Phase-space approach to relativistic quantum mechanics. II. Geometrical aspects

Gerald R. Kaiser

Mathematics Department, University of Lowell, Lowell, Massachusetts 01854
(Received 9 June 1977)

The formalism introduced in Paper I [J. Math. Phys. 18, 952 (1977)] is made manifestly covariant by including as an admissible phase space any $2n$-dimensional submanifold of the forward tube $\mathcal{T}=\{x-iy\in C^{n+1}\,|\,y_0>|y|\}$ which is of the "product" form $\sigma=S-i\Omega_\lambda$, where $\Omega_\lambda=\{y\in R^{n+1}|\,y_0=(\lambda^2+y^2)^{1/2}\}$, $\lambda>0$, and S is any space-or-lightlike submanifold of space–time R^{n+1}. The σ's have natural symplectic structures covariant with respect to the Poincaré group, and a norm $\|\cdot\|_\sigma$ on the space $\mathcal{K}$ of solutions is defined by integrating with respect to the Liouville measure on σ. This automatically gives $\|f\|_\sigma^2$ as the total flux of a conserved space–time vector field, implying that $\|f\|_\sigma$ is independent of σ. Some inconsistencies encountered in the space–time theory of Klein–Gordon particles appear to be resolved in the phase-space framework.

1. INTRODUCTION

In a previous paper,[1] hereafter refered to as I, a phase-space formulation of relativistic quantum mechanics was initiated. A "coherent-state" representation of the Poincaré group $\mathcal{P}_+^\uparrow$ for massive scalar particles was constructed on the space $\mathcal{K}$ of positive-energy solutions of the Klein–Gordon equation in $(n+1)$-dimensional space–time. The elements of $\mathcal{K}$ extend as holomorphic functions to the *forward tube*[2] $\mathcal{T}$:

$$f(z)=(2\pi)^{-n/2}\int_\Omega \exp(-izp)\hat f(p)\,d\Omega(p), \tag{1.1}$$

where

$$z\equiv(\mathbf{z},z_0)\in\mathcal{T}\equiv\{x-iy\in C^{n+1}\,|\,y_0>|\mathbf{y}|\}\equiv R^{n+1}-iV_+,$$

$$\Omega\equiv\{p\in R^{n+1}\,|\,p_0=(m^2c^2+\mathbf{p}^2)^{1/2}\equiv\omega(\mathbf{p})\},\quad m>0,$$

$$d\Omega(p)=dp_1\cdots dp_n/\omega,\quad zp=z_0p_0-\mathbf{z}\cdot\mathbf{p},$$

and $\hat f(p)=\omega[f(\mathbf{x},0)]^\wedge(\mathbf{p})$ where $\wedge$ denotes the Euclidean Fourier transform in R^n. It was shown that the functions

$$e_z(p)=(2\pi)^{-n/2}\exp(i\bar z p) \tag{1.2}$$

[defined so that $f(z)=\langle e_z|\hat f\rangle_{L^2(\Omega)}$; $\bar z$ is the complex conjugate of $z\in\mathcal{T}$] belong to $L^2(\Omega)$ and represent optimal wavepackets (in the sense of Theorem 4 of I) centered in space–time about x (that is, focused at $\mathbf{x}$ at time[3] x_0) and traveling with expected energy–momentum proportional to y. Hence the submanifold

$$P_\lambda=\{x-iy\in\mathcal{T}\,|\,x_0=0,\,y^2=\lambda^2\},\quad\lambda>0, \tag{1.3}$$

can be interpreted as a classical "initial phase space" for the particle. The space of restrictions of $f\in\mathcal{K}$ to P_λ is denoted by $\mathcal{K}_\lambda$. It was shown that the expression

$$\|f\|_\lambda^2\equiv\int_{P_\lambda}|f(z)|^2\,d\mu_\lambda(z), \tag{1.4}$$

$$d\mu_\lambda(z)=C_\lambda\,dx_1\cdots dx_n\,dy_1\cdots dy_n \tag{1.5}$$

(where C_λ is a certain constant) defines a norm on $\mathcal{K}$ and $\mathcal{K}_\lambda$ which satisfies $\|f\|_\lambda=\|\hat f\|_{L^2(\Omega)}$ (Theorem 2 of I). This implies that $\mathcal{K}$ and $\mathcal{K}_\lambda$ are Hilbert spaces under the corresponding inner product $\langle\cdot|\cdot\rangle_\lambda$ and that the map $\hat f\mapsto f$ is unitary from $L^2(\Omega)$ onto $\mathcal{K}$ and $\mathcal{K}_\lambda$. We have a continuous resolution of the identity[4] in terms of the e_z, $z\in P_\lambda$,

$$\int_{P_\lambda}|e_z\rangle\langle e_z|\,d\mu_\lambda(z)=\mathbb{1}. \tag{1.6}$$

Now $L^2(\Omega)$ carries an irreducible unitary representation U of $\mathcal{P}_+^\uparrow$ (characterized by mass m and spin zero) under which the e_z are *covariant*,

$$U_g e_z=e_{gz},\quad g\in\mathcal{P}_+^\uparrow,\quad z\in\mathcal{T}. \tag{1.7}$$

Hence the corresponding representation on $\mathcal{K}$ (also denoted by U) is given by

$$(U_g f)(z)=f(g^{-1}z). \tag{1.8}$$

The set P_λ is not invariant under the action of $\mathcal{P}_+^\uparrow$ on $\mathcal{T}$, hence the above formalism is not manifestly covariant. In this paper we construct a natural class $\mathcal{S}_1$ of "phase spaces" $\sigma\subset\mathcal{T}$ and associated measures μ_σ to which the main results of I extend. $\mathcal{S}_1$ includes P_λ (the corresponding measure being μ_λ) and is invariant under $\mathcal{P}_+^\uparrow$, hence the formalism is freed from its dependence on P_λ and becomes manifestly covariant.

We begin in Sec. 2 by regarding $\mathcal{T}$ as an extended phase space,[5] on which $\mathcal{P}_+^\uparrow$ acts by canonical transformations. Candidates for phase space are $2n$-dimensional symplectic submanifolds[6,7] $\sigma\subset\mathcal{T}$, and $\mathcal{P}_+^\uparrow$ transforms different σ's into one another by canonical transformations. A $2n$-submanifold of the "product" form $S-i\Omega_\lambda$, where S is an n-submanifold of space–time and Ω_λ is a "mass hyperboloid," turns out to be symplectic (with respect to the induced structure) if and only if S is *space-or-lightlike*. Such σ's form a family $\mathcal{S}_1$ which is invariant under $\mathcal{P}_+^\uparrow$.

In Sec. 3 we extend the results of I to arbitrary $\sigma\in\mathcal{S}_1$. Each σ carries a canonically associated (Liouville) measure μ_σ. We show that for each $f\in\mathcal{K}$, $\|f\|_\sigma^2\equiv\|f\|^2_{L^2(\mu_\sigma)}$ is the total flux of a conserved vector field, hence independent of σ.

In Sec. 4 we show how the phase-space formalism can be used to resolve certain inconsistencies in the usual theory of Klein–Gordon particles.

2. SYMPLECTIC STRUCTURE

The Poincaré group acts on $\mathcal{K}$ by simply transforming the underlying space $\mathcal{T}$: $(U_g f)(z)=f(g^{-1}z)$, where $g=(a,\Lambda)\in\mathcal{P}_+^\uparrow$ and $gz=\Lambda z+a$. We wish to supply $\mathcal{T}$ with a symplectic structure[6,7] such that the map $z\mapsto gz$ is a canonical transformation for each $g\in\mathcal{P}_+^\uparrow$. That is, we need a 2-form α on $\mathcal{T}$ which is (a) closed ($d\alpha=0$),

 0022-2488/78/1902-0502$1.00

(b) nondegenerate [the $(n+1)$-fold exterior product α^{n+1} is never zero], and (c) invariant under $\mathcal{P}_+^\dagger$. The last condition means that $g^*\alpha=\alpha$, where g^* is the pullback map on forms induced by g (a brief description of which is given in the Appendix). Since every Poincaré-invariant function $\varphi(z)$ on $\mathcal{T}$ depends on z only through y^2, the most general invariant 2-form is given by

$$\alpha=\varphi(y^2)\,dy_\mu\,dx^\mu+\psi(y^2)y_\mu y_\nu\,dy^\mu\,dx^\nu. \tag{2.1}$$

(We are suppressing the wedge notation; thus, e.g., $dy^\mu\,dx^\nu=-dx^\nu\,dy^\mu$.) Now the action of $\mathcal{P}_+^\dagger$ on $\mathcal{T}$ is not transitive, and $\mathcal{T}$ decomposes into a union of orbits [Eq. (3.8) in I]

$$\mathcal{T}=\bigcup_{\lambda>0}P'_\lambda,\qquad P'_\lambda=\{z=x-iy\in\mathcal{T}\,|\,y^2=\lambda^2\}\approx\mathcal{P}_+^\dagger/\mathrm{So}(n). \tag{2.2}$$

As shown in I, each P'_λ gives rise to an equivalent representation of $\mathcal{P}_+^\dagger$. Our main results in this paper (Sec. 3) will be confined to P'_λ for a fixed λ. Since the second term in (2.1) contains $d(y^2)=2y_\mu\,dy^\mu$ as a factor, its restriction to P'_λ vanishes. Hence we will confine our attention to

$$\alpha=dy_\mu\,dx^\mu \tag{2.3}$$

without essential loss of generality. This form is symplectic as well as invariant, thus each $g\in\mathcal{P}_+^\dagger$ acts on $\mathcal{T}$ by canonical transformations.

$\mathcal{T}$ is an extended phase space, containing the time x^0 and the "energy" y_0 as a pair of free canonical variables. A $2n$-submanifold σ of $\mathcal{T}$ will be a candidate for phase space only if the pullback α_σ of α to σ is a symplectic form. Let σ be given by

$$\sigma=\{z\in\mathcal{T}\,|\,s(z)=h(z)=0\}, \tag{2.4}$$

where s and h are two real-valued, C^∞ functions on $\mathcal{T}$ such that $ds\wedge dh\neq0$ on σ. For example, $\sigma=P_\lambda$ can be obtained from $s(z)=x_0$ and $h(z)=\sqrt{y^2}-\lambda$. The pullback α_σ does not depend on s and h.

Proposition 1: The form α_σ is symplectic if and only if the Poisson bracket

$$\{s,h\}\equiv\frac{\partial s}{\partial x^\mu}\frac{\partial h}{\partial y_\mu}-\frac{\partial s}{\partial y_\mu}\frac{\partial h}{\partial x^\mu}\neq0 \tag{2.5}$$

everywhere on σ.

Proof: α_σ is closed since α is closed. Hence α_σ is symplectic iff it is nondegenerate, i.e., if and only if the nth exterior power α_σ^n of α_σ vanishes nowhere on σ. Now α_σ^n equals the pullback of α^n to σ, and

$$\alpha^n=n!\,d\hat{y}_\mu\,d\hat{x}^\mu, \tag{2.6}$$

where

$$d\hat{y}_\mu=(-)^\mu\,dy_0\cdots dy_{\mu-1}\,dy_{\mu+1}\cdots dy_n,\qquad d\hat{x}^\mu=(-)^\mu\,dx^n\cdots dx^{\mu+1}\,dx^{\mu-1}\cdots dx^0. \tag{2.7}$$

Let $\{u_1,\ldots,u_{2n},v_1,v_2\}$ be a basis for the tangent space $\mathcal{T}_z$ of $\mathcal{T}$ at $z\in\sigma$, with $u_1,\ldots,u_{2n}$ a basis for the subspace σ_z. Then since ds and dh vanish on the u_j,

$$\begin{aligned}(\alpha^n\wedge ds\wedge dh)(u_1,\ldots,u_{2n},v_1,v_2)&=\alpha^n(u_1,\ldots,u_{2n})(ds\wedge dh)(v_1,v_2)\\&=\alpha_\sigma^n(u_1,\ldots,u_{2n})(ds\wedge dh)(v_1,v_2).\end{aligned}$$

By assumption $ds\wedge dh\neq0$ at z, hence $(ds\wedge dh)(v_1,v_2)\neq0$. Thus α_σ is nondegenerate at z if and only if $\alpha^n\wedge ds\wedge dh\neq0$ at z. But by (2.6),

$$\alpha^n\wedge ds\wedge dh=n!\{s,h\}\,dy\,dx,$$

where

$$dy=dy_0\cdots dy_n,\quad dx=dx^n\cdots dx^0.$$

Hence $\alpha_\sigma^n\neq0$ at z if and only if $\{s,h\}\neq0$ at z. ∎

We denote the family of all symplectic $2n$-submanifolds of $\mathcal{T}$ by $\mathcal{S}_0$.

Proposition 2: Let $\sigma\in\mathcal{S}_0$ and $g\in\mathcal{P}_+^\dagger$. Then $g\sigma\in\mathcal{S}_0$ and the restriction $g:\sigma\to g\sigma$ is a canonical transformation from (σ,α_σ) onto $(g\sigma,\alpha_{g\sigma})$.

Proof: Let g^* denote the pullback map defined by g, taking forms on $g\sigma$ to forms on σ. Then the invariance of α implies

$$g^*\alpha_{g\sigma}=\alpha_\sigma, \tag{2.8}$$

thus $\alpha_{g\sigma}$ is nondegenerate, hence symplectic (it is automatically closed since α is closed). Thus $g\sigma\in\mathcal{S}_0$. To say that $g:\sigma\to g\sigma$ is canonical means precisely that α_σ and $\alpha_{g\sigma}$ are related by (2.8). ∎

We will be mainly interested in the special case where $h(z)=(y^2)^{1/2}-\lambda$ for some $\lambda>0$ and $s(z)$ depends only on x. Then $S\equiv\{x\in R^{n+1}\,|\,s(x)=0\}$ is an n-submanifold of space–time R^{n+1}, hence a candidate for configuration space, and $\sigma=S-i\Omega_\lambda$ where Ω_λ is the hyperboloid with $y_0=(\lambda^2+\mathbf{y}^2)^{1/2}$. The following theorem is physically significant in that it relates the pseudo-Euclidean geometry of space–time and the symplectic geometry of phase space.

Theorem 1: Let $\sigma=S-i\Omega_\lambda$ be as above. Then (σ,α_σ) is symplectic if and only if

$$\frac{\partial s}{\partial x_\mu}\frac{\partial s}{\partial x^\mu}\geq0,$$

that is, S is *space-or-lightlike*.

Proof: On σ, we have

$$\{s,h\}=\frac{\partial s}{\partial x^\mu}\frac{y^\mu}{\lambda}\neq0, \tag{2.9}$$

and we may assume $\{s,h\}$ to be positive without loss. For fixed $x\in S$, (2.9) must hold for all $y\in\Omega_\lambda$, hence for all $y\in V_+$. This implies that the vector $(\partial s/\partial x^\mu)$ is in the closure $\bar{V}_+$ of V_+, that is,

$$\frac{\partial s}{\partial x_\mu}\frac{\partial s}{\partial x^\mu}\geq0. \quad ∎$$

We denote the class of $\sigma=S-i\Omega_\lambda$ with

$$\frac{\partial s}{\partial x_\mu}\frac{\partial s}{\partial x^\mu}\geq0$$

by $\mathcal{S}_1$. $\mathcal{S}_1$ is a subfamily of $\mathcal{S}_0$ and is clearly invariant under $\mathcal{P}_+^\dagger$.

3. REPRESENTATION IN $\mathcal{K}_\sigma$, $\sigma\in\mathcal{S}_1$

Every symplectic manifold σ has a canonically associated measure μ_σ, hence an associated complex Hilbert space $L^2(\mu_\sigma)$. Given $\sigma\in\mathcal{S}_0$, we denote the vector space of restrictions of $f\in K$ to σ by K_σ. In general K_σ may not be contained in $L^2(\mu_\sigma)$. In this section we prove that when $\sigma\in S_1$, then $\|f\|_{L^2(\mu_\sigma)}\equiv\|f\|_\sigma=\|f\|_\lambda$ for all $f\in K$ [in particular, K_σ is a closed subspace of $L^2(\mu_\sigma)$ and we have the counterparts of Theorem 2 of I and corollary 1 of I for K_σ]. Hence each $\sigma\in S_1$ is as good as P_λ. The proof suggests that S_1 is the natural class of "phase spaces" for our approach.

It is remarkable, and somewhat surprising, that so large a class of phase spaces are admissible, in particular those with lightlike S. A possible application is suggested in Sec. 4.

Thus let $\sigma=S-i\Omega_\lambda\in S_1$ and define the forms μ (on T) and μ_σ (on σ) by

$$\mu=\frac{1}{n!}C_\lambda\alpha^n=C_\lambda d\hat{y}_\mu d\hat{x}^\mu,\quad \mu_\sigma=\frac{1}{n!}C_\lambda\alpha_\sigma^n, \tag{3.1}$$

where C_λ is given by (3.12) of I. We shall give a concrete expression for μ_σ. Since

$$\frac{\partial s}{\partial x_\mu}\frac{\partial s}{\partial x^\mu}\geq 0$$

and $ds\neq 0$ on σ, we can solve $ds=0$ (satisfied by the *restriction* of ds to σ) for dx^0 and substitute this into $d\hat{x}^k$. This (and a similar procedure for y) gives

$$\left.\begin{aligned} d\hat{x}^\mu&=\left(\frac{\partial s}{\partial x^0}\right)^{-1}\frac{\partial s}{\partial x^\mu}d\hat{x}^0\\ d\hat{y}_\mu&=\left(\frac{\partial h}{\partial y_0}\right)^{-1}\frac{\partial h}{\partial y_\mu}d\hat{y}_0=y_0^{-1}y^\mu d\hat{y}_0\end{aligned}\right\}\text{on }\sigma, \tag{3.2}$$

hence[8]

$$\mu_\sigma=C_\lambda\left(\frac{\partial s}{\partial x^0}y_0\right)^{-1}\left(\frac{\partial s}{\partial x^\mu}y^\mu\right)d\hat{y}_0\,d\hat{x}^0. \tag{3.3}$$

We identify σ with R^{2n} by solving $s(x)=0$ for $x^0=t(\mathbf{x})$ and mapping $(\mathbf{x}-i\mathbf{y},t(\mathbf{x})-i(\lambda^2+\mathbf{y}^2)^{1/2})$ to $(\mathbf{x},\mathbf{y})$. We further identify $d\hat{y}_0 d\hat{x}^0$ with Lebesgue measure $d^ny\,d^nx$ on R^{2n} (this amounts to choosing the nonstandard orientation[9] $dy_1\cdots dy_n dx^n\cdots dx^1$ of R^{2n}). This gives μ_σ as a measure on R^{2n}. Note that when $s(x)=x^0$ we obtain $\sigma=P_\lambda$ and $\mu_\sigma=\mu_\lambda$ [Eq. (3.11) of I]. Now $s(x)\equiv 0$ on σ implies

$$0=\frac{\partial}{\partial x^k}s(\mathbf{x},t(\mathbf{x}))=\frac{\partial s}{\partial x^k}+\frac{\partial s}{\partial x^0}\frac{\partial t}{\partial x^k}, \tag{3.4}$$

which can be substituted into (3.3) to give

$$\begin{aligned}\mu_\sigma&=C_\lambda\left(1-\frac{\partial t}{\partial x^k}\frac{y^k}{y_0}\right)d^ny\,d^nx\\ &=C_\lambda\left(1-\nabla t\cdot\frac{\mathbf{y}}{y_0}\right)d^ny\,d^nx.\end{aligned} \tag{3.5}$$

But

$$\frac{\partial s}{\partial x_\mu}\frac{\partial s}{\partial x^\mu}\geq 0$$

means $|\nabla t|\leq 1$, hence $\lambda>0$ ($|\mathbf{y}/y_0|<1$) implies that μ_σ is nondegenerate as expected. Equation (3.5) also shows that if $|\nabla t|=1$ for some $\mathbf{x}$, μ_σ becomes "asymptotically" degenerate at $(\mathbf{x},\mathbf{y})$ as $|\mathbf{y}|\to\infty$ in the direction of ∇t. That is, if S is lightlike at $(\mathbf{x},t(\mathbf{x}))$, then μ_σ becomes small as the velocity $\mathbf{y}/y_0$ approaches the speed of light in the direction of ∇t. This means that functions in $L^2(\mu_\sigma)$—and in particular, as we will show, in K_σ—are allowed high velocities in the direction of $\nabla t(\mathbf{x})$ at $(\mathbf{x},t(\mathbf{x}))\in S$.

Let $\sigma\in S_1$ and denote the Hilbert space of all complex-valued, measurable functions on σ with

$$\|f\|_\sigma^2\equiv\int_\sigma|f|^2\mu_\sigma<\infty \tag{3.6}$$

by $L^2(\mu_\sigma)$. If f is a C^∞ function on T, we restrict it to σ and define $\|f\|_\sigma$ by (3.6). To prove that $\|f\|_\sigma=\|f\|_\lambda$ for $f\in K$ we first show that each $f\in K$ defines a conserved (probability) current on space–time. Let

$$j^\mu(x)=C_\lambda\int_{\Omega_\lambda}|f(x-iy)|^2 d\hat{y}_\mu, \tag{3.7}$$

where Ω_λ has the orientation defined by $d\hat{y}_0$, so that $j^0(x)$ is positive. Then

$$\|f\|_\sigma^2=\int_S j^\mu(x)\,d\hat{x}^\mu, \tag{3.8}$$

where S is oriented by $d\hat{x}^0$ (the restriction of $d\hat{x}^0$ to S does not vanish since $|\nabla t|\leq 1$).

Theorem 2: Let $\hat{f}(\mathbf{p})$ be C^∞ with compact support. Then $j^\mu(x)$ is C^∞ and

$$\frac{\partial j^\mu}{\partial x^\mu}=0. \tag{3.9}$$

Proof: By (3.2),

$$j^\mu(x)=C_\lambda\int_{\Omega_\lambda}y^\mu|f(x-iy)|^2 d\Omega_\lambda(y), \tag{3.10}$$

where $d\Omega_\lambda(y)=d\hat{y}_0/y_0$. The function

$$F_x^\mu(y,p,q)\equiv y^\mu\exp[ix(p-q)-y(p+q)]\overline{\hat{f}(p)}\hat{f}(q)$$

is in $L^1(\Omega_\lambda\times\Omega\times\Omega)$, hence by Fubini's theorem,

$$\begin{aligned} j^\mu(x)&=(2\pi)^{-n}C_\lambda\int_{\Omega_\lambda}d\Omega_\lambda(y)\int_{\Omega\times\Omega}d\Omega(p)\,d\Omega(q)F_x^\mu(y,p,q)\\ &=(2\pi)^{-n}C_\lambda\int_{\Omega\times\Omega}d\Omega(p)\,d\Omega(q)\exp[ix(p-q)]\\ &\quad\times\overline{\hat{f}(p)}\hat{f}(q)\int_{\Omega_\lambda}d\Omega_\lambda(y)y^\mu\exp[-y(p+q)]\\ &=(2\pi)^{-n}C_\lambda\int_{\Omega\times\Omega}d\Omega(p)\,d\Omega(q)\exp[ix(p-q)]\\ &\quad\times\overline{\hat{f}(p)}\hat{f}(q)[(p^\mu+q^\mu)/\pi](2\pi\lambda/\eta)^{\nu+1}K_{\nu+1}(\lambda\eta),\end{aligned} \tag{3.11}$$

where $\eta=[(p+q)^2]^{1/2}\geq 2m$ and we have used (A6) of I. Differentiation under the integral sign to any order in x still gives an absolutely convergent integral since $\hat{f}$ has compact support; hence j^μ is C^∞. Differentiation with respect to x^μ brings down $i(p_\mu-q_\mu)$ from the exponent, hence (3.9) follows from $p^2=q^2=m^2$. ∎

Remark: Equation (3.9) can also be given a geometrical argument. Let $B_\lambda=\{y\in V_+|y_0>(\lambda^2+\mathbf{y}^2)^{1/2}\}$, oriented by $dy=dy_0\cdots dy_n$. Then $\Omega_\lambda=-\partial B_\lambda$ (Ω_λ is oriented by $d\hat{y}_0$), hence by Stokes' theorem[9]

$$\begin{aligned} j^\mu(x)&=-C_\lambda\int_{\partial B_\lambda}|f(x-iy)|^2 d\hat{y}_\mu\\ &=-C_\lambda\int_{B_\lambda}d(|f|^2 d\hat{y}_\mu)\\ &=-C_\lambda\int_{B_\lambda}\frac{\partial|f|^2}{\partial y_\mu}dy.\end{aligned} \tag{3.12}$$

To justify the use of Stokes' theorem it must be shown that the contribution from $|\mathbf{y}|\to\infty$ to the first integral vanishes. Then (3.9) is obtained by differentiating under the integral sign (which must also be justified) and using

$$\frac{\partial^2 |f|^2}{\partial x^\mu \partial y_\mu}=0, \tag{3.13}$$

which holds for $f\in K$. Equation (3.13) depends upon both the holomorphy (or antiholomorphy) of f and the fact that f satisfies the Klein–Gordon equation—that is, it holds for positive-energy (or negative-energy) solutions only. For such solutions (3.13) states that $(\partial|f|^2/\partial y_\mu)$ is a "microlocal" (local in phase space) conserved space–time probability current for each fixed $y\in V_+$. Hence the scalar function $|f(z)|^2$ is a kind of "potential" for the probability current. We can now prove our main result.

Theorem 3: Let $\sigma = S - i\Omega_\lambda \in \mathcal{S}_1$ and $f\in K$. Then $\|f\|_\sigma = \|f\|_\lambda$.

Remarks:

1. Properties (a)–(c) of Theorem 2 of I have counterparts for K_σ.

2. As before, we obtain a resolution of the identity by polarization.

3. Let $\hat f\in L^2(\Omega)$, let f be the corresponding function in K, and let f_σ be its restriction to $\sigma\in\mathcal{S}_1$. Then $\|\hat f\| = \|f\|_\sigma = \|f_\sigma\|_\sigma$. We will always regard K and K_σ as *Hilbert* spaces and *identity* $K_\sigma \approx K \approx L^2(\Omega)$.

Proof: We will prove that $\|f\|_\sigma = \|f\|_\lambda$ when $\hat f(\mathbf{p}) \in \mathcal{D}(R^n)$, which implies the result for arbitrary $\hat f\in L^2(\Omega)$ by continuity. Let S be given by $x_0 = t(\mathbf{x})$, and let

$$D_R=\{x\in R^{n+1}\mid |\mathbf{x}|<R,\ x_0\in[0,t(\mathbf{x})]\},$$
$$E_R=\{x\in R^{n+1}\mid |\mathbf{x}|=R,\ x_0\in[0,t(\mathbf{x})]\},$$
$$S_{0R}=\{x\in R^{n+1}\mid |\mathbf{x}|<R,\ x_0=0\},$$
$$S_R=\{x\in R^{n+1}\mid |\mathbf{x}|<R,\ x_0=t(\mathbf{x})\},$$

where $[0,t(\mathbf{x})]$ means $[t(\mathbf{x}),0]$ if $t(\mathbf{x})<0$. We orient S_{0R} and S_R by $d\hat x_0$, E_R by the "outward normal"

$$\hat r = \frac{1}{R}\sum_{k=1}^{n} x^k d\hat x^k, \tag{3.14}$$

and D_R so that $\partial D_R = S_R - S_{0R} + E_R$. Now let $\hat f(\mathbf{p})\in\mathcal{D}(R^n)$. Then j^μ is C^∞, hence by Stokes' theorem,

$$\left[\int_{S_R} - \int_{S_{0R}} + \int_{E_R}\right] j^\mu d\hat x^\mu = \int_{D_R} d(j^\mu d\hat x^\mu) = (-)^n \int \frac{\partial j^\mu}{\partial x^\mu} dx = 0.$$

We will show that

$$\Delta(R)\equiv \int_{E_R} j^\mu d\hat x^\mu \to 0 \text{ as } R\to\infty, \tag{3.15}$$

which implies that

$$\|f\|_\sigma^2 \equiv \lim_{R\to\infty}\int_{S_R} j^\mu d\hat x^\mu = \lim_{R\to\infty}\int_{S_{0R}} j^\mu d\hat x^\mu \equiv \|f\|_\lambda^2.$$

To prove (3.15), note that on E_R, $d\hat x^0 = 0$ and

$$d\hat x^k = x^k \frac{d\hat x^1}{x^1} = x^k\frac{d\hat x^2}{x^2} = \cdots = x^k \frac{d\hat x^n}{x^n}, \tag{3.16}$$

each form being defined except on a set of measure zero; hence $\hat r = R\, d\hat x^1/x^1$. By (3.10), $|j^k(x)|\le j^0(x)$, hence

$$|\Delta(R)| = \left|\sum_{k=1}^{n}\int_{E_R} j^k d\hat x^k\right| = \left|\sum_{k=1}^{n}\int_{E_R} j^k x^k \frac{d\hat x^1}{x^1}\right| \le n\int_{E_R} j^0(x) R\frac{d\hat x^1}{x^1} = n\int_{E_R} j^0(x)\hat r \equiv a(R). \tag{3.17}$$

Now by (3.11)

$$j^0(x) = \int_{R^{2n}} d^n p\, d^n q\, \exp[ix(p-q)]\phi(\mathbf{p},\mathbf{q}), \tag{3.18}$$

where

$$\phi(\mathbf{p},\mathbf{q}) = (2\pi)^{-n} C_\lambda \overline{\hat f(\mathbf{p})}\hat f(\mathbf{q})\left(\frac{p_0+q_0}{\pi p_0 q_0}\right)\left(\frac{2\pi\lambda}{\eta}\right)^{\nu+1} K_{\nu+1}(\lambda\eta),$$

with $\eta \equiv [(p+q)^2]^{1/2}\ge 2m$. Let D be the operator $\hat x\cdot\nabla_p$, where $\hat x = \mathbf{x}/R$, and observe that (for $x\in E_R$)

$$D\exp(ixp) = -iR\left(1-\frac{x_0}{R}\hat x\cdot \mathbf{v}\right)\exp(ixp) = -iR\xi(x,\mathbf{p})\exp(ixp), \tag{3.19}$$

where $\mathbf{v} = \mathbf{p}/p_0$. Since $\phi\in\mathcal{D}(R^{2n})$, there is a constant $\alpha<1$ such that $|\mathbf{v}|\le\alpha$ for all $\mathbf{p}\in\operatorname{supp}\phi$. Furthermore, since $|\nabla t|\le 1$, given any $\epsilon>0$ we have $|x_0|<R(1+\epsilon)$ for $x\in E_R$ with R large enough; hence

$$|\xi(x,\mathbf{p})|\ge 1-\alpha(1+\epsilon),\quad x\in E_R,\quad \mathbf{p}\in\operatorname{supp}\phi. \tag{3.20}$$

Choose $0<\epsilon<1/\alpha - 1$, substitute

$$\exp(ixp) = \frac{i}{R\xi(x,p)} D\exp(ixp),\quad x\in E_R \tag{3.21}$$

into (3.18) and integrate by parts:

$$j^0(x) = \frac{1}{iR}\int_{R^{2n}} d^n p\, d^n q\, \exp[ix(p-q)] D\left(\frac{\phi(\mathbf{p},\mathbf{q})}{\xi(x,\mathbf{p})}\right) \equiv \frac{1}{iR}\int_{R^{2n}} d^n p\, d^n q\, \exp[ix(p-q)]\phi_x^1(\mathbf{p},\mathbf{q}). \tag{3.22}$$

This process can be continued, giving (for $x\in E_R$)

$$j^0(x) = (iR)^{-N}\int_{R^{2n}} d^n p\, d^n q\,\exp[ix(p-q)]\phi_x^N(\mathbf{p},\mathbf{q}),\quad N=1,2,\cdots, \tag{3.23}$$

where

$$\phi_x^N(\mathbf{p},\mathbf{q}) = \left(D\frac{1}{\xi}\right)^N \phi(\mathbf{p},\mathbf{q}) \equiv \left(\hat x\cdot\nabla_p\left(1-\frac{x_0}{R}\hat x\cdot\mathbf{v}\right)^{-1}\right)^N \phi(\mathbf{p},\mathbf{q}). \tag{3.24}$$

Now $[D(1/\xi)]^N$ is a partial differential operator in $\mathbf{p}$ whose coefficients are polynomials in $D^k(1/\xi)$, $k=0,1,\ldots,N$. We will show that for $x\in E_R$ with R sufficiently large, there are constants b_k such that

$$|D^k(1/\xi)| < b_k,\quad k=0,1,2,\cdots \tag{3.25}$$

which implies that

$$\|\phi_x^N\|_{L^1(R^{2n})} < C_N,\quad x\in E_R,\quad N=1,2,\cdots \tag{3.26}$$

for constants C_N, so that by (3.17) and (3.23),

$$a(R) = n\int_{E_R} j^0 \hat r \le nR^{-N}\int_{E_R}\|\phi_x^N\|_{L^1(R^{2n})}\hat r(x) \le nR^{-N}C_N\cdot\frac{2\pi^{n/2}}{\Gamma(n/2)}R^{n-1}\int_0^{R(1+\epsilon)} dx_0 = \frac{2n\pi^{n/2}}{\Gamma(n/2)}C_N R^{n-N}(1+\epsilon)\to 0 \text{ as } R\to\infty$$

if we choose $N>n$. To prove (3.25), note that it holds for $k=0$ by (3.20) and let $u=\hat{x}\cdot\mathbf{v}$. Then

$$Du\equiv(\hat{x}\cdot\nabla_p)(\hat{x}\cdot\mathbf{p}/p_0)=\frac{1-u^2}{p_0},$$

and if

$$D^k u=P_k(u)/p_0^k, \qquad (3.27)$$

where P_k is a constant-coefficient polynomial, then

$$D^{k+1}u=\frac{P_k'(u)Du}{p_0^k}-\frac{kP_k(u)}{p_0^{k+1}}\cdot u\equiv\frac{P_{k+1}(u)}{p_0^{k+1}};$$

hence (3.27) holds for $k=1,2,\cdots$ by induction. Thus

$$D^k\xi=-\frac{x_0}{R}D^k u=-\frac{x_0}{R}\frac{P_k(u)}{p_0^k},\quad k=1,2,\cdots,$$

which implies

$$|D^k\xi|\leq\frac{1+\epsilon}{m^k}\max_{|u|\leq1}|P_k(u)|. \qquad (3.28)$$

But $D^k(1/\xi)$ is a polynomial in $1/\xi$ and $D\xi, D^2\xi,\ldots,D^k\xi$, hence (3.25) follows from (3.20) and (3.28). ∎

4. DISCUSSION

1. The phase-space approach appears to resolve some difficulties[10,11] encountered in the usual (space—time) theory of Klein—Gordon particles, which goes as follows: The counterpart of $\mathcal{K}$ is the space $\mathcal{H}$ of *boundary values* $f(x)$ of $f(x-iy)$, $f\in\mathcal{K}$, as $y\to0$ in V_+. For a given spacelike surface $S\subset R^{n+1}$, the norm in $\mathcal{H}$ is defined by[12]

$$\|f\|_S^2=\int_S J^\mu(x)\,d\hat{x}^\mu, \qquad (4.1)$$

$$J^\mu(x)=-\operatorname{Im}\left\{\overline{f(x)}\frac{\partial f(x)}{\partial x_\mu}\right\}. \qquad (4.2)$$

The current J^μ satisfies the continuity equation (3.9), ensuring that $\|f\|_S$ is independent of S. Now the Newton—Wigner postulates[13] for "localized states" uniquely determine these states (at time $x_0=0$) to be

$$\psi_{\mathbf{x}}(\mathbf{p})=(2\pi)^{-n/2}\sqrt{\omega}\exp(-i\mathbf{x}\cdot\mathbf{p}) \qquad (4.3)$$

[these are the generalized eigenvectors of the position operators X_k of Eq. (4.1) of I]. Hence the configuration-space probability density at time $x_0=0$ is given by

$$\rho(\mathbf{x})=|\langle\psi_{\mathbf{x}}|\hat{f}\rangle_{L^2(\Omega)}|^2$$
$$=(2\pi)^{-n}\iint\frac{d^n p\,d^n p'}{\sqrt{\omega\omega'}}\overline{\hat{f}(p)}\hat{f}(p')\exp[i\mathbf{x}\cdot(\mathbf{p}'-\mathbf{p})]. \qquad (4.4)$$

This expression does not coincide with

$$J^0(x)=\tfrac{1}{2}(2\pi)^{-n}\iint d^n p\,d^n p'\left(\frac{1}{\omega}+\frac{1}{\omega'}\right)\overline{\hat{f}(p)}\hat{f}(p')\times\exp[i\mathbf{x}\cdot(\mathbf{p}'-\mathbf{p})], \qquad (4.5)$$

which is the probability density associated with the current $J^\mu(x)$. In fact, $J^0\rho(x)$ cannot be the time component of *any* space—time vector field (which shows once more[13] that sharp localization, in the sense of Newton and Wigner, is incompatible with relativistic covariance). This is the first difficulty. The second difficulty is that even if one gives up the notion of localized states as a fundamental concept, $J^\mu(x)$ cannot be accepted as the probability current since it turns out that $J^0(x)$ can be negative.[11] (Even for positive-energy solutions, i.e., $f\in\mathcal{H}$, for which $\|f\|_S$ is actually positive definite!)

By contrast, our expression $j^0(x)$ is nonnegative as well as being the time component of a vector field; hence the second difficulty is clearly resolved in the phase-space framework. As for the first, note that if "sharp" localization in space is replaced with "soft" localization in phase space—i.e., replace the $\psi_{\mathbf{x}}$, which satisfy $\langle\psi_{\mathbf{x}'}|\psi_{\mathbf{x}}\rangle_{L^2(\Omega)}=\delta(\mathbf{x}'-\mathbf{x})$, by the e_z, $z\in\sigma$ for some $\sigma\in\mathcal{S}_1$—then we obtain

$$\tilde{\rho}(z)\equiv|\langle e_z|\hat{f}\rangle_{L^2(\Omega)}|^2=|f(z)|^2,\quad f\in\mathcal{K} \qquad (4.6)$$

as the probability density (with respect to μ_σ) in phase space. This, as we have seen, *is* compatible with (in fact, gives rise to!) the current $j^\mu(x)$. The price of this compatibility is that $j^0(x)$ can no longer be regarded as a *sharp* probability density but, rather, is the average of the "soft" density $\tilde{\rho}(z)$ over the "mass shell" Ω_λ. Accordingly, $j^\mu(x)$ depends on the parameter λ which measures, very roughly, the extent of spatial smearing associated with the continuous basis e_z, $z\in\sigma\subset P_\lambda'$.

2. In I we have seen that $\lambda=0$ can be admitted as a limiting value in the relation $\|f\|_\lambda=\|\hat{f}\|$, provided f is interpreted as a boundary-value function. Similar considerations apply to general $\sigma\subset P_\lambda'$, $\sigma\in\mathcal{S}_1$, when $\lambda\to0$. Thus the family $\mathcal{S}_1$ could be slightly enlarged. Note that the expression (3.5) for μ_σ shows that even when $\lambda=0$ (i.e., $|\mathbf{y}/y_0|=1$), μ_σ is nondegenerate except at those $(\mathbf{x},\mathbf{y})$ for which (a) $|\nabla t|=1$ (i.e., S is lightlike at $\mathbf{x}$) and (b) $\mathbf{y}/y_0=\nabla t$. This set is of measure zero in R^{2n}, and on it $f\in\mathcal{K}$ may develop singularities (caustics).

3. In recent years there has been progress in the quantization of field theories on surfaces other than $x_0=\text{const}$.[14,15] In particular, the so-called "lightlike quantization" uses surfaces which are everywhere lightlike, and appears to have practical applications.[14] The transition from $x_0=\text{const}$ to lightlike surfaces appears to present mathematical problems[15]; possibly the present formalism, when extended to quantum field theory, can be of help.

ACKNOWLEDGMENTS

I wish to express my deep appreciation to my advisor Lon Rosen and to Ivan Kupka. Both have generously given me their time and provided invaluable technical assistance.

APPENDIX

We give here a brief description of the pullback map, used throughout this paper.

Given two manifolds M and N of dimensions m and n respectively, a differentiable mapping $g: M\to N$ can be expressed locally as $g: U\to V$ where U and V are open subsets in R^m and R^n, respectively. Then the differential map g_* maps each tangent space M_x, $x\in M$, to the tangent space N_{gx} at $gx\in N$ and is given in local coordinates by

$$g^i_*(\xi) = A^i_j(x)\xi^j, \quad A^i_j \equiv \frac{\partial g^i}{\partial x^j}. \tag{A1}$$

The pullback map g^* takes the dual N^*_{gx} of N_{gx} to the dual M^*_x of M_x as follows: The linear form $p : N_{gx} \to R$ is mapped to the linear form $g^*p : M_x \to R$ defined by

$$(g^*p)(\xi) = p(g_*\xi). \tag{A2}$$

A 2-form on N is a bilinear, skew-symmetric mapping $\alpha : N_y \times N_y \to R$ on each tangent space N_y, $y \in N$. The map g defines a map (also denoted by g^*) taking 2-forms on N to 2-forms on M, as follows,

$$(g^*\alpha)(\xi, \xi') = \alpha(g_*\xi, g_*\xi'), \quad \xi, \xi' \in M_x. \tag{A3}$$

In case M is a submanifold of N and g denotes the inclusion map, we have $A^i_j(x) \equiv \delta^i_j$, $x \in M$; hence $g^*\alpha$ is the restriction of α to vectors tangent to M.

[1]G. Kaiser, J. Math. Phys. 18, 952 (1977).

[2]R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics and All That* (Benjamin, New York, 1964).

[3]In computing $\langle X_k\rangle = x_k \equiv \mathrm{Re}\ z_k$ in the state e_z [Eq. (I.4.2)], we neglected to mention the assumption that $x_0 \equiv \mathrm{Re}\ z_0 = 0$.

[4]J.R. Klauder, J. Math. Phys. 4, 1055, 1058 (1963); 5, 177 (1964); with J. McKenna, J. Math. Phys. 5, 878 (1964).

[5]H.C. Corben and P. Stehle, *Classical Mechanics* (Wiley, New York, 1960), 2nd ed.

[6]R. Abraham and J.E. Marsden, *Foundations of Mechanics* (Benjamin, New York, 1967).

[7]S. MacLane, "Geometrical Mechanics" I, II, Univ. of Chicago lecture notes (1968).

[8]When no confusion can result, we denote forms, their pullbacks, and the measures they define by the same symbol. Also, we write μ_σ instead of $d\mu_\sigma$ to avoid confusion with the exterior derivative.

[9]F. Warner, *Foundations of Differentiable Manifolds and Lie Groups* (Scott, Foresman & Co., London, 1971).

[10]A.O. Barut and S. Malin, Rev. Mod. Phys. 40, 632 (1968).

[11]B. Gerlach, D. Gromes, and J. Petzold, Z. Phys. 202, 401 (1967).

[12]S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper & Row, New York, 1961).

[13]T.D. Newton and E.P. Wigner, Rev. Mod. Phys. 21, 400 (1949).

[14]R. Jackiw, in *Springer Tracts in Modern Physics*, edited by G. Höhler (Springer, Berlin, 1972), Vol. 62.

[15]S. Fubini, A.J. Hanson, and R. Jackiw, Phys. Rev. D 7, 1732 (1973).