International Journal of Theoretical Physics, Vol. 8, No. 6 (1973), pp. 391–408

Hamiltonians Allowing Wave Equations

RICHARD ARENS

Received: 15 *September* 1972

Abstract

A general criterion of when a Hamiltonian system has a wave equation is set up, and all such Hamiltonian systems (and hence all wave equations) are found. It is shown that the correspondence is one-to-one.

1. *Introduction*

Ideas about the relation between the Hamiltonian of a system as conceived classically and the wave equation of its quantum-theoretic analogue have been current since the beginning of the theory. It is also well known that a correspondence between classical observables and operators cannot be set up if too many algebraic relations are to be preserved by this correspondence (see, for example, Joseph (1970) and Arens & Babbitt (1965)). Here we formulate a criterion of correspondence between Hamiltonians and wave equations which is sufficiently close to allow exactly one wave equation for each of the standard systems. The correspondence is so explicitly formulated that one can discover essentially all classical systems which have (in our sense of 'have') wave equations. We can show in fact that the Schrödinger equation, together with a rather interesting generalization of the Klein–Gordon equation (in which the mass is provided by a scalar field) already provide the list of all possible wave equations.

A wave equation $\mathscr{W}\varphi = 0$ is any equation satisfied by e^{if} whenever f is a solution to the Hamilton–Jacobi equation whose associated field of extremals represents a beam of *parallel* rays. In this case we call f a *planar* phase function. In Section 2 this condition of planarity is presented. It is shown that in a Hamiltonian (or Lagrangian) system there is an intrinsic *connection* giving rise to a generalized covariant differentiation, and planarity is defined in terms of the vanishing of the associated divergence of the field of extremals.

The construction of the wave equation does not proceed by attempting to assign operators to classical observables and this is to be expected in view of the inherent difficulties in the latter program, as already mentioned. Thus for some Hamiltonians there is no wave equation and, as we have already said, we can exhibit explicitly all Hamiltonians which do allow wave equations.

The construction presented here is related to but not equivalent to a method presented earlier (Arens, 1965). In the present scheme, for example, in Lorentz formulation of the motion of a charged particle in an electromagnetic field, the customary Klein–Gordon equation does not qualify for a wave equation (indeed there is none) unless the field is purely magnetic. This is in harmony with the fact that only then does the Klein–Gordon equation allow a decomposition into energy, and negative energy, submanifolds.

2. *Planar Phase Functions*

A familiar feature found in most text-books on quantum mechanics (e.g. Schiff, 1955, pp. 17–21) is a sketchy derivation or 'development' of the Schrödinger equation. These developments are made by combining the momentum–energy relation of the classical conceptualization of the system under consideration with some rather elementary and intuitive ideas about complex-valued waves.

Our intent here is to show that such derivations can be made absolutely rigorous provided one recognizes explicitly that the waves involved in the argument are planar. The waves involved in Schiff (1955) are indeed planar but one could get the impression that this is just done for simplicity. Moreover, for systems with more complicated Hamiltonians than those in the usual examples, the dynamically appropriate concept of planarity might not be the obvious one. In any case we will give a definition of planarity for phase functions. This definition will involve only the Hamiltonian of the system.

The concept of a Hamiltonian system involves a differentiable† manifold Q (the *configuration* space), the *phase space* (or cotangent bundle $T_1(Q)$ of Q (Sternberg, 1964, p. 144) and a function H defined on $T_1(Q) \times \mathbb{R}$. Thus this *Hamiltonian H* is a function of the position, 'momenta' and 'time'.

The dynamics of the system is defined as follows. Consider the *fundamental linear differential* form θ (Sternberg, 1964, p. 143, Th. 7.1) which is defined on $T_1(Q)$. This can be used to define a linear differential form $\lambda = \theta - Hdt$ on $T_1(Q) \times \mathbb{R}$. In terms of *canonical coordinates* $(x^1,\ldots,x^n, p_1,\ldots,p_n)$ in $T_1(Q)$, θ takes the form (Sternberg, 1964, p. 144, Th. 7.14)

$$\theta = p_i \, dx^i \qquad \text{(summation convention)}$$

and thus

$$\lambda = p_i \, dx^i - Hdt$$

One regularly requires that (Sternberg, 1964, p. 152)

$$\det\left(\frac{\partial^2 H}{\partial p_i \, \partial p_j}\right) \neq 0 \tag{2.1}$$

† We will henceforth omit this word 'differentiable' and tacitly assume that all manifolds, functions, etc., introduced have a suitable order of differentiability.

at every point of $T_i(Q) \times \mathbb{R}$. It then follows that, at each (α, τ) in $T_1(Q) \times \mathbb{R}$, there is exactly one vector ξ such that (1) $\xi(t) = 1$ and (2) $\langle d\lambda; \xi, \eta\rangle = 0$ for all η. The vectors thus obtained by considering all (α, τ) form a vector field which in coordinates has the appearance

$$\frac{\partial}{\partial t} + H^k \frac{\partial}{\partial x^k} - H_k \frac{\partial}{\partial p_k} \tag{2.2}$$

where here, as well as later, H^k means $\partial H/\partial p_k$ and H_k means $\partial H/\partial x^k$. We will refer to it as the dynamical flow, and denote it by F_H.

Let π_1 be the mapping of $T_1(Q) \times \mathbb{R}$ onto $Q \times \mathbb{R}$ wherein (α, τ) is mapped onto the pair (q, τ) where q in Q is the base point of the covector α. So now if f is a function defined on $Q \times \mathbb{R}$ then $f \circ \pi_1$ is a function defined on $T_1(Q) \times \mathbb{R}$.

Definition. Let f be a real-valued function defined on $Q \times \mathbb{R}$. Then f will be called a *phase function* if $F_H(f \circ \pi_1) = \langle F_H; \lambda\rangle$. When expressed in coordinates, this condition says more or less literally that the level surface $\{f = C + \varepsilon\}$ in $Q \times \mathbb{R}$ is obtained from $\{f = C\}$ by the Huyghens construction (Courant & Hilbert, 1931, p. 105).

Let f be expressed in terms of coordinates $x^1, \ldots, x^n$ in Q and t in $\mathbb{R}$, as $f = \varphi(x^1, \ldots, x^n, t)$. Then $f \circ \pi_1$ has formally the same form $\varphi(x^1, \ldots, x^n, t)$ where here $x^1, \ldots, x^n$ have to be lifted to $T_1(Q)$ and $p_1, \ldots, p_n$ are the conjugate canonical coordinates. From (2.2) and the coordinate form of λ (or from Courant & Hilbert, 1931, pp. 103–4) we have a familiar result.

