A Pseudo-Quantum Triad: Schrödinger’s Equation, the Uncertainty Principle, and the Heisenberg Group

Maurice A. de Gosson
University of Vienna, Faculty of Mathematics, NuHAG, Nordbergstraße 15, A-1090 Vienna
AUSTRIA
E-mail: maurice.de.gosson@univie.ac.at

Abstract. We show that the paradigmatic quantum triad “Schrödinger equation–Uncertainty principle–Heisenberg group” emerges mathematically from classical mechanics. In the case of the Schrödinger equation, this is done by extending the metaplectic representation of linear Hamiltonian flows to arbitrary flows; for the Heisenberg group this follows from a careful analysis of the notion of phase of a Lagrangian manifold, and for the uncertainty principle it suffices to use tools from multivariate statistics together with the theory of John’s minimum volume ellipsoid. Thus, the mathematical structure needed to make quantum mechanics emerge already exists in classical mechanics.

1. Introduction
The Schrödinger equation, the uncertainty principle, and the Heisenberg group are three basic paradigms of Quantum Mechanics. The aim of this contribution is to show that this “triad” can actually be rigorously (i.e. mathematically) constructed within classical mechanics in its Hamiltonian formulation; