

# Imprints of the Quantum World in Classical Mechanics

Maurice A. de Gosson  
*Universität Wien, NuHAG  
Fakultät für Mathematik  
A-1090 Wien*

Basil Hiley  
*TPRU, Birkbeck  
University of London  
London, WC1E 7HX*

May 17, 2010

## Abstract

The imprints left by quantum mechanics in classical (Hamiltonian) mechanics are much more numerous than is usually believed. We show that the Schrödinger equation for a nonrelativistic spinless particle is a classical equation which is equivalent to Hamilton's equations. Our discussion is quite general, and incorporates time-dependent systems. This gives us the opportunity of discussing the group of Hamiltonian canonical transformations which is a non-linear variant of the usual symplectic group.

## 1 Introduction

*“Where did that [the Schrödinger equation] come from? Nowhere. It came out of the mind of Schrödinger, invented in his struggle to find an understanding of the experimental observations in the real world. ” (Richard Feynman in [4].)*

Similar statements abound in the physical literature; they are found in both introductory and advanced text on quantum mechanics, and we can read them on the web in various blogs and forums. However, they are strictly speaking *not true*. Schrödinger [20] was led to his equation from his knowledge of the classical Hamilton-Jacobi approach which has a close connection with the eikonal of classical wave theory. His original proposal was not firmly based on rigorous mathematics as he himself acknowledged but his intuition was correct. Attempts to provide a detailed relationship between classical

and quantum mechanics has remained somewhat of a problem because we have been left with two widely different mathematical formulations, one involving particles evolving under dynamical laws in a phase space, the other involving operators and waves operating in a Hilbert space. This leaves the impression that there are two very different worlds, the classical and the quantum. However it is clear that we do not inhabit two different worlds and so we are left with the puzzle as how to relate the two formalisms?

Already in studying the properties of light, we see in one phenomena both aspects. Geometric optics gives us light rays travelling in straight lines, while wave optics gives us interference and diffraction. In this phenomena we see the essential mathematics emerging. Ray optics emerges from symplectic geometry, while wave optics emerges from the geometry of the covering group of the symplectic group,  $Sp(2n)$ , namely the metaplectic group,  $Mp(2n)$ . Ray tracing involves symplectic flows, while the wave evolution is identified with the metaplectic flows. These are not separate flows but deeply related as the metaplectic flows can be mathematically ‘lifted’ from symplectic flows and the metaplectic flows can be projected onto symplectic flows.

The fact that classical mechanics has the same mathematical structure as ray optics, namely the symplectic geometry, suggest that quantum mechanics in its wave description could be related to such a covering structure. Indeed if we restrict ourselves to quadratic Hamiltonians, we find that the metaplectic flow is determined exactly by Schrödinger’s equation, in which the classical Hamiltonian is replaced by the Weyl Hamiltonian,  $H \xleftrightarrow{\text{Weyl}} \widehat{H}$ . Unfortunately the generalisation to all physically relevant Hamiltonians has always floundered on the mathematical difficulties involved in investigating the covering properties of general group of canonical transformations which involve non-linear transformations. However such difficulties should not force us into a ‘two world’ situation.

Indeed we have shown recently [6, 7] that Hamiltonian mechanics, in its symplectic formulation, in a sense reproduces what is considered as being one of the hallmarks of quantum mechanics, the uncertainty principle in its strong form (the Schrödinger–Robertson inequalities); the argument is based on a deep and new property of canonical transformations, Gromov’s [9] non-squeezing theorem, alias “the principle of the symplectic camel”.

In the present paper we will show that one can *mathematically* derive *riorously* from Hamiltonian mechanics, the Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = H(x, -\hbar \nabla_x, t) \psi$$

and this without any recourse to any physical ‘quantum-mechanical’ hy-

potheses. The aim of this paper is to show how this can be done.

In fact we will show that there is a one-to-one and onto correspondence between Hamiltonian flows and the quantum evolution group, which only becomes apparent if one uses a deep property of symplectic covariance together with Stone's theorem on one-parameter groups of unitary operators.

Up to some technical assumptions on the Hamiltonian function ensuring us that the flows they generate exist for all times, we will in fact show that:

*There is a one-to-one and onto correspondence between Hamiltonian flows generated by a Hamiltonian  $H$  and strongly continuous unitary one-parameter groups satisfying Schrödinger's equation with Hamiltonian operator  $\widehat{H} = H(x, -i\hbar\nabla_x, t)$  obtained from  $H$  by Weyl quantization. Equivalently, the Hamilton equations*

$$\dot{x} = \nabla_p H(x, p, t) \text{ , } \dot{p} = -\nabla_x H(x, p, t)$$

*are mathematically rigorously equivalent to Schrödinger's equation*

$$i\hbar \frac{\partial \psi}{\partial t} = H(x, -i\hbar\nabla_x, t)\psi.$$

The proof relies on two deep mathematical results

- The first is that the Weyl correspondence is the *only* quantization (*that is*, pseudodifferential calculus) which is symplectically covariant. We emphasize the word “only”, because while it is reasonably well-known, both in mathematics and physics that Weyl operators are symplectically covariant, the converse, namely the only operators that are symplectically covariant are Weyl operators, is largely ignored in the literature; this uniqueness is the key to our argument.
- The second is Stone's theorem from functional analysis which concerns the infinitesimal generators of strongly continuous unitary one-parameter groups. It states that each such one-parameter group determines a self-adjoint operator.

In Section 2.2 where we define and study the group  $\text{Ham}(2n, \mathbb{R})$  of all Hamiltonian canonical transformations following Banyaga [2] and Polterovich [18]. It turns out that by systematically considering general Hamiltonians that depend on  $t$  makes our task is considerably easier than if we had limited ourselves to the case where  $H$  only depends on the phase space variables,  $z = (x, p)$ .

The hypothesis of a bounded phase space (implying a fortiori a bounded Universe) is not essential; it is added as an ad hoc assumption in order to avoid mathematical difficulties related to self-adjointness and the fact that Hamiltonian vector fields need not be complete.

**CAVEAT LECTOR:** We are *not* claiming that we are deriving quantum mechanics from classical mechanics; what we are doing is the following: knowing that quantum mechanics exists, we show that the mathematical formulation of quantum mechanics in its Schrödinger formulation lies within Hamiltonian mechanics. This does not imply that quantum mechanics –as a *physical* theory– can be reduced to classical mechanics. For as Mackey (who used to see quantum mechanics as a refinement of Hamiltonian mechanics) stresses in [15] (p. 106), quantum mechanics is not just an algorithm for attaching a self-adjoint operator to every classical Hamiltonian, because such a program would overlook many facts: first, quantum mechanics rules out a large number of conceivable Hamiltonians, and secondly there are features of quantum mechanics (such as spin) which do not manifest themselves in the classical limit. However, adds Mackey, one can argue convincingly that classical mechanics must have independent existence since it is only what quantum mechanics looks like under certain limiting conditions. Our program does not contradict Mackey’s insight because our approach shows that the imprint left by quantum mechanics in classical mechanics is more than could be expected at first sight from the limiting conditions.

## 2 The Dynamical Groups $Sp(2n, \mathbb{R})$ , $Mp(2n, \mathbb{R})$ and $Ham(2n, \mathbb{R})$

### 2.1 Motivation from Optics

We recall that ray tracing in Gaussian optics uses an identical formulation of classical mechanics, which for the purposes of this paper can be written in the form

$$z_t = f_t^H(z_0)$$

where the phase variables  $z = (x_1 \dots x_n, p_1 \dots p_n)$  and  $f_t^H$  is the flow determined by Hamilton’s equations of motion

$$\dot{x} = \nabla_p H(x, p) \quad , \quad \dot{p} = -\nabla_x H(x, p). \quad (1)$$

where  $H(x, p)$  is some physically relevant Hamiltonian.

We now ask the question, “Is it possible to lift this flow onto the covering group of  $Sp(2n, \mathbb{R})$  so as to display possible wave properties of the symplectic flow?” The answer is, “Yes, provided we restrict ourselves to quadratic Hamiltonians.”