Theorem. *f is a phase function if and only if*

$$\frac{\partial f}{\partial t} + H(\nabla f) = 0 \tag{2.3}$$

Here by ∇f we mean that the p_i are to be set equal to the $\partial f/\partial x_i$. The notation ∇f is intended to suggest something like df. However, df is a mapping from $Q \times \mathbb{R}$ to $T_1(Q \times \mathbb{R})$ and ∇f is the mapping from $Q \times \mathbb{R}$ to $T_1(Q) \times \mathbb{R}$ in which (q, τ) goes into $(d(f(\cdot, \tau))|_q, \tau)$. Here $f(\cdot, \tau)$ is the function for which $f(\cdot, \tau)(q) = f(q, \tau)$. Roughly speaking, ∇f is the differential of f-with-t-constant.

Equation (2.3) is the Hamilton–Jacobi relation. The condition that f be a phase function implies that the submanifold of all (α, τ) in $T_1(Q) \times \mathbb{R}$ such that $\alpha = \nabla f$ is preserved under F_H.

For a given τ, the Legendre transformation (Sternberg, 1964, p. 152) transforms the covector field defined by df (or ∇f) on $Q \times \{\tau\}$ into a vector field $L_f|_{t=\tau}$ on $Q \times \{\tau\}$. In terms of coordinates,

$$L_f|_{t=t} = X^i \frac{\partial}{\partial x^i}$$

where

$$X^i(q, \tau) = H^i(\nabla f(q, t)) \tag{2.4}$$

If one adds on a unit t-component, one obtains a vector field

$$\frac{\partial}{\partial t} + L_f$$

on $Q \times \mathbb{R}$ whose integral curves form the *field of extremals* associated with a solution of the Hamilton–Jacobi equation in Courant & Hilbert (1931), p. 102.

We are going to define a divergence for vector fields in Q, and when $L_f|_{t=\tau}$ has divergence 0 at some q in Q we will say that f is *planar* at (q, τ).

Say $x^1, \ldots, x^n$ are coordinates in Q. Then $x^1, \ldots, x^n, p_1, \ldots, p_n, t$ are coordinates in $Q \times \mathbb{R}$. Using (2.1) one can replace $p_1, \ldots, p_n$ by $u^1, \ldots, u^n$ where $u^i = H^i$. Thus we obtain a new local coordinate system $x^1, \ldots, x^n$, $u^1, \ldots, u^n, t$ for $T_1(Q) \times \mathbb{R}$. In these coordinates the dynamical flow F_H (see (2.2)) has the form

$$\frac{\partial}{\partial t} + u^i \frac{\partial}{\partial x^i} - \Delta^i \frac{\partial}{\partial u^i} \qquad \text{(summation convention)} \tag{2.5}$$

We will say that Δ^i is the ith coefficient of the *dynamic connection* relative to the coordinates $x^1, \ldots, x^n$ in Q.

Now let $y^1, \ldots, y^n$ be another coordinate system in Q. It leads canonically to a coordinate system $y^1, \ldots, y^n, q_1, \ldots, q_n, t$ in $T_1(Q) \times \mathbb{R}$ where $q_\alpha dy^\alpha = p_i dx^i$. We introduce the analogues of the u_i, namely $v^\beta = \partial H/\partial q_\beta$.

Proposition. *Let $\bar{\Delta}^\alpha$ be the coefficients of the dynamic connection relative to the new coordinate system. Then*

$$\bar{\Delta}^\alpha \frac{\partial x^k}{\partial y^\alpha} = \frac{\partial^2 x^k}{\partial y^\alpha \partial y^\beta} v^\alpha v^\beta + \Delta^k \tag{2.5}$$

The proof is as follows.

From $p_i dx^i = q_\alpha dy^\alpha$ we get $q_\alpha = p_i(\partial x^i/\partial y^\alpha)$. Now

$$u^k = \frac{\partial H}{\partial p_k} = \frac{\partial q_\alpha}{\partial p_k}\frac{\partial H}{\partial q_\alpha},$$

so $u^k = (\partial x^k/\partial y^\alpha) v\alpha$. We apply F_H to both sides of this equation, using $(\partial/\partial t) + u^i(\partial/\partial x^i) - \Delta^i(\partial/\partial u^i)$ on the left, and $(\partial/\partial t) + v^\alpha(\partial/\partial y^\alpha) - \bar{\Delta}^\alpha(\partial/\partial v^\alpha)$ on the right. This gives (2.5).

In the following discussion we will be dealing with some particular fixed value τ of time t.

Let Y be a vector field on Q. We will say that Y is *dynamically presentable* if it is the Legendre transform of some covector field α. This means in coordinates that $Y^i = H^i \circ \alpha$ where, as we just said, t is held fixed. If X and Y are vector fields in Q and Y is dynamically presentable, then we can form a new† vector field $\nabla_X Y$, using in any coordinate system the

† The symbol ∇ here is a rather modern one and not related to 'gradient' and thus not related to the ∇f in (2.4) either.

formula

$$\nabla_X Y = X^j \left[\frac{\partial Y^i}{\partial x^j} + \Delta_j{}^i \circ \alpha \right] \frac{\partial}{\partial x^i} \tag{2.6}$$

where

$$\Delta_j{}^i = \frac{1}{2} \frac{\partial \Delta^i}{\partial u^j}$$

Proposition. (1) *The right-hand side of* (2.6) *is the same in all coordinate systems.* (2) $\nabla_{fX+gY} = f\nabla_X + g\nabla_y$. (3) *If $\Delta_j{}^i$ depends linearly on the coordinates $u^1, \ldots, u^n$ then ∇ is an affine connection.*

Assertion (1) follows immediately from (2.5). (2) is of course obvious and is mentioned merely for comparison with the defining properties ∇_1 and ∇_2 for affine connections (Helgason, 1962, p. 26). The truth of (3) is obvious once one has these properties ∇_1, ∇_2 in mind.

We feel that this relation to affine connections warrants the term dynamic connection.

The relation to the older notation (cf. Eisenhart, 1962, p. 26) of covariant derivatives is established by recognizing that if Y is a dynamically presentable vector field, being indeed the Legendre transform of a covector field α then (a) in a given coordinate system, the Y^i are the contravariant components of Y and the components of α are the covariant components of Y;

(b) $$\frac{\partial Y^i}{\partial x^j} + \Delta_j{}^i \circ \alpha$$

are the components of a mixed tensor of type $T_1{}^1$: the *dynamical derivative* of Y.

The *dynamic divergence* of Y is now (consistently) the function on Q (involving however the parameter τ)

$$\operatorname{div} Y = \frac{\partial Y^i}{\partial x^i} + \Delta_i{}^i \circ \alpha \tag{2.7}$$

where the summation convention is adhered to.

For the dynamic divergence there is another, often more convenient, formula,

$$\operatorname{div} Y = \frac{\partial Y^i}{\partial x^i} - \left[\frac{1}{2\delta} F_H(\delta) \right] \circ \alpha \tag{2.8}$$

where δ is the determinant (2.1), F_H is the dynamic flow (perhaps in the form (2.2)) and $\circ\alpha$ means that after the differentiation (F_H) is performed the p_i are replaced by the covariant components of Y (compare Eisenhart, 1962, p. 32, Ex. 8).

