generalizes the notion of a complete set of states at a fixed moment of time used in canonical formulations [3].

In this paper we have investigated a particular aspect of the correspondence between these two different approaches to quantum theory. Namely, we demonstrated the emergence of the composition law from the sum-over-histories approach, in the context of relativistic quantum mechanics in Minkowski space. Quite generally, such a derivation will be an important step in the route from a sum-over-histories formulation to a canonical formulation in a reparametrization invariant theory. We have admittedly not determined the exact status of the composition law along this route. In particular, it is not clear whether the existence of the composition law alone is a *sufficient* condition for the recovery of a canonical formulation. This would be an interesting question to pursue, perhaps taking as a starting point the comments at the end of Sec. II C, on the recovery of the canonical inner product given the propagator. However, as argued in the Introduction, it is at least a *necessary* condition. It is therefore of interest to find a situation in which this necessary condition is not satisfied.

Such a situation is provided by the case of relativistic quantum mechanics in curved spacetime backgrounds with a spacetime-dependent mass term (i.e., a potential). Let us consider the generalization of our results to this case.

The path decomposition expansion (3.2) is a purely kinematical result. As we have shown, it arises solely from partitioning the paths in the sum over histories, and does not depend on the detailed dynamics. We would therefore expect it to hold in a very general class of configuration spaces, including curved ones. It follows that for the relativistic particle, one would *always* expect a composition law of the form

$$\mathcal{G}(x''|x')=\int d\sigma^\mu \mathcal{G}(x''|x)\partial_\mu \mathcal{G}^{(r)}(x|x') \ , \tag{5.1}$$

where $\mathcal{G}^{(r)}$ is the restricted relativistic propagator [26]. In the case of flat backgrounds, with constant potential, it was possible to express the restricted Green functions in terms of unrestricted ones, using (3.20) and (3.21). The important point, however, is that in general backgrounds, and with arbitrary potentials, the steps (3.20) and (3.21) are not possible, and a composition law of the desired type (1.4) is not recovered. Of course, (5.1) is still a composition law of sorts, but $\mathcal{G}$ and $\mathcal{G}^{(r)}$ are quite different types of object, and (5.1) is not compatible with regarding $\mathcal{G}(x''|x')$ as a canonical expression of the form $\langle x''|x'\rangle$, since there is no known canonical representation for $\mathcal{G}^{(r)}(x''|x')$.

What is needed for steps (3.20) and (3.21) to work? The main issue is understanding how the method of images can be generalized. First of all, consider the case of one dimension with a potential. The method of images yields the restricted propagator for any potential which is symmetric about the factoring surface (actually a point in one dimension). For example, the restricted propagator in $x>0$ for the harmonic oscillator is readily obtained in this way. However, we would like to obtain the restricted propagator on one side of *any* factoring surface. The only potential invariant under reflections about any point is a constant. So in one dimension, (3.21) follows if the potential is constant. Similarly, it is easy to see that in flat spaces of arbitrary dimension, with a flat factoring surface, (3.21) will follow if the potential is constant in the direction normal to the surface.

Now consider the case of curved spacetimes with a Lorentzian signature (although our conclusions will not be restricted to this situation). From the above, we have seen that the method of images will work if the propagator is symmetric about each member of a family of factoring surfaces. We will now argue that this will be true if there is a timelike Killing vector.

Consider first the case of static spacetimes. This means there is a timelike Killing vector field normal to a family of spacelike hypersurfaces. It is therefore possible to introduce coordinates such that

$$g_{\mu\nu}dx^{\mu}dx^{\nu}=g_{00}(x^{i})(dx^{0})^{2}+g_{kl}(x^{i})dx^{k}dx^{l} , \qquad (5.2)$$

where $i,k,l=1,2,3$. The action in the sum-over-histories representation of $g(x'',T|x',0)$ is

$$S=\int_{0}^{T}dt[g_{\mu\nu}\dot{x}^{\mu}\dot{x}^{\nu}-V(x^{\mu})] . \qquad (5.3)$$

If the metric is of the form (5.2) and if the potential is independent of x^0, then the action (5.3) will be invariant under reflections about any surface of constant x^0. It is reasonable to expect that the path-integral measure will be similarly invariant, and hence the method of images may be used to construct the restricted propagator in a region bounded by $x^0=$const. We therefore expect (3.21) to hold in static spacetimes in which the potential is invariant along the flow of the Killing field. We anticipate that this argument may be generalized to stationary spacetimes (for which there is a Killing field that is not hypersurface orthogonal), but we have not proved this.