To see how this works let us write the quadratic Hamiltonian in the form

$$H(z, t) = \frac{1}{2} z^T M(t) z = \frac{1}{2} \begin{pmatrix} x \\ p \end{pmatrix}^T M(t) \begin{pmatrix} x \\ p \end{pmatrix} \quad (2)$$

where  $M(t)$  is a  $2n \times 2n$  symmetric matrix depending in a  $C^\infty$ .

Now the metaplectic group  $Mp(2n, \mathbb{R})$  can be realised as a group of unitary operators acting on the space  $L^2(\mathbb{R}^n)$  of square integral functions on  $\mathbb{R}^n$ . The metaplectic group is generated by the Fourier quadratic transformation  $F_t^W$  defined as follows (see e.g. de Gosson [5], Chapter 7): let  $W$  be a real quadratic form of the type

$$W(x, x') = \frac{1}{2} P x \cdot x - L x \cdot x' + \frac{1}{2} Q x' \cdot x' \quad (3)$$

with  $P = P^T$ ,  $Q = Q^T$ , and  $\det L \neq 0$ ; we are writing  $Px \cdot x$  for  $x^T P x$ , etc. Denoting by  $m\pi$  the choice of an argument for  $\det L$  modulo  $2\pi$  we have

$$F_t^W \psi(x) = \left( \frac{1}{2\pi i \hbar} \right)^{n/2} i^m \sqrt{|\det L|} \int_{\mathbb{R}^n} e^{\frac{i}{\hbar} W(x, x')} \psi(x') dx'. \quad (4)$$

where  $m$  is the Maslov index.

The homomorphism  $Sp_2(2n, \mathbb{R}) \approx Mp(2n, \mathbb{R})$  allows us to associate to the curve  $t \mapsto f_t$  in  $Sp(2n, \mathbb{R})$  a unique curve  $t \mapsto F_t$  in  $Mp(2n, \mathbb{R})$  such that  $F_0$  is the identity operator. This means that if we introduce a projection  $\Pi^{Mp} : Mp(2n, \mathbb{R}) \rightarrow Sp(2n, \mathbb{R})$  then  $\Pi^{Mp}(F_t^W) = f_t^H$ .

Now choose a function  $\psi_0$  in some subspace of infinitely differentiable functions in  $L^2(\mathbb{R}^n)$  and set  $\psi_t = F_t \psi_0$ . We have  $\psi_t \in L^2(\mathbb{R}^n)$  for each  $t$  and the function  $\psi_t$  satisfies the Schrödinger-like equation

$$i \frac{\partial \psi_t}{\partial t} = \widehat{H} \psi_t \quad (5)$$

where  $\widehat{H}$  is a partial differential operator obtained from  $H$  using the Weyl ordering. That is, we perform the substitution  $p \rightarrow -i\nabla_x$  in the Hamiltonian (2) after having written it in the form

$$H(z, t) = \frac{1}{2} A(t) x \cdot x + B(t) p \cdot x + B^T(t) x \cdot p + \frac{1}{2} C(t) p \cdot p$$

where  $M(t) = \begin{pmatrix} A(t) & B(t) \\ B^T(t) & C(t) \end{pmatrix}$  with  $A(t) = A^T(t)$  and  $C(t) = C^T(t)$ . Using the obvious identity  $B\nabla_x \cdot x = x \cdot B\nabla_x + i \operatorname{Tr} B$  we thus obtain

$$\widehat{H} = \frac{1}{2}A(t)x \cdot x - ix \cdot B(t)\nabla_x + \frac{1}{2}C(t)\nabla_x \cdot \nabla_x - \frac{i}{2} \operatorname{Tr} B(t)$$

from which it is easy to verify that the operator  $\widehat{H}$  is symmetric (in fact self-adjoint) for the  $L^2$  scalar product. One formally writes

$$\widehat{H} = H(x, -i\nabla_x, t). \tag{6}$$

which is obtained from  $H$  by the substitution  $p_j \longrightarrow -i\hbar\partial/\partial x_j$ ; it is of course *formally* the usual “quantization” of the Hamiltonian function  $H$  of standard quantum mechanics.

It looks as if we have gained nothing by the process we have gone through, but this conclusion is not correct. We have shown that there is a deep relation between classical and quantum mechanics in this instance. It is the homomorphism  $Sp(2n, \mathbb{R}) \approx Mp(2n, \mathbb{R})$  that allows us to associate the classical flow  $f_t^H$  with the quantum flow  $F_t^W$ . In this sense we have been able to derive the Schrödinger equation from the classical equations of motion.

## 2.2 The Group $\operatorname{Ham}(2n, \mathbb{R})$

The group  $\operatorname{Ham}(2n, \mathbb{R})$  of Hamiltonian canonical transformations is the non-linear analogue of the symplectic group  $\operatorname{Sp}(2n, \mathbb{R})$ . What we would be looking for if we are to generalise the quadratic Hamiltonian result is a unitary representation of some covering structure of  $\operatorname{Ham}(2n)$ . This possibility has not been explored in the physical literature (to the best of the knowledge of the present authors). It is the purpose of this paper to show how these results can be extended to all physically relevant Hamiltonians. To this aim, the next two sections discuss the symplectic covariance of Hamiltonian mechanics and the corresponding symplectic covariance of Weyl quantisation. It is through these two notions that we can show how the Schrödinger equation is directly obtained from classical Hamiltonian mechanics.

## 3 Symplectic covariance of Hamiltonian mechanics

Let  $H$  now be an arbitrary Hamiltonian function, that is, a function  $H(x, p, t)$  which we assume to be at least once continuously differentiable in the variables  $x_1, \dots, x_n; p_1, \dots, p_n$  and  $t$ . The flow  $(f_t^H)$  determined by these equations

(one also says “generated by  $H$ ”) consists of the mappings  $f_t^H : \mathbb{R}^{2n} \longrightarrow \mathbb{R}^{2n}$  which associate to a point  $z = (x, p)$  at initial time  $t = 0$  the value  $z_t = (x_t, p_t)$  of the solutions at time  $t$  of the differential equations

$$\dot{x} = \nabla_p H(x, p, t) \quad , \quad \dot{p} = -\nabla_x H(x, p, t). \quad (7)$$

Thus, by definition,  $z_t = f_t^H(z_0)$ . Generically the flow is only defined in a neighborhood of  $t = 0$ ; we will later see that the consideration of compactly supported Hamiltonian functions eliminates these difficulties.

One of the problems we face is that for time dependent Hamiltonians, the corresponding flows  $f_t^{H(t)}$  are such that

$$f_t^{H(t)} f_t^{H(t')} \neq f_{t+t'}^{H(t+t')}$$

and therefore does not have the required group property. In subsection 2.3 we will show how a group  $\text{Ham}(2n, \mathbb{R})$  can be constructed. In the mean time we note that when  $H$  does not depend on time:  $\partial H / \partial t = 0$ , the system of differential equations (7) is autonomous and the flow  $(f_t^H)$  satisfies the one-parameter group property  $f_t^H f_{t'}^H = f_{t+t'}^H$  whenever the mappings  $f_t^H$ ,  $f_{t'}^H$ , and  $f_{t+t'}^H$  exist.

A fundamental property of Hamiltonian flows is that each  $f_t^H$  is a canonical transformation (or symplectomorphism). This means that for every  $z = (x, p)$  the Jacobian matrix

$$Df_t^H(z) = \frac{\partial z_t}{\partial z} = \frac{\partial(x_t, p_t)}{\partial(x, p)}$$

is symplectic, i.e.  $Df_t^H(z) \in \text{Sp}(2n, \mathbb{R})$ . (In some texts the term “canonical transformation” refers more generally to mappings that preserve the form of Hamilton’s equations; the definition we use here is thus more restrictive). The simplest proof of this property consists in showing that the matrix  $S_t = Df_t^H(z)$  satisfies the first order matrix equation

$$\frac{d}{dt} S_t = JH''(z_t, t) S_t \quad (8)$$

where  $H'' = D^2H$  is the Hessian matrix of the Hamiltonian, i.e. the matrix of its second derivatives in the variables  $x$  and  $p$  (see [5] §2.3.2 for the derivation of the equation (8), which is sometimes called the “variational equation” in the literature). To show that  $S_t$  is symplectic, we first set

$A_t = (S_t)^T J S_t$ . It follows from this equation that

$$\begin{aligned} \frac{dA_t}{dt} &= \frac{d(S_t)^T}{dt} J S_t + (S_t)^T J \frac{dS_t}{dt} \\ &= (S_t)^T H''(z_t) S_t - (S_t)^T H''(z_t) S_t \\ &= 0 \end{aligned}$$

hence  $A_t$  is constant. Thus  $A_t = A_0 = (S_0)^T J S_0 = J$  which shows that  $S_t \in \text{Sp}(2n, \mathbb{R})$ .

Let now  $s$  be an arbitrary symplectic matrix; by the chain rule, the Jacobian matrix of  $s f_t^H s^{-1}$  is  $s S_t s^{-1}$  hence it is also a canonical transformation; it is in fact the flow of a certain Hamiltonian. Let us now construct this Hamiltonian:

**Proposition 1** *The family of canonical transformations  $s f_t^H s^{-1}$  is the Hamiltonian flow determined by the function  $K(z, t) = H(s^{-1}z, t)$ . That is,*

$$s f_t^H s^{-1} = f_t^K, \quad K(z) = H(s^{-1}z). \quad (9)$$

**Proof.** Writing Hamilton equations for  $K$  as  $\dot{z} = J \nabla_z [H(s^{-1}z, t)]$  we get, using by the chain rule

$$\nabla_z [H(s^{-1}z)] = (s^T)^{-1} (\nabla_z H)(s^{-1}z)$$

hence, since  $s J s^T = J$  because  $s$  is symplectic,

$$\dot{z} = J (s^T)^{-1} (\nabla_z H)(s^{-1}z) = s J (\nabla_z H)(s^{-1}z).$$

Thus  $s^{-1}z(t)$  is the solution of Hamilton's equations for  $H$  with initial datum  $s^{-1}z(0)$ ; it follows that  $f_t^H(s^{-1}z) = s^{-1} f_t^K(z)$  which is equivalent to formula (9). ■

We will call the property just proved the *symplectic covariance property of Hamiltonian mechanics*. It will be generalized in the next subsection to the nonlinear case, where we study products of Hamiltonian flows. Formula (9) remains true if we replace  $s$  by an arbitrary symplectic transformation  $f$ . The result is actually well-known in standard Hamiltonian mechanics in the following form: set  $(x', p') = f(x, p)$  and  $K = H \circ f$ ; if  $f$  is a canonical transformation then we have the equivalence

$$\begin{aligned} \dot{x}' = \nabla_{p'} K(x', p', t) \quad \text{and} \quad \dot{p}' = -\nabla_{x'} K(x', p', t) \\ \iff \\ \dot{x} = \nabla_p H(x, p, t) \quad \text{and} \quad \dot{p} = -\nabla_x H(x, p, t). \end{aligned} \quad (10)$$

(see de Gosson [5], §2.3.2, for a proof). Let  $X_H$  the Hamilton vector field of  $H$ ; by definition

$$X_H(z, t) = (\nabla_p H(z, t), -\nabla_x H(z, t)).$$

Hamilton's equations are thus equivalent to  $\dot{z} = X_H(z, t)$ . Note that  $X_H(z, t)$  is not strictly speaking a vector field on  $\mathbb{R}^{2n}$  because of the dependence in  $t$ . In terms of  $X_H$  we can restate the theorem above as the transformation law

$$X_{H \circ f}(z) = [Df(z)]^{-1}(X_H \circ f)(z) \quad (11)$$

for Hamilton vector fields.

### 3.1 Operations on Hamiltonian Flows

We assume as before that all Hamiltonians have flows defined for all values of time  $t$ .

**Proposition 2** *Let  $(f_t^H)$  and  $(f_t^K)$  be Hamiltonian flows. Then:*

$$f_t^H f_t^K = f_t^{H \# K} \quad \text{with} \quad H \# K(z, t) = H(z, t) + K((f_t^H)^{-1}(z), t). \quad (12)$$

$$(f_t^H)^{-1} = f_t^K \quad \text{with} \quad K(z, t) = -H(f_t^H(z), t). \quad (13)$$

**Proof.** Let us first prove (12). Using successively the product and chain rules, we have

$$\begin{aligned} \frac{d}{dt}(f_t^H f_t^K) &= \left( \frac{d}{dt} f_t^H \right) f_t^K + (Df_t^H) f_t^K \frac{d}{dt} f_t^K \\ &= X_H(f_t^H f_t^K) + (Df_t^H) f_t^K \circ X_K(f_t^K) \end{aligned}$$

and it thus suffices to show that

$$(Df_t^H) f_t^K \circ X_K(f_t^K) = X_{K \circ (f_t^H)^{-1}}(f_t^K). \quad (14)$$

Writing

$$(Df_t^H) f_t^K \circ X_K(f_t^K) = (Df_t^H)((f_t^H)^{-1} f_t^H f_t^K) \circ X_K((f_t^H)^{-1} f_t^H f_t^K)$$

the equality (14) follows from the transformation formula (11) for Hamilton vector fields. Formula (13) is now an easy consequence of (12), noting that if  $K$  is given by the second formula (13) then  $(f_t^H f_t^K)$  is the flow determined by the Hamiltonian

$$K(z, t) = H(z, t) + K((f_t^H)^{-1}(z), t) = 0;$$

$f_t^H f_t^K$  is thus the identity, so that  $(f_t^H)^{-1} = f_t^K$  as claimed. ■

The formulae above show why we cannot avoid considering time-dependent Hamiltonians: even if  $H$  and  $K$  do not depend explicitly on time, the product and inverse of their flows is generated by time-dependent Hamiltonians! This remark will be important below when we define the group of Hamiltonian canonical transformations.

### 3.2 The group of Hamiltonian canonical transformations

From now on we assume that the solutions for the Hamilton equations with arbitrary initial data at time  $t = 0$  are uniquely determined by these data and exist for all times. This may seem to be a strong technical restriction, but it can actually easily be implemented (see e.g. Polterovich [18]). If  $H$  is such that  $f_t^H$  is not defined for all values of  $t$ , we just replace  $H$  by the function  $H\Theta$  where  $\Theta = \Theta(x, p)$  is a compactly supported infinitely differentiable function equal to one on some arbitrarily chosen subset  $\Omega$  of phase space.

Now proposition 2 allows us to define the group  $\text{Ham}(2n, \mathbb{R})$  of Hamiltonian canonical transformations. To do this, let us introduce the following terminology. We will say that a canonical transformation  $f$  of  $\mathbb{R}^{2n}$  is *Hamiltonian* if it is the time-one flow of some Hamiltonian function, i.e. if there exists a (time-dependent) Hamiltonian  $H$  such that we have  $f = f_{t=1}^H$ . Note the set of all Hamiltonian canonical transformation of  $\mathbb{R}^{2n}$  is denoted by  $\text{Ham}(2n, \mathbb{R})$ . The choice  $t = 1$  is arbitrary, and can be replaced by any other time  $t_0$ , as a consequence of the following simple observation: define the Hamiltonian function  $K(z, t) = t_0 H(z, t_0 t)$ ; the flow it determines is given by  $f_{t/t_0}^K = f_t^H$ , hence  $f_{t_0}^H = f_1^K$  is also a Hamiltonian canonical transformation.

It follows from Proposition 2 that  $\text{Ham}(2n, \mathbb{R})$  is indeed a group for the product operation: first, the product is clearly associative (because composition of mappings is). Assume that  $H$  is a constant; then  $f_t^H$  is the identity  $I$  on  $\mathbb{R}^{2n}$  and hence  $I$  is in  $\text{Ham}(2n, \mathbb{R})$ . Suppose now  $f$  and  $g$  are two Hamiltonian canonical transformations:  $f = f_{t=1}^H$  and  $g = f_{t=1}^K$ . In view of formula (12) we have  $fg = f_{t=1}^{H\#K}$  hence  $fg$  is also a Hamiltonian canonical transformation. That the inverse  $f^{-1}$  also is in  $\text{Ham}(2n, \mathbb{R})$  follows similarly from formula (13).

### 3.3 The key role of the Banyaga Theorem.

The group  $\text{Ham}(2n, \mathbb{R})$  contains the symplectic group  $\text{Sp}(2n, \mathbb{R})$  as a subgroup. Equivalently and more important for us here: *every symplectic matrix is the time-one flow of some Hamiltonian function.* This is obvious if  $s$  is in the range of the exponential  $\exp : \mathfrak{sp}(2n, \mathbb{R}) \rightarrow \text{Sp}(2n, \mathbb{R})$ , writing  $s = \exp(X)$ : the formula  $s_t = \exp(tX)$  then defines a one-parameter group of symplectic matrices, which is the propagator for the Hamilton equations for  $H(z) = -\frac{1}{2}JXz \cdot z$ . The general case easily follows by induction since every symplectic matrix can be written as a product  $\exp(X_1) \cdots \exp(X_k)$  where  $X_1, \dots, X_k$  are in the Lie algebra  $\mathfrak{sp}(2n, \mathbb{R})$ . There is actually a more constructive way to prove this result as a particular case of the following far-reaching property of  $\text{Ham}(2n, \mathbb{R})$ , proven in full generality by Banyaga [2] in the context of symplectic manifolds:

**Theorem 3 (Banyaga)** *Let  $(f_t)$  be a family of Hamiltonian canonical transformations depending smoothly on the parameter  $t$ , and such that  $f_0 = I$  (the identity). Then  $f_t = f_t^H$  for some Hamiltonian function  $H(z, t)$  given by the formula*

$$H(z, t) = H(z, 0) - \int_0^1 \sigma(X(uz, t), z) du \quad (15)$$

where  $X = \left(\frac{d}{dt}f_t\right) f_t^{-1}$  and  $\sigma$  is the symplectic form  $\sigma(z, z') = (z')^T Jz = Jz \cdot z'$  on the phase space  $\mathbb{R}^{2n}$ .

Let us sketch the proof (for a complete proof see Polterovich [18] or Banyaga's original paper [2]). One begins by noting that if  $X_H$  is a Hamiltonian vector field, one can reconstruct  $H$  by the following method: first write

$$\begin{aligned} H(z, t) &= H(z, 0) + \int_0^1 \frac{d}{du} X_H(uz, t) du \\ &= H(z, 0) + \int_0^1 [\nabla_z H(uz, t) \cdot z] du \end{aligned}$$

(the second equality in view of the chain rule). Next observe that since  $\nabla_z H(uz, t) = -J^2 \nabla_z H(uz, t) = -JX_H(uz, t)$ , we have

$$H(z, t) = H(z, 0) - \int_0^1 \sigma(X_H(uz, t), z) du$$

where  $\sigma$  is the standard symplectic form. One then proves the proposition by setting  $X = \left(\frac{d}{dt}f_t\right) f_t^{-1}$  and showing that the flow determined by the Hamiltonian functions is precisely  $(f_t)$ .

This result is remarkable, and even in some sense surprising, because it shows that the datum of a family of time-one Hamiltonian canonical transformations, each coming from a different Hamiltonian function, is itself the flow of some Hamiltonian.

This has several non-trivial consequences. First, it shows that the group  $\text{Ham}(2n, \mathbb{R})$  is arcwise connected; secondly it immediately implies that  $\text{Sp}(2n, \mathbb{R}) \subset \text{Ham}(2n, \mathbb{R})$ : let  $s \in \text{Sp}(2n, \mathbb{R})$  and choose a path  $(s_t)$  joining the identity to  $s$  in  $\text{Sp}(2n, \mathbb{R})$ . Then  $X = \left(\frac{d}{dt}s_t\right) s_t^{-1}$  belongs to the symplectic Lie algebra  $\mathfrak{sp}(2n, \mathbb{R})$ ; in particular it is a matrix so that formula (15) defines a quadratic Hamiltonian whose associated flow is precisely  $(s_t)$ .

### 3.4 Reduction to the time-independent case: the extended phase space

It turns out that the study of time-dependent Hamiltonian function can always be reduced to that of Hamiltonians which do not explicitly contain time. As we will see, the price we have to pay for this is that we have to work in a phase space with dimension  $n + 2$  instead of  $n$ . Thus, in a sense, there are no time-dependent Hamiltonians!

The trick is the following, which seems to go back to [25, 11] (see [24] for an analysis in depth). We define a new Hamiltonian function  $\tilde{H}$  by the formula

$$\tilde{H}(x, p, t, E) = H(x, p, t) - E \tag{16}$$

where  $E$  is a new variable, viewed as conjugate to the time  $t$  (the latter now has the status of a “position variable”; we could as well write the definition of  $\tilde{H}$  in the form

$$\tilde{H}(x, p, x_{n+1}, p_{n+1}) = H(x, p, x_{n+1}) - p_{n+1}$$

but we will however stick to the notation (16) if only for economy. The function  $\tilde{H}$  is defined on the extended phase space  $\mathbb{R}^{2n+2} \equiv \mathbb{R}_{x,p}^{2n} \times \mathbb{R}_E \times \mathbb{R}_t$  and the associated Hamilton equations are, expressed in terms of the original Hamiltonian  $H$ :

$$\begin{aligned} \frac{dx}{dt'} &= \nabla_p H \quad , \quad \frac{dp}{dt'} = -\nabla_x H \\ \frac{dE}{dt'} &= \frac{\partial H}{\partial t} \quad , \quad \frac{\partial t}{\partial t'} = 1 \end{aligned}$$

where the parameter  $t'$  plays the role of a new “time”; since  $\tilde{H}$  does not contain explicitly that parameter it is indeed a “time-independent” Hamiltonian

on extended phase space. Notice that in view of the fourth equation above we may choose  $t' = t$  so that the two first equations are just the Hamiltonian equations for  $H$ ; as a bonus the third equation is just the familiar law for the variation of energy of a time-dependent Hamiltonian system:

$$\frac{dE}{dt} = \frac{d}{dt}H(x, p, t) = \frac{\partial}{\partial t}H(x, p, t)$$

(the second equality because of the chain rule and using the fact that  $x$  and  $p$  satisfy Hamilton's equations).

We can now define the “extended Hamiltonian flow” ( $\tilde{f}_t^H$ ) of  $H$  by the formula  $\tilde{f}_t^H = f_t^{\tilde{H}}$ . Notice that since  $(f_t^{\tilde{H}})$  is the flow determined by a time-independent Hamiltonian ( $\tilde{f}_t^H$ ), it enjoys the one-parameter group property  $\tilde{f}_t^H \tilde{f}_{t'}^H = \tilde{f}_{t+t'}^H$  and  $\tilde{f}_0^H = I$  (the identity operator on the extended phase space  $\mathbb{R}^{2n+2}$ ).

Denote now by  $(f_{t,t'}^H)$  the two-parameter family of canonical transformations of  $\mathbb{R}^{2n}$  defined as follows: for fixed  $t'$  the function  $z = f_{t,t'}^H(z')$  is the solution of Hamilton's equations for  $H$  taking the value  $z'$  at time  $t'$ . Clearly  $f_{t,t}^H$  is the identity operator on  $\mathbb{R}^{2n}$  and  $f_{t,t'}^H f_{t',t''}^H = f_{t,t''}^H$ ,  $(f_{t,t'}^H)^{-1} = f_{t',t}^H$ . The two-parameter family  $(f_{t,t'}^H)$  is sometimes called the “time-dependent flow”; it is related to the extended flow defined above by the simple (and useful) formula

$$\tilde{f}_t^H(z', t', E') = (f_{t,t'}^H(z'), t + t', E_{t,t'}) \quad (17)$$

with

$$E_{t,t'} = E' + H(f_{t,t'}^H(z'), t) - H(z', t'). \quad (18)$$

We will see below that the procedure above extends, suitably modified, to the case of the Schrödinger equation.

## 4 The Weyl Formalism and Schrödinger's Equation

In classical physics, observables are taken to be functions on phase space, limited by whatever conditions of regularity that are required by the physics of the problem. A class of observables having been chosen, a state is a continuous functional on this class (the continuity being imposed for some specified topology). We will cast a wide net and accept as observable not only functions, but any tempered distribution on the configuration space  $\mathbb{R}^d$ ; the space of such distributions is traditionally denoted by  $\mathcal{S}'(\mathbb{R}^n)$ . It is the dual of the Schwartz space  $\mathcal{S}(\mathbb{R}^n)$  of all complex functions defined on

$\mathbb{R}^n$  which decrease, together with their derivatives, faster than the inverse of any polynomial in the variables  $x_1, \dots, x_n$ .