We proceed to establish (2.8). In the first place we surely have $-\Delta^i = F_H(u^i) = F_H(H^i)$. Now

$$\frac{\partial}{\partial u^j} = g_{jk} \frac{\partial}{\partial p_k}$$

where (g_{ij}) is the inverse (matrix) of (H^{jk}). Thus

$$2\Delta_j{}^i = -g_{jk}\frac{\partial}{\partial p_k}[H_0{}^i + H^m H_m{}^i - H_m H^{im}]$$

where f^i means $\partial f/\partial p_i$, f_i means $\partial f/\partial x^i$, and f_0 means $\partial f/\partial t$. Proceeding,

$$2\Delta_j{}^i = -g_{jk}[H_0^{ik} + H^{mk} H^m H_m^{ik} - H_m{}^k H^{im} - H_m H^{imk}]$$
$$2\Delta_j{}^i = -g_{jk}[H_0^{ik} + H^m H_m^{ik} - H_m H^{imk}] + g_{jk}[H_m{}^k H^{im} - H_m{}^i H^{mk}]$$

The second bracket is antisymmetric in (i,k), while for $j = i$, its coefficient is symmetric. Thus in $\Delta_i{}^i$ this term drops out and we are left with

$$\begin{aligned} 2\Delta_i{}^i &= -g_{ik}[H_0^{ik} + H^m H_m^{ik} - H_m H^{imk}] \\ &= -g_{ik} F_H(H^{ik}) \end{aligned}$$

Now the operator F_H satisfies the rule $F_H(AB) = F_H(A)B + AF_H(B)$ and, therefore, whenever g and h are matrices inverse to each other then

$$F_H(\det h) = \det(h)\,\mathrm{trace}(gF_H(h)) \tag{2.9}$$

This formula results from the fact that $\det(\exp(h)) = \exp(\mathrm{trace}(h))$. Applying it to g_{ik} and $h_{ik} = H^{ik}$, we obtain

$$2\Delta_i{}^i = -(2\delta)^{-1} F_H(\delta)$$

Thereupon (2.7) gives (2.8).

3. *Two Examples*

Theorem. *Suppose $H = \frac{1}{2}g^{ij}(p_i + A_i)(p_j + A_j) + A_0$ where g^{ij}, A^i, V depend on $x^1, \ldots, x^n$ and t; and g^{ij} is a non-singular symmetric matrix. Then the dynamical divergence is the same as the Riemannian one based on the quadratic form $g_{ij}x^i x^j{}_j$, (g_{ij}) being the inverse to (g^{ij}).* (3.1)

Perhaps 'pseudo-Riemannian' would be a better term here, since we are not requiring (g^{ij}) to be positive definite.

The proof of (3.1) is as follows. We note $H^{ij} = g^{ij}$ and then we observe that (2.8) has exactly the same form as Eisenhart (1962, p. 32, Ex. 8).

While the divergence is the same as the Riemannian one, the dynamical connection is different from the Riemannian connection. Suppose in fact $Q = \mathbb{R}^n$ with the usual Cartesian metric $(g^{ij} = \delta^{ij})$. Then $\Delta_j{}^i$ turns out to be

$$\frac{1}{2}\left(\frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j}\right)$$

The dynamics of a charged particle is governed by a Hamiltonian of this type, and the $\Delta_j{}^i$ essentially constitute the components of the magnetic field.

The next theorem mentions normal coordinates (Eisenhart, 1962, p. 55) in a Riemannian manifold. Normal coordinates at P_0 are coordinates $x^1, \ldots, x^n$ such that if g_{ij} are the components of the metric then $g_{ij} - \delta_{ij}$ and its first-order derivatives vanish at P_0.

Theorem. *Let* $H = A_0 + V^k P_k + R$ *where* R *is the positive square root of* $m^2 + g^{ij}(p_i + A_i)(p_j + A_j)$. A_0, v^k, m, g^{ij}, A_i *are functions of* $x^1, \ldots, x^n$ *and* t *such that* (3.2)

$$m \text{ is always positive} \tag{3.3}$$

$$\text{the matrix } (g^{ij}) \text{ is symmetric and positive definite} \tag{3.4}$$

Then

$$\delta = M^2 R^{-n-2} \det(g^{ij}) \tag{3.5}$$

The remainder of the theorem will be formulated for coordinates normal at P_0. Partial derivation with respect to time is indicated by suffixing a '0', and with respect to x^k, by suffixing a 'k'. We also abbreviate $p_k + A_k$ by q_k. Moreover, we will always sum over repeated indices.

Then (see (2.8)):

The term $-(1/2\delta)F_H(\delta)$ *is the sum of three quantities*:

$$-\frac{1}{m}\left[m_0 + \left(v^k + \frac{q_k}{R}\right)m_k\right] \tag{3.6}$$

$$\frac{n+2}{2R^2}\left[\tfrac{1}{2}g_0^{ij} q_i q_j - v_k{}^s q_s q_k + q_i(A_{i0} - A_{0i} + v^k A_{ik} + v_i{}^s A_s) + m(m_0 + v^k m_k)\right] \tag{3.7}$$

and

$$-\tfrac{1}{2}g_0^{ii} \tag{3.8}$$

We indicate first the proof of (3.5). First of all,

$$H^{ij}\left(= \frac{\partial^2 H}{\partial p_i \, \partial p_j}\right) = \frac{g^{ij}}{R} - \frac{g^{ik} q_k g^{js} q_s}{R^3}$$

Let M_{ij} be the inverse of the matrix g^{ij}. Then

$$RH^{ij} M_{js} = \delta_s{}^i + R^i q_s \qquad \text{where } R^i = -\frac{g^{ik} q_k}{R^2}.$$

Now $\det(\delta_s{}^i + R^i q_s)$ is the value at -1 of $p(\lambda) = \det(R^i q_s - \lambda\delta_s{}^i)$, the characteristic polynomia of $(R^i q_s)$. This latter has rank one and so its non-zero eigenvalue is its trace, whence $p(\lambda) = (-\lambda)^{n-1}(R^i q_i - \lambda)$ and so

$$\det(RH^{ij} M_{js}) = R^i q_i + 1 \tag{3.9}$$

The latter is $m^2 R^{-2}$. Of course $\det(M_{js})$ is the reciprocal of $\det(g^{ij})$ from which (3.23) follows at once from $\det(RH^{ij} M_{js}) = R^n \delta \det(M_{js})$.

The calculation of $-(1/2\delta)F_H(\delta)$ is naturally broken into a sum of three parts, because it is $-\frac{1}{2}$ times a logarithmic derivative. Hence we calculate $-(1/2\delta)F_H(\,)$ of m^2, R^{-n-2}, and $\det(g^{ij})$. The results, we claim, are (3.6), (3.7) and (3.8) respectively. In each case we use (2.2) for example, (3.6) results easily from the fact that

$$H^i = v^i + \frac{q^i}{R}$$

For (3.7) one has essentially to compute $F_H(R^2)$. This is $(\partial/\partial t)R^2 + \{H, R^2\}$ where $\{\ ,\ \}$ is the Poisson bracket (cf. (2.2)). Here $\{H, R^2\} = \{A_0 + v^k p_k, R^2\}$ because $\{R, R^2\} = 0$. At this point we leave the remaining calculations to the reader, remarking that those p_k have eventually to be replaced by $q_k - A_k$, and also that $g^{ij} = \delta^{ij}$ and $g_k^{ij} = 0$ at P_0.

For (3.8) we use (2.9) with $h_{ij} = g^{ij}$.

4. *The Differential Equations for Planar Phase Functions*

A function f on $Q \times \mathbb{R}$ is a phase function if and only if the Hamilton–Jacobi relation (2.3) holds. It is planar at a point (p_0, τ) in $Q \times \mathbb{R}$ if for $Y = L_f$, the Legendre transform of ∇f (see Section 2, particularly (2.4)), the dynamical divergence (2.8) vanishes at (p_0, τ). The α in formula (2.8) is in this application ∇f, which is d (f restricted to $t = \tau$).

The vanishing of the divergence we will call the *divergence condition*. It takes the form of a partial differential equation. The purpose of this section is to exhibit this equation for each of the two systems exhibited and studied in Section 3.

In each case we assume normal coordinates at P_0. When the metric is indefinite, as it may be in (3.1), normal coordinates means that $g^{ij} = 0$ at P_0 for $i \neq j$ and $= \lambda_k$ for $i = j = k$ where $\lambda_k = \pm 1$, and also $g_k^{ij} = 0$ at P_0. However, it is often better to leave g^{ij} evaluated at P_0 in the form g^{ij} rather than attempting to express it in terms of λ_i (or λ_j!).

Again, new suffixes on any tensor mean partial derivatives (and at P_0 this is the same as covariant differentiation).

Theorem. *Let H be as in* (3.1) *and suppose f is a phase function. Then f is planar at P_0 if and only if* (4.1)

$$g^{ij}(f_{ij} + A_{ij}) = 0 \qquad \text{at } P_0 \tag{4.2}$$

Proof. In normal coordinates, the divergence of Y is $Y_i^{\ i}$. Now Y^i for the vector field associated with f is, by (2.4),

$$g^{ij}(p_j + A_j)|_{p_k = f_k} = g^{ij}(f_j + A_j)$$

We take $\partial/\partial x^i$ of this and evaluate at P_0. The result is (4.2).

Theorem. *Let H be as in* (3.2) *and suppose f is a phase function. Then f is planar at P_0 if and only if, at P_0* (4.3)

$$-R^{-3}(f_{ij} + A_{ij})q_i q_j - R^{-3} m m_j q_j + v_j^{\ j} + R^{-1}(f_{jj} + A_{jj}) + \mathscr{E} = 0 \tag{4.4}$$

where $q_i = f_i + A_i$ and

$$\begin{aligned}\mathscr{E} = &-m^{-1}[R^{-1} q_j m_j + m_0 + v^k m_k] + R^{-2}(\tfrac{1}{2}n + 1) \\ &\times [\tfrac{1}{2} g_0^{ij} q_i q_j - v_j^{\ i} q_i q_j + q_i(A_{i0} - A_{0i} + v^k A_{ik} + v_i^{\ k} A_k) + m(m_0 + v^k m_k)] \\ &- \tfrac{1}{2} g_0^{ii}\end{aligned} \tag{4.5}$$

Here we sum over repeated indices. The term especially exhibited as $\mathscr{E}$ is the 'extra' term of the divergence provided by $-(1/2\delta)F_H(\delta)$ (see (3.6)–(3.8)).

We account for (4.4) as follows. According to (2.4),

$$Y^i = H^i|_{p_k=f_k} = v^i + \frac{g^{ij}(f_i + A_i)}{R}$$

Here R now stands for the old R with p_k replaced by f_k. Although this Y^i is still going to be differentiated, we can already assume

$$Y^i = v^i + \frac{f_i + A_i}{R}$$

where $R^2 = m^2 + (f_i + A_i)(f_i + A_i)$ because the g_k^{ij} are 0 at P_0. Upon taking $\partial/\partial x_i$ of this Y^i we obtain (4.4) except for the $\mathscr{E}$, as we are supposed to do. We emphasize that the R in (4.4) and (4.5) is the positive square root of $m^2 + (f_i + A_i)(f_i + A_i)$.

5. *Wave Equations*

Let Q be the configuration space and H the Hamiltonian of a dynamical system. Let $\mathscr{W}$ be a linear at most second-order partial differential operator on $Q \times \mathbb{R}$:

$$\mathscr{W} = W + W^\alpha \frac{\partial}{\partial x^\alpha} + W^{\alpha B} \frac{\partial^2}{\partial x^\alpha \, \partial x^B} \tag{5.1}$$

where $\alpha = 0, 1, \ldots, x^n$ and $x^0 = t$ or

$$\mathscr{W} = W + W^0 \frac{\partial}{\partial t} + W^k \frac{\partial}{\partial x^k} + W^{00} \frac{\partial^2}{\partial t^2} + 2W^{0k} \frac{\partial^2}{\partial t \, \partial x^k} + W^{jk} \frac{\partial^2}{\partial x^i \, \partial x^k} \tag{5.1}$$

Here W, W^α, $W^{\alpha\beta}$ are functions on $Q \times \mathbb{R}$ and $W^{\alpha\beta} = W^{\beta\alpha}$.

Definition. We will say that $\mathscr{W}\varphi = 0$ is a *wave equation* for the system if $\mathscr{W}(e^{if}) = 0$ at (P_0, τ) whenever f is a local phase function in a neighborhood of (P_0, τ) and planar at (P_0, τ), and $i = \sqrt{-1}$.

Here the term 'local phase function in a neighborhood' is used and must be defined. It shall mean that f is defined in some neighborhood of (P_0, τ) and satisfies the Hamilton–Jacobi equation (2.3) on that neighborhood. A reader interested in basic principles might then well wonder what the point of the theorem involving (2.3) is, if one would just as well have defined phase function as meaning (2.3). Our answer is a combination of two remarks: (1) One wants to emphasize the relation of phase functions to the Huyghens construction; (2) it is possible to define dynamical system in a more general way, a purely local way, so that any open neighborhood in $Q \times \mathbb{R}$ is a (sub) system and then a local phase function is a phase function for the subsystem.

Another feature of the definition is that the condition of planarity (e.g. (4.4)) is required only at (P_0, τ) rather than, say, in a neighborhood. This would lead to serious difficulties inasmuch as we would have to insure the existence of many simultaneous solutions (in a neighborhood) of (4.4) and (2.3).

Lemma. *$\mathscr{W}\varphi = 0$ is a wave equation if and only if* (5.2)

$$\begin{aligned} &W - iW^0 H + iW^k f_k + W^{00}(-iH_0 + iH^j H_j - HH) \\ &\quad + 2W^{0j}(-iH_j + Hf_j) - W^{jk} f_j f_k \\ &\quad + \{iW^{00} H^j H^k - iW^{0j} H^k - iW^{0k} H^j + iW^{jk}\} f_{jk} = 0 \end{aligned} \tag{5.3}$$

at (P_0, τ) whenever the divergence condition holds at (P_0, τ). It is understood that each p_k in H and its derivatives is replaced by f_k before evaluation at (P_0, τ).

HH is merely the square of H.

The proof of (5.2) is as follows. In the definition of wave equation one can certainly test $\mathscr{W}$ only with f such that $f(P_0, \tau) = 0$, since (1) $\mathscr{W}$ is *linear* and (2) only the derivatives of f enter into the Hamilton–Jacobi equation (2.3) and the divergence condition. Now for such f, $\mathscr{W}(e^{if}) = \mathscr{W}(1 - if - \frac{1}{2}f^2)$ at (P_0, τ). Using (2.3) to get rid of f_0, f_{00}, and f_{0j} results in $\mathscr{W}(e^{if})$ being at (P_0, τ), the left side of equation (5.3).

In Section 4 we defined the divergence condition. In general, this is a quasi-linear differential equation of the form

$$D^{jk} f_{jk} + D = 0 \tag{5.4}$$

where the D^{jk} and D contain the f_i. (The D is used† to remind us of 'divergence condition'.) Examples are (4.2) and (4.4). We can assume $D^{jk} = D^{kj}$.

Lemma. *$\mathscr{W}\varphi = 0$ is a wave equation if and only if there is for each (P_0, τ) an expression λ involving the indeterminates $f_1, \ldots, f_n$ such that* (5.5)

$$A + B^{jk} f_{jk} = \lambda(D^{jk} f_{jk} + D) \tag{5.6}$$

is an identity in the indeterminates $f_1, \ldots, f_n$ and $f_{11}, \ldots, f_{nn}$ (where however $f_{ij} = f_{ji}$), the coordinates having been set equal to those for (P_0, τ).

The sufficiency of (5.5) is easy to see. The necessity is equally important, especially when we want to show that sometimes no wave equation is possible. The necessity depends primarily on the fact that the left-hand sides of (5.3) and (5.4) depend linearly on the f_{jk}. Suppose that $\mathscr{W}\varphi = 0$ is a wave equation. We can obtain a solution to (2.3) (Hamilton–Jacobi equation) with any preassigned values of $f_1, \ldots, f_{11}, \ldots, f_{nn}$ at (P_0, τ). Fix some values for $f_1, \ldots, f_n$ and regard the f_{jk} $(j \leqslant k)$ as the Cartesian coordinates in real $\frac{1}{2}n(n+1)$-dimensional space. By hypothesis, $D^{jk} f_{jk} + D = 0$

† These D^{jk} and the D are functions defined on $Q \times \mathbb{R}$, just as the W, H, etc., are, and as the A and B^{jk}, about to be introduced, are. Incidentally, this A is unrelated to the field whose components are $A_1, \ldots, A_n$.

implies both the real and the imaginary parts of $A + B^{jk} f_{jk} = 0$. Hence these are each real multiples of the real $D^{jk} f_{jk} + D$ and from this the required (complex) λ can be assembled. Its value depends on the values chosen for $f_1, \ldots, f_n$ at (P_0, τ). This completes the proof of (5.5).

We are now prepared to consider examples.

Theorem. *Let Q be a manifold and let g be a 1-parameter family of non-degenerate contravariant symmetric tensor fields of order 2. Let A be a 1-parameter family of covariant vector fields on Q and let A_0 be a function on $Q \times \mathbb{R}$. Then the system with Hamiltonian*

$$H = \tfrac{1}{2} g^{jk}(p_j + A_j)(p_k + A_k) + A_0$$

has a unique wave equation

$$\frac{1}{i}\frac{\partial \varphi}{\partial t} + \tfrac{1}{2} g^{jk}\left(\frac{1}{i}\frac{\partial}{\partial x^j} + A_j\right)\left(\frac{1}{i}\frac{\partial}{\partial x^k} + A_k\right)\varphi + A_0 \varphi = 0 \tag{5.7}$$

Here the operators $\partial/\partial x^j$ are to be interpreted as *covariant* differentiations appropriate to whatever kind of tensor field on which they are about to operate.

The proof of (5.7) begins by a statement of (5.6) assuming that the coordinates are normal at (P_0, τ). (The relation expressed by (5.6) is the same in all coordinate systems.) So we write (5.6), using (4.2) but merely referring to (5.3) for A and B^{jk}.

$$A + B^{jk} f_{jk} = \lambda g^{jk}(f_{jk} + A_{jk}) \tag{5.8}$$

Thus $B^{jk} = \lambda g^{jk}$, and $B^{jk} = 0$ for $j \neq k$. This is an identity in $f_1, \ldots, f_n$, and $H^j = g^{jk}(f + A_k)$. Therefore (one must look at (5.3)) W^{00} must be 0. After that one sees that W^{0k} must be 0, too. Thus λ does not really depend on $f_1, \ldots, f_n$ in this example and is constant. (Everything has been evaluated at P_0, τ in these equations.) The relation $A = \lambda g^{jk} A_{jk}$, written out with $W^{00} = W^{0k} = 0$ reads

$$W - iW^0[\tfrac{1}{2} g^{jk}(f_j + A_j)(f_k + A_k) + A_0] - i\lambda g^{jk} f_j f_k = \lambda g^{jk} A_{jk}$$

This yields

$$W^0 = 2\lambda, \quad W^j = 2\lambda g^{jk} A_k, \quad \text{and} \quad W = i\lambda(g^{jk} A_j A_k + 2A_0) + \lambda g^{jk} A_{jk}.$$

Except for the common factor λ, the equation is now uniquely determined (and is obviously all one can mean by 'unique equation'). If we let $\lambda = -i/2$ it comes out exactly as asserted with the partial derivatives as *written*. Considering that we have normal coordinates, it must have the covariant differentiation form in other coordinate systems. So much then for the uniqueness. The existence follows from the fact that (5.8) does hold with the values $W, \ldots, \lambda$ just found.

This is of course the Schrödinger equation.

Theorem. *Let Q, g, A, A_0 be as in* (5.7), *but require g positive definite. Let v be a* 1*-parameter family of vector fields on Q, and let m be a positive real-valued function on $Q \times \mathbb{R}$. Then the system with Hamiltonian*†

$$H = A_0 + v^j p_j + R$$

where

$$R = [m^2 + g^{jk}(p_j + A_j)(p_k + A_k)]^{1/2}$$

has at most one wave equation; and it has none at all unless all of the following hold. For each (P_0, τ), in terms of coordinates normal at (P_0, τ) (5.9)

$$A_{0k} - A_{k0} = v_k{}^j A_j + v_j A_{kj} \tag{5.9.1}$$

$$g_0^{jk} - v_k{}^j - v_j{}^k = 0 \qquad \text{for } j \neq k \tag{5.9.2}$$

$$g_0^{11} - 2v_1{}^1 = g_0^{22} - 2v_2{}^2 = \cdots = g_0^{nn} - 2v_n{}^n = -\frac{m_0 + v^j m_j}{m} \tag{5.9.3}$$

Here suffixed zero means partial derivative with respect to t. A suffix not 0 indicates covariant differentiation with respect to g of the type appropriate to the tensor field involved.

To prove (5.9) we begin by observing that $H^j = v^j + (q_j/R)$, where $g_j = f_j + A_j$. If $f_1, \ldots, f_n$ are regarded as independent variables, as we will (and did in (5.7)) we can express them in terms of $H^1, \ldots, H^n$ and take them as the indeterminates. (The additional indeterminates will be the $f_{jk}, j \leqq k$, as before.)

As in the proof of (5.7), we have to verify the definition of wave equation at each point (P_0, τ) of $Q \times \mathbb{R}$. We will assume that the coordinates $x^1, \ldots, x^n$ are normal at P_0 with respect to the metric g, since it suffices to do this in one coordinate system. Therefore relations involving partial derivatives with respect to these coordinates at P_0 correspond to relations involving the corresponding covariant derivatives at P_0 in any other coordinate system.

Let us consider the identity $B^{jk} = \lambda D^{jk}$. It helps to let $\lambda = i\mu$. We consult (5.3) and (4.4). After dividing out the i and using $q_j = R(H^i - v^i)$, we get

$$W^{00} H^j H^k - W^{0j} H^k - W^{0k} H^j + W^{jk} = \frac{\mu}{R}[\delta^{jk} - (H^j - v^j)(H^k - v^k)]$$

Consider this first for $j = k = 1$ and then for $j = 2$, $k = 1$. Use that relation to eliminate μ/R. This results in

$$\begin{aligned}(W^{00} H^1 H^1 - 2W^{01} H^1 + W^{11})[-(H^2 - v^2)(H^1 - v^1)] \\ = (W^{00} H^2 H^1 - W^{02} H^1 - W^{01} H^2 + W^{21})[1 - (H^1 - v^1)(H^1 - v^1)]\end{aligned}$$

We consider this an identity in $H^1, \ldots, H^n$. The quartic terms cancel, so we equate the coefficients of $(H^1)^3$: $W^{00} v^2 = W^{02}$. Hence

(1) $$W^{0j} = W^{00} v^j \quad \textit{for all } j$$

† The theorem holds also if $H = A_0 + v^j p_j - R$, with the same wave equation.

Equating coefficients of $(H^1)^2$ yields $W^{12} = W^{00}v^1v^2$. Equating coefficients of H^1H^2 yields $W^{11} = W^{00}(-1 + v^1v^1)$. This clearly implies

$$W^{jk} = W^{00}(-\delta^{jk} + v^j v^k) \tag{2}$$

Putting these into our original equation shows that it is an identity if and only if $-W^{00} = \mu/R$.

We now consult (5.3) and (4.4), and consider the relation $A = \lambda D$. We note (5.3)

$$\begin{aligned} A &= W - iW^0 H + iW^k f_k \\ &\quad + W^{00}[-iH_0 + iH^j H_j - HH + 2v^j(-iH_j + Hf_j) - (\delta^{jk} + v^j v^k) f_j f_k] \end{aligned}$$

From $A = \lambda D$ and $\lambda = -iRW^{00}$ it follows that W^{00} cannot ever be 0, for this would make W^0, W^k, and $W = 0$. So we can let $W^{00} = -1$ and $\lambda = iR$.

We now write out $A = \lambda D$ in full (see (4.4)).

$$\begin{aligned} &W - iW^0 H + iW^k f_k + iH_0 - iH^j H_j + HH + 2iv^j H_j - 2v^j f_j H \\ &+ (\delta^{jk} - v^j v^k) f_j f_k - iR\{-R^{-3} A_{ij} q_i q_j - R^{-3} m m_j q_j + v_j{}^j + R^{-1} A_{jj} + \mathscr{E}\} \\ &= 0 \end{aligned} \tag{3}$$

For $\mathscr{E}$, see (4.5). When the f_k are expressed as $q_k - A_k$, this is an algebraic identity in $q_1, \ldots, q_n$. It actually is equivalent to two polynomial identities.

The first is obtained by taking from (3) those terms containing an odd power of R. The second equation consists of those terms having only even powers of R. This second equation (after multiplying by i and various simplifyings) is

$$\begin{aligned} &iW + W^0[A_0 + v^j(q_j - A_j)] - W^j(q_j - A_j) - A_{00} - v_0{}^j(q_j - A_j) - v^j A_{0j} \\ &\quad - v^k v_k{}^j(q_j - A_j) + i(A_0{}^2 - A_j A_j + m^2 + 2A_j q_j) - m^{-1} m_j q_j + A_{jj} = 0 \end{aligned}$$

This implies that

$$W^j = W^0 v^j - v_0{}^j - v^k v_k{}^j + 2iA_j - m_j m^{-1} \tag{4}$$

and

$$\begin{aligned} -iW &= W^0(A_0 - v^k A_k) + W^k A_k - A_{00} + v_0{}^k A_k \\ &\quad - v^k A_{0k} + v^k v_k{}^j A_j + i(A_0{}^2 - A_j A_j + m^2) + A_{jj} \end{aligned} \tag{5}$$

These just tell us what the W^j and W are, in terms of W^0 and the Hamiltonian.

The first (odd powers of R) equation is

$$\begin{aligned} R^2\left\{W^0 + 2iA_0 - \frac{m_0 + v^k m_k}{m} + v_j{}^j - \tfrac{1}{2}g_0^{jj}\right\} + \frac{n}{2}[\tfrac{1}{2}g_0^{jk} q_j q_k - v_j{}^k q_k q_j \\ + q^j(A_{j0} - A_{0j} + v^k A_{jk} + v_j{}^k A_k) + m(m_0 + v^k m_k)] = 0 \end{aligned}$$

This requires (5.9.1), (5.9.2), (5.9.3) to hold and also tells us that

$$W^0 = -2iA_0 + \frac{m_0 + v^k m_k}{m} \tag{6}$$

If the reader will accept our report on the results of laborious computations, then (5.9) is now established.

This (5.9) differs from (5.7) in that no wave equation exists unless some relations hold between the given fields A_j, A_0, v^j, m: namely (5.9.1)–(5.9.3).

These conditions are not automatically fulfilled even when m is constant, v is 0, and $g^{jk} = \delta^{jk}$. This is of course the case of a charged particle in an electromagnetic field. (5.9.1) remains, and says that there is a wave equation if and only if the electric field $A_{k0} - A_{k0}$ is 0.

The actual list of coefficients of the operator $\mathscr{W}$ is as follows.

$$W^{00} = -1, \qquad W^{0j} = W^{j0} = -v^j, \qquad W^{jk} = \delta^{jk} - v^j v^k$$

We introduce

$$\kappa = \frac{m_0 + v^j m_j}{m}$$

Then

$$W^0 = -2iA_0 + \kappa,$$
$$W^j = 2i(A_j - A_0 v^j) - v_0{}^j - v^k v_k{}^j - m_j m^{-1} + \kappa v^j$$
$$W = (A_0)^2 - A_k A_k - m^2 + i(\kappa A_0 - A_k m_k m^{-1} - A_{00} - v^k A_{0k} + A_{kk})$$

With our previous understanding about covariant derivatives the operator $\mathscr{W}$ can be written

$$\begin{aligned}\left(\frac{1}{i}\frac{\partial}{\partial t} + \frac{v^j}{i}\frac{\partial}{\partial x^j} + A_0 + \frac{i\kappa}{2}\right)^2 + \delta^{jk}\frac{\partial^2}{\partial x^j\,\partial x^k} \\ + (2iA_k - m^k m^{-1})\frac{\partial}{\partial x^k} - m^2 - A_k A_k \\ + i(A_{kk} - A_k m_k m^{-1}) - \frac{1}{2}\frac{\partial\kappa}{\partial t} - \tfrac{1}{4}\kappa^2\end{aligned} \tag{5.9.5}$$

Thus the wave equation can be written in the form $T^2\varphi = S\varphi$ where S contains only spatial derivatives and T is first order in the time derivative. The natural next step is to consider the relation of the solutions of $T^2\varphi = S\varphi$ to the solutions of $T\Psi = S^{\frac{1}{2}}\Psi$ and $T\Psi = -S^{\frac{1}{2}}\Psi$, as is done for the Klein–Gordon equation (Schweber, 1961, pp. 63–4). To have any formal relation between the φ and the Ψ, the operators T and S must commute. For this reason, we have calculated the commutator of the T and S in our $\mathscr{W}$ (5.9.5). We present the result only for $\kappa = 0$, when in fact the wave equation is quite reminiscent of Klein–Gordon:

$$\left(\frac{1}{i}\frac{\partial}{\partial t} + \frac{v^j}{i}\frac{\partial}{\partial x^j} + A_0\right)^2 \varphi - g^{jk}\left(\frac{1}{i}\frac{\partial}{\partial x^j} + A_j\right)\left(\frac{1}{i}\frac{\partial}{\partial x^k} + A_k\right)\varphi = m^2\varphi \tag{5.9.6}$$

(The m here may still be non-constant.)

Theorem. *The commutator of*

$$T = \frac{1}{i}\frac{\partial}{\partial t} + \frac{v^k}{i}\frac{\partial}{\partial x^k} + A_0$$

and

$$S = g^{jk}\left(\frac{1}{i}\frac{\partial}{\partial x^j} + A_j\right)\left(\frac{1}{i}\frac{\partial}{\partial x^k} + A_k\right) + m^2$$

is

$$-i(g_0^{jk} - v_k{}^j - v_j{}^k)\frac{\partial}{\partial x^j\,\partial x^k}$$
$$+ [-2g_0^{jk}A_j + 2(A_{0k} - A_{k0} - v^j A_{kj} + A_j v_j{}^k) - v^k_{jj}]\frac{\partial}{\partial x^k}$$
$$- 2mi(m_0 + v^k m_k) - ig_0^{jk}(A_jA_k - iA_{jk}) + 2iA_k(A_{0k} - A_{k0} - v^j A_{kj})$$

The relevance of this is as follows. If S and T commute, we obtain, after a little rearrangement, two conditions (among others) *one of which is precisely* (5.9.1) and the other

$$g_0^{jk} - v_k{}^j - v_j{}^k = 0 \tag{5.9.7}$$

obviously related to (5.9.2).

It appears, in other words, that the side-conditions (5.9.1)–(5.9.3) resulting from our concept of wave equation are very similar to those conditions which are needed to relate the solutions of $T^2\varphi = S\varphi$ to those of $T\Psi = S^{\frac{1}{2}}\Psi$ and $T\Psi = -S^{\frac{1}{2}}\Psi$.

6. *What Systems Have Wave Equations?*

The question above is equivalent to the following: what wave equations are there? The answer is: practically no others than the two examples given by (5.7) and (5.9).

For our main theorem here we require some real analyticity. A real analytic function is representable by its Taylor series.

Theorem. *Suppose Q is a connected real analytic manifold and suppose H is a real analytic function on $T_1(Q) \times \mathbb{R}$. Suppose H is a Hamiltonian for which there is a wave equation whose coefficients do not all vanish at any point of $Q \times \mathbb{R}$.* (6.1)

Suppose that in coordinates $x^1, \ldots, x^n, t$ for $Q \times \mathbb{R}$ the equation is (cf. (5.1))

$$W^{00}\frac{\partial^2\varphi}{\partial t^2} + 2W^{0j}\frac{\partial^2\varphi}{\partial t\,\partial x^j} + W^{jk}\frac{\partial^2\varphi}{\partial x^j\,\partial x^k} + \text{lower order terms} = 0 \tag{6.2}$$

Then in the coordinates $x^1, \ldots, x^n, p_1, \ldots, p_n, t$ for $T_1(Q) \times \mathbb{R}$, the Hamiltonian H satisfies an equation

$$W^{00}H^2 - 2(c + W^{0j}p_j)H + W^{jk}p_jp_k = 2a^kp_k + 2b$$

where c, a^k, b (as well as W^{00}, W^{0j}, W^{jk}) *are functions of $x^1, \ldots, x^n, t$.*

Before proving this we state a corollary.

Theorem. *If the coefficient W^{00} in* (6.2) *is identically zero, then the system is one of those presented in* (5.7). *If the coefficient W^{00} is never* 0, *then the* *system is one of those presented in* (5.9). *If* W^{00} *is not identically* 0 *but has some zeros, then the Hamiltonian still looks like that of* (5.9) *but the tensor fields* g, A, A_0, m *may have singularities at those zeros.* (6.3)

Let us show now how (6.2) follows from (6.1). Suppose W^{00} is 0. Then

$$(c + W^{0j}p_j)H = \tfrac{1}{2}W^{jk}p_jp_k + a^kp_k + b$$

This shows that $W^{0j} = 0$ for all j. For if W^{0j} is not 0, then the right-hand side would have to be divisible (as a polynomial in p) by $c + W^{0j}p_j$. This would make H linear in the p's forcing the determinant of H^{jk} to vanish. But for a Hamiltonian, this is not allowed.

If we now knew that c is not identically zero, we would be able to conclude that H has the desired form because of the analyticity assumption. So suppose c were identically 0. Then W^{jk} would be 0, and the wave equation would be entirely a first-order equation. We would then have, according to our fundamental relation (5.6), the identity

$$W - iW^0H + iW^kf_k = \lambda(H^{jk}f_{jk} + D).$$

Now the H^{jk} cannot all vanish, so $\lambda = 0$ and $W - iW^0H + iW^kf_k = 0$. If $W^0 = 0$ then immediately all the coefficients of the wave equation would vanish. So W^0 is not 0, thus making H a first-degree polynomial in $p_1, \ldots, p_n$. This, as we have remarked, cannot be.

This finishes the case $W^{00} = 0$. If W^{00} is never 0, we may as well take it to be 1. Then our equation may be written as

$$(H - c - W^{0j}p_j)^2 = g^{jk}p_jp_k + a^kp_k + b \tag{6.4}$$

where these g's, a's, and b are easily calculated. By a change of coordinates we may diagonalize g^{jk}. Of course it will be a *sum* of squares, but will there be n terms? If there were less than n terms, or in other words if some p's were missing, then the corresponding a^k would also be zero (by the evident positivity in (6.4)). Thus if there were less than n terms we would have $H = c + W^{0j}p_j +$ a function of only part of the p's. Such an H would have $\det H^{jk} = 0$ at the (P_0, τ) in question. This contradiction shows that g^{jk} is positive definite and we can complete the square on the right:

$$(H - c - W^{0j}p_j)^2 = g^{jk}(p_j + A_j)(p_k + A_k) + d$$

This function d can obviously never be negative, but can it ever be 0? If d were 0 then H would not be differentiable with respect to p_k at $p_k = -A_k$. So d is never 0 and H has the form $A_0 + v^jp_j \pm R$ where R is as in (5.9).

Finally, if W^{00} has both zero and non-zero values, let G be the subset of $Q \times \mathbb{R}$ where W^{00} is not 0. By the previous reasoning, H has on the part of $T_1(Q) \times \mathbb{R}$ lying above G the form announced in (6.2), whose proof is now complete. It is interesting to note that the points of $Q \times \mathbb{R}$ not included in G constitute only a set of dimension lower than that of $Q \times \mathbb{R}$.

We now proceed to prove (6.1). The variables of interest are the $p_1, \ldots, p_n$ and functions on $T_1(Q) \times \mathbb{R}$ which depend only on $x^1, \ldots, t$ will simply be called constants. Thus W, W^0, W^k, etc., are constants. By H^j, etc., we mean as before $\partial H/\partial p_j$. For functions F and K soon to be introduced we use the same superscript notation for partial derivatives.

Consider the relation $B^{jk} = \lambda D^{jk}$ (see (5.6)). Let $\lambda = iy$. Then

$$W^{00} H^j H^k - W^{0j} H^k - W^{0k} H^j + W^{jk} = yH^{jk}$$

Let $F = \frac{1}{2} W^{jk} p_j p_k - W^{0j} p_j H + \frac{1}{2} W^{00} H^2$. (This last H^2 is really the square of H!) It is easy to verify that $F^{jk} = KH^{jk}$ where

$$K = y - W^{0j} p_j + W^{00} H$$

Suppose K were a constant c. Then $(F - cH)^{jk} = 0$ and $F - cH$ has the form $a^k p_k + b$ where a^k and b are constants and after multiplying everything by 2 this is precisely the assertion of (6.1).

Therefore it remains only to show K is constant. From $F^{jk} = KH^{jk}$ we get $F^{jki} = K^i H^{jk} + KH^{ijk}$ and so $K^i H^{jk} = K^j H^{ik}$. Multiplying by dp_i (and of course summing) shows that $H^{jk} dK = K^j dH^k$.

Now suppose K were not constant. Then there is a j and a point A in $\mathbb{R}^n$ where K^j is not 0 and hence in a neighborhood of A we have $dH^k = LdK$ for some function L. This of course implies that each partial derivative H^k has the form $f^k(K)$ where $f^1, \ldots, f^n$ are some n functions of a single variable. From $H^{kj} = H^{jk}$ one gets $f^{k\prime}(K) K^j = f^{j\prime}(K) K^k$. Moreover, the functions $(f^{1\prime}(K), f^{2\prime}(K), \ldots, f^{n\prime}(K))$ are never all zero at any point, since then $\det H^{jk}$ would vanish.

Consider the level surface Z where K has some value c. We have $f^k(c) K^j = f^j(c) K^k$ and this implies that the normal $(K^1, \ldots, K^n)$ to Z is always parallel to $(f^{1\prime}(c), \ldots, f^{n\prime}(c))$. This obviously implies that Z is a piece of a hyperplane (i.e., flat).

Since each H^k depends only on K, it follows that each H^k is constant on Z. Therefore if $(v_1, \ldots, v_n)$ is a vector orthogonal to $(f^{1\prime}(c), \ldots, f^{n\prime}(c))$ then $v_i(\partial/\partial p_i)$ must kill H^k. That is to say $v_i H^{ki} = 0$. Since $n > 1$, such a relation implies $\det H^{ki} = 0$, which by the definition of Hamiltonian is, of course, impossible.

Thus (6.1) is proved.

In a way (6.1) is not very surprising, for if the wave equation is at most second order in $\partial/\partial t$, then H ought to satisfy a second-degree equation. Thus the conclusion of (6.1) might be regarded as a consequence of our having made 'second order' part of the definition of 'wave equation'. It is our guess that (6.1) would still be true if the restriction to second order were dropped from the definition (cf. Arens, 1965, p. 158). The degree 2 in (6.1) is probably a consequence just of the fact that the divergence condition involves only the second- and first-order derivatives of the phase function f.

References

Arens, R. and Babbitt, D. (1965). Algebraic difficulties of preserving dynamical relations when forming quantum-mechanical operators, *Journal of Mathematical Physics*, **6**, 1071.

Arens, R. (1965). Wave mechanization of dynamical systems, *Journal of Mathematical Analysis and Applications*, **10**, 135.

Courant, R. and Hilbert, D. (1931). *Methoden der Mathematischen Physik*, Erster Band, zweite verbesserte Auflage. Julius Springer Verlag, Berlin.

Eisenhart, L. P. (1962). *Riemannian Geometry*. Princeton University Press, Princeton, N.J.

Helgason, S. (1962). *Differential Geometry and Symmetric Spaces*. Academic Press, New York.

Joseph, A. (1970). Derivations of lie brackets and canonical quantization, *Communications in Mathematical Physics*, **17**, 210.

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

Schweber, S. S. (1961). *Introduction to Relativistic Quantum Field Theory*. Row-Peterson, Evanston, Ill.

Sternberg, S. (1964). *Differential Geometry*. Prentice Hall, Englewood Cliffs, New Jersey.