What we find, therefore, is that the existence of a timelike Killing vector field, along which the potential is constant, is a *sufficient condition* for the existence of a composition law for the sum over histories. We cannot conclude from the above argument that it is also a necessary condition, although this is plausibly true for a general class of configuration spaces, with the possible exception of a limited number of cases in which special properties of the space avoid the need for a Killing vector [27]. Modulo these possible exceptions, we have therefore achieved our desired aim: we have found a situation, spacetimes with no Killing vectors, in which the necessary condition for the recovery of a canonical formulation from a sum over histories is generally not satisfied.

This is a desirable conclusion: the existence of a timelike Killing vector field is the sufficient condition for a consistent one-particle quantization in the canonical theory (see Sec. 9 of Ref. [1] and references therein). Again, it is not obviously a necessary condition because there could be spacetimes with no Killing vectors but some special properties permitting quantization in them. We therefore find close agreement (although not an exact correspondence) between our approach, in which the canonical formulation is regarded as derived from a sum over histories, and standard lore, in which it is constructed directly.

Turn now to quantum cosmology. As noted in the Introduction, relativistic quantum mechanics is frequently used as a model for quantum cosmology. In quantum cosmology, the wave function for the system, the Universe, obeys the Wheeler-DeWitt equation. This is a functional differential equation which has the form of a Klein-Gordon equation in which the four x^{α}'s are replaced by the three-metric field $h_{ij}(\mathbf{x})$, the "mass" term is dependent on the three-metric, and the "background" (superspace, the space of three-metrics) is curved.

As outlined in the Introduction, one may construct the propagator between three-metrics. The object obtained is most closely analogous to either the Feynman or the Hadamard propagators, as noted above. One can then ask whether it obeys a composition law. An important result due to Kuchař [1] is that there are no Killing vectors associated with the Wheeler-DeWitt equation. We therefore find that there is no composition law for the propagator between three-metrics generated by a sum over histories [28]. It follows that we do not expect to recover a canonical formulation. Again this is in agreement with standard lore on the canonical quantization of quantum cosmology, which holds that there is no consistent "one-universe" quantization [1].

Our final conclusions on the existence of a canonical scheme for quantum cosmology are therefore not new. However, what has not been previously appreciated, as far as we are aware, is the close connection of this question with the question of the existence of a composition law for the sum over histories.

Finally, we may comment on the suggestion of Hartle discussed in the Introduction—that the sum over histories is more general than the canonical scheme. Our results are not inconsistent with this claim: the absence of Killing vectors associated with the Wheeler-DeWitt equation probably rules out a canonical quantization, but does not obviously prevent the construction of sums over histories. Of course, there still remains the question of how the sums over histories are to be used to construct probabilities, i.e., the question of *interpretation.* This is a difficult question and will not be addressed here.

We emphasize that these arguments are intended to be suggestive, rather than rigorous. These issues will be considered in greater detail in future publications.

ACKNOWLEDGMENTS

We would like to thank Arlen Anderson, Franz Embacher, Eddie Farhi, Jeffrey Goldstone, Sam Gutmann, Jim Hartle, Chris Isham, Samir Mathur, and Claudio

Teitelboim for useful conversations. We would particularly like to thank Larry Schulman for useful conversations, and for introducing us to the path decomposition expansion. J.J.H. was supported in part by the Royal Society, and M.O. by the SERC (UK). This work was supported in part by funds provided by the U.S. Department of Energy (D.O.E.) under Contract No. DE-AC02-76ER03069.

---

[1] K. Kuchař, in *Proceedings of the 4th Canadian Conference on General Relativity and Relativistic Astrophysics,* edited by G. Kunstatter, D. Vincent, and J. Williams (World Scientific, Singapore, 1992).

[2] C. Isham, "Canonical quantum gravity and the problem of time," Imperial College Report No. TP/91-92/25, gr-qc/9210011 (unpublished).

[3] The idea that the sum over histories may supply more a general formulation of quantum mechanics has been emphasized by Hartle. See, for example, J. B. Hartle, Phys. Rev. D **38**, 2985 (1988); **44**, 3173 (1991); in *Quantum Cosmology and Baby Universes,* Proceedings of the Jerusalem Winter School on Theoretical Physics, Jerusalem, Israel, 1990, edited by S. Coleman, J. Hartle, T. Piran, and S. Weinberg (World Scientific, Singapore, 1991); in Proceedings of the International Symposium on Quantum Physics and the Universe, Tokyo, Japan, 1991 (unpublished).

[4] J. J. Halliwell and J. B. Hartle, Phys. Rev. D **43**, 1170 (1991).