An observable will usually be denoted by a small Latin letter, such that  $a, b, \dots$ ; tradition provides us with the exception confirming the rule: when an observable is viewed as a Hamiltonian function it will always be denoted by  $H, K$ , etc.

#### 4.1 Pseudodifferential operators

We begin by briefly discussing some well-known aspects from the theory of pseudodifferential operators (see for instance Shubin [21]). Let  $A$  be a linear operator  $\mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$ . If we assume that  $A$  is continuous from  $\mathcal{S}(\mathbb{R}^n)$  to  $\mathcal{S}'(\mathbb{R}^n)$  (when  $\mathcal{S}'(\mathbb{R}^n)$  is equipped with the weak-\* topology, that is the topology of pointwise convergence) then a deep result from functional analysis, Schwartz's kernel theorem, tells us that there exists a distribution  $K_A$  in  $\mathcal{S}'(\mathbb{R}^n \times \mathbb{R}^n)$  such that for all  $\psi, \phi$  in  $\mathcal{S}(\mathbb{R}^n)$ , we have  $\langle A\psi, \phi \rangle = \langle \langle K_A, \psi \otimes \phi^* \rangle \rangle$  where  $\langle \cdot, \cdot \rangle$  and  $\langle \langle \cdot, \cdot \rangle \rangle$  are the distributional brackets for  $\mathcal{S}'(\mathbb{R}^n)$  and  $\mathcal{S}'(\mathbb{R}^n \times \mathbb{R}^n)$ , respectively (see Gröchenig [8], §14.3, for an elegant proof of the kernel theorem in the broader context of Feichtinger's modulation spaces). We will informally write

$$A\psi(x) = \int_{\mathbb{R}^n} K_A(x, y)\psi(y)dy \quad (19)$$

where the integral should be interpreted in the distributional sense.

Let now  $\tau$  be an arbitrary fixed real number and define a distribution  $a_\tau(x, p)$  by the formula

$$a_\tau(x, p) = \int_{\mathbb{R}^n} e^{-\frac{i}{\hbar}p \cdot y} K_A(x + \tau y, x - (1 - \tau)y)dy; \quad (20)$$

where we have introduced a scaling factor which for convenience, we have called  $\hbar$ . This ensures that the exponential is dimensionless. The magnitude of this factor is left open in this paper since we are concentrating only on mathematical relationships.

Using the Fourier inversion formula, the kernel is then expressed in terms of  $a$  by

$$K_A(x, y) = \left(\frac{1}{2\pi\hbar}\right)^n \int_{\mathbb{R}^n} e^{\frac{i}{\hbar}p \cdot (x-y)} a_\tau((1 - \tau)x + \tau y, p)dp. \quad (21)$$

It follows that formula (19) can be rewritten in pseudodifferential form as

$$A\psi(x) = \left(\frac{1}{2\pi\hbar}\right)^n \iint_{\mathbb{R}^n \times \mathbb{R}^n} e^{\frac{i}{\hbar}p \cdot (x-y)} a_\tau((1 - \tau)x + \tau y, p)\psi(y)dydp \quad (22)$$

Thus we have shown how to associate an operator  $A$  with an observable  $a_\tau$ .

Now we can use formula (22) to associate to given observable  $a$  with an operator  $A_\tau$  for every value of the constant  $\tau$ : by definition

$$A_\tau \psi(x) = \left(\frac{1}{2\pi\hbar}\right)^n \iint_{\mathbb{R}^n \times \mathbb{R}^n} e^{\frac{i}{\hbar}p \cdot (x-y)} a((1-\tau)x + \tau y, p) \psi(y) dy dp \quad (23)$$

the “double integral” being again interpreted in some reasonable sense (for instance as a repeated distributional bracket; it is actually no restriction in our discussion to assume that the observable is in  $\mathcal{S}(\mathbb{R}^{2n})$  in which case the right hand side of (23) strictly makes sense since the integral is absolutely convergent).

Let us dignify the discussion above in compact form as a Theorem:

**Theorem 4** *Every continuous linear operator  $A : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}'(\mathbb{R}^n)$  can be written in pseudodifferential form (22), and this for every real value of the parameter  $\tau$ .*

The (admittedly somewhat technical) discussion above shows that for each value of  $\tau$  we have a bijective (= one-to-one and onto) correspondence  $a \longleftrightarrow A_\tau$  between observables and continuous operators  $\mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}'(\mathbb{R}^n)$  given by (23). We emphasize again that each choice of  $\tau$  is a priori equally good; for instance if we take  $\tau = 0$  then we get

$$\begin{aligned} A_1 \psi(x) &= \left(\frac{1}{2\pi\hbar}\right)^n \iint_{\mathbb{R}^n \times \mathbb{R}^n} e^{\frac{i}{\hbar}p \cdot (x-y)} a(x, p) \psi(y) dy dp \\ &= \left(\frac{1}{2\pi\hbar}\right)^n \int_{\mathbb{R}^n} e^{\frac{i}{\hbar}p \cdot x} a(x, p) F\psi(p) dy \end{aligned}$$

( $F\psi$  the  $\hbar$ -Fourier transform) which is the standard pseudodifferential representation widely used in partial differential equations. There is however a privileged choice in quantum mechanics: it consists in taking  $\tau = \frac{1}{2}$ . The corresponding operator  $A_{1/2}$  will be denoted  $\widehat{A}$  and called the Weyl operator corresponding to the observable (or symbol)  $a$ . Explicitly

$$\widehat{A}\psi(x) = \left(\frac{1}{2\pi\hbar}\right)^n \iint_{\mathbb{R}^n \times \mathbb{R}^n} e^{\frac{i}{\hbar}p \cdot (x-y)} a\left(\frac{1}{2}(x+y), p\right) \psi(y) dy dp$$

and a few elementary manipulations show that this equivalent to the following expression

$$\widehat{A}\psi(x) = \left(\frac{1}{2\pi\hbar}\right)^n \int_{\mathbb{R}^n} a_\sigma(z') \widehat{T}(z') \psi(x) dz'$$

which is more familiar to physicist; here

$$a_\sigma(x', p') = \left(\frac{1}{2\pi\hbar}\right)^n \int_{\mathbb{R}^{2n}} e^{-\frac{i}{\hbar}\sigma(z', z'')} a(z'') dz''$$

(the symplectic Fourier transform of  $a$ ) and  $\widehat{T}(z')$  is the Heisenberg–Weyl operator:

$$\widehat{T}(z')\psi(x) = e^{\frac{i}{\hbar}(p'\cdot x - \frac{1}{2}p'\cdot x')}\psi(x - x');$$

see Littlejohn [12] for an analysis of these operators of which we have given a detailed study in [5].

The Weyl correspondence associates in a unique way to every observable  $a \in \mathcal{S}'(\mathbb{R}^n)$  a continuous linear operator  $\widehat{A}$  whose domain always contains  $\mathcal{S}(\mathbb{R}^n)$  and whose range is contained in  $\mathcal{S}'(\mathbb{R}^n)$ . This correspondence actually is not only one-to-one, but also onto: every continuous linear operator  $\widehat{A} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}'(\mathbb{R}^n)$  is associated to an observable  $a$  via the Weyl correspondence. We will symbolically write the Weyl correspondence between an observable  $a$  and the corresponding operator  $\widehat{A}$  is  $a \overset{\text{Weyl}}{\longleftrightarrow} \widehat{A}$  or  $\widehat{A} \overset{\text{Weyl}}{\longleftrightarrow} a$ . (In the mathematical literature  $a$  is called the *Weyl symbol* of  $\widehat{A}$ ).

The Weyl correspondence has the following very agreeable feature, which distinguishes it from other  $\tau$ -correspondences, and which already motivated Weyl’s interest: the transition to the adjoint operator is particularly simple. If  $a \overset{\text{Weyl}}{\longleftrightarrow} \widehat{A}$  then  $\bar{a} \overset{\text{Weyl}}{\longleftrightarrow} \widehat{A}^*$ . In particular, if  $a$  is a real observable then  $\widehat{A}$  is formally self-adjoint. Thus the Weyl correspondence associates to real observables formally self-adjoint operators. This is one of the reasons why the “Weyl ordering” was adopted at an early stage (but more or less explicitly) in quantum mechanics. The second reason is that Weyl operators are characterized by a symplectic covariance property, an important characterisation for the discussions in this paper. We will now discuss the details of this symplectic covariance.

## 4.2 Symplectic covariance of Weyl calculus

The fundamental result, which is at the very heart of our derivation of Schrödinger’s equation in next subsection is the following:

**Theorem 5** *Let  $\mathcal{L}(\mathcal{S}(\mathbb{R}^n), \mathcal{S}'(\mathbb{R}^n))$  be the space of all continuous linear mappings  $\mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}'(\mathbb{R}^n)$ . Let*

$$\mathcal{M} : \mathcal{S}'(\mathbb{R}^{2n}) \longrightarrow \mathcal{L}(\mathcal{S}(\mathbb{R}^n), \mathcal{S}'(\mathbb{R}^n))$$

be a continuous mapping associating to each observable  $a$  a continuous operator  $\mathcal{M}(a)$ . If  $\mathcal{M}$  has the property that for every  $s$  in  $\mathrm{Sp}(2n, \mathbb{R})$  we have  $\mathcal{M}(a \circ s^{-1}) = S\mathcal{M}(a)S^{-1}$  where  $S$  in  $\mathrm{Mp}(2n, \mathbb{R})$  is such that  $\pi^{\mathrm{Mp}}(S) = s$  then  $\mathcal{M}$  is the Weyl correspondence, that is  $a \xleftrightarrow{\mathrm{Weyl}} \mathcal{M}(a)$ .

Note that if  $a \xleftrightarrow{\mathrm{Weyl}} \widehat{A}$  then we have  $a \circ s^{-1} \xleftrightarrow{\mathrm{Weyl}} S\widehat{A}S^{-1}$ : this is the well-known property of symplectic covariance of Weyl calculus; we have given a detailed proof of this property in [5], §7.1.3. There is also a proof in Littlejohn [12].

The real impact of the theorem above is that it says that, conversely, among all possible correspondences  $a \longleftrightarrow A_\tau$  between observables and continuous operators (23), the Weyl correspondence is the *only* one enjoying the property of symplectic covariance. The proof of this result is alluded to in Stein's book [22] (pp. 578–579) and proven in detail in the last Chapter of Wong's book [26]. Wong's proof actually relies on the following additional assumption (which is also found in Stein's statement): the mapping  $\mathcal{M}$  should associate to every essentially bounded observable  $a$  which depends only on  $x$ , the operation of multiplication by  $a$  viz: if  $a = a(x)$  then  $Q(a)\psi = a\psi$ . However this property is automatically satisfied by the operators  $A_\tau$  and thus superfluous in our context: we have

$$\begin{aligned} A_\tau\psi(x) &= \left(\frac{1}{2\pi\hbar}\right)^n \iint_{\mathbb{R}^n \times \mathbb{R}^n} e^{\frac{i}{\hbar}p \cdot (x-y)} a((1-\tau)x + \tau y) \psi(y) dy dp \\ &= \left(\frac{1}{2\pi\hbar}\right)^n \int_{\mathbb{R}^n} \left[ \int_{\mathbb{R}^n} e^{\frac{i}{\hbar}p \cdot (x-y)} dp \right] a((1-\tau)x + \tau y) \psi(y) dy \\ &= \int_{\mathbb{R}^n} \delta(x-y) a((1-\tau)x + \tau y) \psi(y) dy \\ &= a(x) \psi(x). \end{aligned}$$

## 5 Derivation of Schrödinger's equation

### 5.1 Statement of the main result

Now we come to the main result in the paper, namely, to show that the Schrödinger equation can be derived from Hamilton's equations of motion by lifting the flows  $F_t^H$  of  $\mathrm{Ham}(2n, \mathbb{R})$  onto a unitary representation of a covering structure.

In order to motivate our approach let us now briefly return to the property of the metaplectic representation of  $\mathrm{Sp}(2n, \mathbb{R})$  which shows that to every

family  $(s_t)$  of symplectic matrices depending smoothly on  $t$  and such that  $s_0 = I$ , we can associate, in a unique way, a family  $(S_t)$  of unitary operators on  $L^2(\mathbb{R}^n)$  belonging to the metaplectic group  $\text{Mp}(2n, \mathbb{R})$  such that  $S_0 = 1$  and  $S_t S_{t'} = S_{t+t'}$ .

We have seen from Banyaga's theorem, that for every flow  $(s_t)$  there exists some Hamiltonian  $H$ , so that we can write  $(s_t) = (f_t^H)$ . Now if we use the correspondence  $H \xleftrightarrow{\text{Weyl}} \widehat{H}$  we have  $(S_t) = (F_t^H)$  where

$$S_t = e^{-i\widehat{H}t/\hbar}$$

which is the solution of the Schrödinger equation

$$i\hbar \frac{d}{dt} S_t = \widehat{H} S_t$$

Here we have again introduced a scaling parameter having the dimensions of action which, for convenience, we have denoted by  $\hbar$ . Note that generically  $\widehat{H}$  is not a bounded operator on  $L^2(\mathbb{R}^n)$  so that the exponential has to be defined using some functional calculus (see Reed and Simon [19] §VIII.3 for a discussion of these technicalities).

We now ask whether this property has an analogue for paths in the group  $\text{Ham}(2n, \mathbb{R})$  of Hamiltonian canonical transformations. That is can we find a  $(F_t^H)$  corresponding to  $(f_t^H)$  for every physically relevant Hamiltonian. To sharpen up the discussion, let us introduce the following notation:

- Denote by  $\mathcal{P}\text{Ham}(2n, \mathbb{R})$  the set of all one-parameter families  $(f_t)$  in  $\text{Ham}(2n, \mathbb{R})$  depending smoothly on  $t$  and passing through the identity at time  $t = 0$ . As we have remarked above, such a family of canonical transformations is always the flow  $(f_t^H)$  of some (usually time-dependent) Hamiltonian  $H$  (Banyaga's theorem).
- Denote by  $\mathcal{P}U(L^2(\mathbb{R}^n))$  the set of all strongly continuous one-parameter families  $(F_t)$  of unitary operators on  $L^2(\mathbb{R}^n)$  depending smoothly on  $t$  and such that  $F_0$  is the identity operator, and having the following property: the domain of the infinitesimal generator  $\widehat{H}$  of  $(F_t)$  contains the Schwartz space  $\mathcal{S}(\mathbb{R}^n)$ .

Recall that strong continuity for a one-parameter group  $(F_t)$  means that we have

$$\lim_{t \rightarrow t_0} F_t \psi = F_{t_0}^H \psi \tag{24}$$

and the infinitesimal generator of  $(F_t)$  is the operator defined by

$$A = i\hbar \left. \frac{d}{dt} F_t \psi \right|_{t=0} = i\hbar \lim_{\Delta t \rightarrow 0} \frac{F_{\Delta t} \psi - \psi}{\Delta t} \quad (25)$$

for every  $\psi \in L^2(\mathbb{R}^n)$  and every real number  $t_0$ ; formally  $F_t = e^{-i\hbar A/t}$ . Formally we have Stone's theorem [23]:

**Theorem 6 (Stone)** *For every strongly continuous one-parameter group  $(F_t)$  of unitary operators on a Hilbert space  $\mathcal{H}$  there exists a self-adjoint operator  $A$  on  $L^2(\mathbb{R}^n)$  such that  $F_t = e^{itA/\hbar}$ ; in particular  $A$  is closed and densely defined in  $\mathcal{H}$ . Conversely, if  $A$  is a self-adjoint operator on  $\mathcal{H}$  then there exists a unique one-parameter unitary group  $(F_t)$  whose infinitesimal generator is  $A$ , that is  $F_t = e^{itA/\hbar}$ .*

For self-contained proofs we refer Reed and Simon [19], §VIII.4, or to Abraham et al. [1], Supplement 7.4B, pp.529–535.

We will now use Stone's theorem to prove the hard part of our main result:

**Theorem 7** *There exists a bijective correspondence*

$$\mathcal{C} : \mathcal{P}\mathcal{U}(L^2(\mathbb{R}^n)) \leftrightarrow \mathcal{P}\text{Ham}(2n, \mathbb{R})$$

*whose restriction to families  $(s_t)$  of symplectic matrices reduces to the metaplectic representation, and which has the following symplectic covariance property: for every  $(f_t)$  in  $\mathcal{P}\text{Ham}(2n, \mathbb{R})$  and for every  $s$  in  $\text{Sp}(2n, \mathbb{R})$  we have*

$$\mathcal{C}(s f_t s^{-1}) = (S F_t S^{-1}) \quad (26)$$

*where  $S$  is any of the two operators in  $\text{Mp}(2n, \mathbb{R})$  such that  $\pi^{\text{Mp}}(S) = s$ . This correspondence  $\mathcal{C}$  is bijective and we have*

$$i\hbar \frac{d}{dt} F_t = \widehat{H} F_t. \quad (27)$$

*where  $\widehat{H} \xrightarrow{\text{Weyl}} H$ , the Hamiltonian function  $H$  being determined by  $(f_t)$ .*

It is perhaps worth observing that it is always preferable to take the family  $(F_t)$  as the fundamental object, rather than  $\widehat{H}$  (and hence Schrödinger's equation). This remark has already been made by Weyl who noticed that  $(F_t)$  is everywhere defined and consists of bounded operators, while  $\widehat{H}$  is generically unbounded and only densely defined (see the discussion in Mackey [15] for a discussion of related questions).

## 5.2 Proof: the time-independent case

Let us begin with the case where  $(f_t)$  is the Hamiltonian flow determined by a time-independent Hamiltonian function  $H = H(x, p)$ , in which case  $(f_t) = (f_t^H)$  is a one-parameter group, that is  $f_t^H f_{t'}^H = f_{t+t'}^H$ . We thus want to associate to  $(f_t^H)$  a strongly continuous one-parameter group  $(F_t) = (F_t^H)$  of unitary operators on  $L^2(\mathbb{R}^n)$  satisfying some additional conditions. We proceed as follows: let  $\widehat{H}$  be the operator associated to  $H$  by the Weyl correspondence:  $\widehat{H} \xleftrightarrow{\text{Weyl}} H$  and define  $\mathcal{C}(f_t^H) = (F_t)$  by  $F_t = e^{-it\widehat{H}/\hbar}$ . The Weyl operator  $\widehat{H}$  is self-adjoint and its domain obviously contains  $\mathcal{S}(\mathbb{R}^n)$ . Let us show that the covariance property (26) holds. We have

$$\mathcal{C}(s f_t^H s^{-1}) = \mathcal{C}(f_t^{H \circ s^{-1}})$$

in view of formula (9) in Proposition 1, that is, by definition of  $\mathcal{C}$ ,

$$\mathcal{C}(s f_t^H s^{-1}) = (e^{-it\widehat{H \circ s^{-1}}/\hbar}).$$

In view of the symplectic covariance property  $a \circ s^{-1} \xleftrightarrow{\text{Weyl}} S\widehat{A}S^{-1}$  of Weyl operators we have  $\widehat{H \circ s^{-1}} = S\widehat{H}S^{-1}$ , and hence

$$\mathcal{C}(s f_t^H s^{-1}) = (e^{-itS\widehat{H}S^{-1}/\hbar}) = (S e^{-it\widehat{H}/\hbar} S^{-1})$$

which is property (26).

Let conversely  $(F_t)$  be in  $\mathcal{P}U(L^2(\mathbb{R}^n))$ ; we must show that we can find a unique  $(f_t)$  in  $\mathcal{P}\text{Ham}(2n, \mathbb{R})$  such that  $\mathcal{C}(f_t) = (F_t)$ . By Stone's theorem and our definition of  $\mathcal{P}U(L^2(\mathbb{R}^n))$  there exists a unique self-adjoint operator  $A$ , densely defined, and whose domain contains  $\mathcal{S}(\mathbb{R}^n)$ . Thus  $A$  is continuous on  $\mathcal{S}(\mathbb{R}^n)$  and in view of Theorem 4 for each value of the parameter  $\tau$  there exists an observable  $a$  such that  $A \longleftrightarrow a_\tau$ . Choose  $\tau = \frac{1}{2}$ ; then  $A = \widehat{H} \xleftrightarrow{\text{Weyl}} H$  for some function  $H = H(x, p)$  and we have  $\mathcal{C}(f_t^H) = (F_t)$ .

There remains to show that the correspondence  $\mathcal{C}$  restricts to the metaplectic representation for semigroups  $(f_t) = (s_t)$  in  $\text{Sp}(2n, \mathbb{R})$ ; but this is clear since  $(s_t)$  is generated, as a flow, by a quadratic Hamiltonian, and that the unitary one-parameter group of operators determined by such a function precisely consists of metaplectic operators.

## 5.3 Proof: the general case

We now no longer assume that  $(f_t)$  and  $(F_t)$  are one-parameter groups. Recall that we have reduced the study of a time-dependent Hamiltonian

$H = H(z, t)$  to by introducing

$$\widetilde{H}(x, p, t, E) = H(x, p, t) - E \quad (28)$$

which is a time-independent Hamiltonian on  $\mathbb{R}^{2n+2} \equiv \mathbb{R}_{x,p}^{2n} \times \mathbb{R}_E \times \mathbb{R}_t$  where  $E$  is viewed as conjugate variable to  $t$ . The flow  $(\widetilde{f}_t^H) = (f_t^{\widetilde{H}})$  on  $\mathbb{R}^{2n+2}$  generated by  $\widetilde{H}$  is related to the time-dependent flow  $(f_{t,t'}^H)$  by the formula

$$\widetilde{f}_t^H(z', t', E') = (f_{t,t'}^H(z'), t + t', E_{t,t'}) \quad (29)$$

where  $E_{t,t'} - E'$  is the variation of the energy in the time interval  $[t', t]$ . The advantage of this reformulation of the dynamics associated with  $H$  is that  $(\widetilde{f}_t^H)$  is a one-parameter group of canonical transformations of  $\mathbb{R}^{2n+2}$ . In the operator case we can proceed in a quite similar way, noting that the Weyl operator associated with  $\widetilde{H}$  is given by

$$\widehat{\widetilde{H}} = \widehat{H}(x, -i\hbar\nabla_x, t) - i\hbar\frac{\partial}{\partial t}.$$

Of course  $\widehat{\widetilde{H}}$  is self-adjoint if and only if  $\widehat{H}$  is, which is the case since  $H$  is real. We will need the following elementary fact, which is a variant of the method of separation of variables:

**Lemma 8** *Let  $E$  be an arbitrary real number. The function*

$$\Psi(x, t; t') = \psi(x, t)e^{\frac{i}{\hbar}E(t-t')} \quad (30)$$

*is a solution of the extended Schrödinger equation*

$$i\hbar\frac{\partial\Psi}{\partial t'} = \widehat{\widetilde{H}}\Psi \quad (31)$$

*if and only if  $\psi = \psi(x, t)$  is a solution of the usual Schrödinger equation*

$$i\hbar\frac{\partial\psi}{\partial t} = \widehat{H}\psi. \quad (32)$$

**Proof.** We first note the obvious identity

$$i\hbar\frac{\partial\Psi}{\partial t'} = E\Psi. \quad (33)$$

Writing for short  $\widehat{H}(t) = \widehat{H}(x, -i\hbar\nabla_x, t)$  we have, after a few calculations

$$\left(\widehat{H}(t) - i\hbar\frac{\partial}{\partial t}\right)\Psi = \left[\widehat{H}(t)\psi - i\hbar\frac{\partial\psi}{\partial t}\right]e^{\frac{i}{\hbar}E(t-t')} + E\Psi \quad (34)$$

hence (31) is equivalent to (32) in view of (33). ■

This result shows the following: choose an initial function  $\psi_0 = \psi_0(x)$  at time  $t = 0$  and solve the usual Schrödinger equation (32), which yields the solution  $\psi = \psi(x, t)$ . Then  $\Psi = \Psi(x, t; t')$  defined by (30) is the solution of the extended Schrödinger equation (31) with initial datum  $\Psi(x, t; t) = \psi(x, t)$  at time  $t' = t$ . In terms of flows we can rewrite this as

$$\tilde{F}_{t'-t}(F_t\psi_0) = (F_t\psi_0)e^{\frac{i}{\hbar}E(t-t')}$$

## 6 Discussion

### 6.1 Why Planck's constant?

We have shown that the mathematical formalism of the theory of Schrödinger's equation is already present in classical mechanics, and is in fact a reformulation of Hamiltonian dynamics in terms of operators. So where does quantum *physics* enter the scene? The most obvious question is of course how to give a physical meaning to the constant  $\hbar$ . That is: *Why do we need Planck's constant?* Let us explore a few possibilities; the list is certainly not exhaustive, and the choice has been done in accordance with the present author's tastes (and knowledge!).

**Spectral properties of operators** In a sense the most obvious and naive way to give a physical motivation for the need of Planck's constant is, no doubt, empirical. We have seen that there is a one-to-one correspondence between Hamiltonian flows and unitary evolution group; the following are solutions of Schrödinger equations. Now, the spectral properties of the involved Hamiltonian operators are well-understood; a basic postulate of traditional quantum mechanics is that the eigenvalues of an operator are the values that the corresponding observable can take. Using this postulate one can thus put in the right value of the constant  $\hbar$  by hand. This is an empirical motivation, but it seems hard to refute because we know that it works! It is irrefutable from an epistemological point of view: putting in  $\hbar$  that way we are sure to get the right physical answers!

**The Narcowich–Wigner spectrum of a mixed state** Consider a density matrix  $\hat{\rho}$ : by definition it is a self adjoint non-negative trace class operator with trace unity:  $\hat{\rho} = \hat{\rho}^*$ ,  $\hat{\rho} \geq 0$ ,  $\text{Tr } \hat{\rho} = 1$ . As is well-known from operator theory it is the positivity property  $\hat{\rho} \geq 0$  which is generally the most difficult to check; one explicit tool is provided by the Kastler–Loupias–Miracle Sole

(KLM) conditions [10, 13, 14], but these imply the simultaneous verification of infinitely many conditions (mathematically the KLM conditions are a symplectic variant of Bochner’s positivity criteria for the Fourier transform of a measure). Narcowich [16] and Narcowich and O’Connell [17] have noticed the following property of mixed quantum states: the positivity property depends in a crucial way on the numerical value of  $\hbar$ . That is, if we change  $\hbar$  some classical states become “quantum” (that is, representable by a density matrix), and some quantum states lose this property (their density matrix is no longer non-negative). The set of all values of  $\hbar$  for which  $\hat{\rho}$  remains positive is then called the Narcowich–Wigner spectrum of  $\hat{\rho}$ .

**The uncertainty principle** We mention, in passing, that there is a related property which shows that  $\{z_1, z_2, \dots, z_k\}$  and  $\{z'_1, z'_2, \dots, z'_k\}$  be two arbitrary sets of  $k$  distinct points in phase space. We can always find a Hamiltonian motion taking the first set to the other in some finite time  $t$ . This result (“Boothby’s theorem”; for a proof see Banyaga [3], p. 125, Lemma 5.1.1) shows that spreading is by no means a typical quantum phenomenon since the points  $z_j$  can be clustered in a small region of phase space, while the points  $z'_j$  can be disseminated through a large region (that spreading is not characteristic of quantum theory has been known for a long time; see e.g. Littlejohn’s seminal paper [12]).

**Information theory** Perhaps, after all, the answer ultimately lies in information theory. It might very well be that the discrete nature of information is the key for the passage from classical mechanics to quantum theory. Paraphrasing Anton Zeilinger:

*In conclusion it may very well be said that information is the irreducible kernel from which everything else flows. The question why Nature appears quantized is simply a consequence of the fact that information itself is quantized. It might even be fair to observe that the concept that information is fundamental is very old knowledge of humanity, witness for example the beginning of gospel according to John: “In the beginning was the Word and the Word was with God, and the Word was God”.*

## References

- [1] Abraham, R., Marsden, J.E., Ratiu, T.: Manifolds, Tensor Analysis, and Applications. Applied Mathematical Sciences **75**, Springer (1988)

- [2] Banyaga, A.: Sur la structure du groupe des difféomorphismes qui préservent une forme symplectique. *Comm. Math. Helv.* **53**, 174–227 (1978)
- [3] Banyaga, A.: *The structure of classical diffeomorphism groups*, Kluwer Academic Publishers (1997)
- [4] Feynman, R.P., Leighton, R.B., Sands, M.: *The Feynman Lectures on Physics, III*, p.16-12, Addison-Wesley, Reading, MA (1965)
- [5] de Gosson, M.: *Symplectic Geometry and Quantum Mechanics*, Birkhäuser, Basel, (2006)
- [6] de Gosson, M.: The Symplectic Camel and the Uncertainty Principle: The Tip of an Iceberg? *Found. Phys.* **99**, 194–214 (2009)
- [7] de Gosson, M., Luef, F.: Symplectic Capacities and the Geometry of Uncertainty: the Irruption of Symplectic Topology in Classical and Quantum Mechanics. *Physics Reports* **484**, 131–179 (2009)
- [8] Gröchenig, K: *Foundations of Time-Frequency Analysis*. Birkhäuser, Boston (2001)
- [9] Gromov, M.: Pseudoholomorphic curves in symplectic manifolds. *Invent. Math.* **82**, 307–47 (1985)
- [10] Kastler, D.: The  $C^*$ -Algebras of a Free Boson Field. *Commun. math. Phys.* **1**, 114–48 (1965)
- [11] Lanczos, C.: *The Variational Principles of Mechanics*. Toronto, Ontario: University of Toronto Press (1949). Reprint 4th edn. New York: Dover Publications (1986)
- [12] Littlejohn, R.G.: The semiclassical evolution of wave packets. *Physics Reports* **138**(4–5), 193–291 (1986)
- [13] Louprias, G., Miracle-Sole, S.:  $C^*$ -Algèbres des systèmes canoniques, I. *Commun. math. Phys.* **2**, 31–48 (1966)
- [14] Louprias, G., Miracle-Sole, S.:  $C^*$ -Algèbres des systèmes canoniques, II. *Ann. Inst. Henri Poincaré* **6**(1), 39–58 (1967)
- [15] Mackey, G.W.: The Relationship Between Classical and Quantum Mechanics. In *Contemporary Mathematics* **214**, Amer. Math. Soc., Providence, RI (1998)

- [16] Narcowich, F.J., O’Connell, R.F.: Necessary and sufficient conditions for a phase-space function to be a Wigner distribution. *Phys. Rev. A* **34**(1), 1–6 (1986)
- [17] Narcowich, F.J.: Geometry and uncertainty. *J. Math. Phys.* **31**(2) (1990)
- [18] Polterovich, L.: The Geometry of the Group of Symplectic Diffeomorphisms. *Lectures in Mathematics*, Birkhäuser (2001)
- [19] Reed M., Simon B.: *Methods of Modern Mathematical Physics*. Academic Press, New York (1972)
- [20] Schrödinger, E. , Quantisierung als Eigenwertproblem, *Ann. der Physik*, **384**, (1926), 361-376.
- [21] Shubin, M.A.: *Pseudodifferential Operators and Spectral Theory*, Springer-Verlag (1987) [original Russian edition in Nauka, Moskva (1978)]
- [22] Stein, E.M.: *Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals*. Princeton University Press (1993)
- [23] Stone, M.H.: Linear transformations in Hilbert space, III: operational methods and group theory. *Proc. Nat. Acad. Sci. U.S.A.*, 172–175 (1930)
- [24] Struckmeier, J.: Hamiltonian dynamics on the symplectic extended phase space for autonomous and non-autonomous systems. *J. Phys. A: Math. Gen.* **38**, 1257–1278 (2005)
- [25] Synge, J.L.: *Encyclopedia of Physics*, vol **3/1** ed. S. Flügge, Berlin: Springer (1960)
- [26] Wong, M.W.: *Weyl Transforms*, Springer-Verlag (1998).