[5] J. Govaerts, *Hamiltonian Quantisation and Constrained Dynamics* (Leuven University Press, Leuven, Belgium, 1991).

[6] A. Auerbach and S. Kivelson, Nucl. Phys. **B257**, 799 (1985).

[7] M. Henneaux and C. Teitelboim, Ann. Phys. (N.Y.) **143**, 127 (1982).

[8] E. S. Fradkin and G. A. Vilkoviski, Phys. Lett. **55B**, 224 (1975); CERN Report No. TH-2332, 1977 (unpublished); I. A. Batalin and G. A. Vilkoviski, Phys. Lett. **69B**, 309 (1977).

[9] C. Teitelboim, Phys. Rev. D **25**, 3159 (1982).

[10] J. J. Halliwell, Phys. Rev. D **38**, 2468 (1988).

[11] J. B. Hartle and K. V. Kuchař, Phys. Rev. D **34**, 2323 (1986).

[12] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1955).

[13] That this is somewhat puzzling was, to our knowledge, first noticed by H. Ikemori, Phys. Rev. D **40**, 3512 (1989).

[14] P. van Baal, "Tunneling and the path decomposition expansion," Utrecht Report No. THU-91/19, 1991 (unpublished).

[15] R. H. Cameron, J. Math. Phys. **39**, 126 (1960).

[16] C. Itzykson and J. M. Drouffe, *Statistical Field Theory Vol. 1, From Brownian Motion to Renormalization and Lattice Gauge Theory* (Cambridge University Press, Cambridge, England, 1989).

[17] The validity of the PDX for arbitrary surfaces has been demonstrated in Ref. [14] by the use of a generalized Green's theorem.

[18] M. Kac, *Probability and Related Topics in the Physical Sciences* (Interscience, New York, 1959).

[19] N. Yamada, Sci. Rep. Tôhoku Univ., Ser. 8 **12**, 177 (1992).

[20] It is also possible to prove the path decomposition expansion for nonzero potential by using the Wiener measure on the set of all continuous Brownian paths [Sam Gutmann (private communication)]. For a description of the Wiener measure, and its role in the sum over histories, see, for example, L. S. Schulman, *Techniques and Applications of Path Integration* (Wiley, New York, 1982).

[21] L. Schulman and R. W. Ziolkowski, in *Path Integrals from meV to MeV,* edited by V. Sa-yakanit, W. Sritrakool, J. Berananda, M. C. Gutzwiller, A. Inomata, S. Lundqvist, J. R. Klauder, and L. S. Schulman (World Scientific, Singapore, 1989).

[22] J. B. Hartle, "Spacetime quantum mechanics and the quantum mechanics of spacetime," UCSB Report No. TH 92-21, 1993 (unpublished).

[23] The naive composition law, in the context of quantum gravity, has appeared in numerous papers over the years. Just a few of them are S. W. Hawking, in *General Relativity, An Einstein Centenary Survey,* edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979); J. A. Wheeler, in *Relativity, Groups and Topology,* edited by B. S. DeWitt and C. M. DeWitt (Gordon and Breach, New York, 1963); S. Coleman, Nucl. Phys. **B310**, 643 (1988); S. Giddings and A. Strominger, *ibid.* **B321**, 481 (1988). Many of the above authors appreciated that it may be incorrect, but not for the reasons given in this paper.

[24] J. J. Halliwell, in *Conceptual Problems of Quantum Gravity,* edited by A. Ashtekhar and J. Stachel (Birkhäuser, Boston, 1991).

[25] C. Teitelboim (private communication).

[26] One might reasonably ask which Green function is involved in (5.1), and to what extent it is defined in general spacetimes. One can think of $\mathcal{G}$ as being the Feynman or Hadamard propagators. These may be formally defined in *any* spacetime, using the proper time representation (4.1), although their interpretation in terms of positive and negative frequencies is generally not possible [C. Teitelboim (private communication)].

[27] Note that none of these claims are in contradiction with the fact that the causal Green function is well defined and obeys a composition law on any globally hyperbolic spacetime, even those possessing no Killing vectors [see S. Fulling, *Aspects of Quantum Field Theory in Curved Spacetime* (Cambridge University Press, Cambridge, England, 1989)]. The causal Green function does not appear to have a sum-over-histories representation, while the above conclusions specifically concern propagators generated by sums over histories.

[28] This conclusion does not exclude the possibility of the existence of a propagator not given by a sum over histories which obeys a composition law, in analogy with the causal Green function. Note also that the propagation amplitude generated by a sum over histories will obey the Wheeler-DeWitt equation [4], but this does not imply the existence of a composition law or a canonical formulation.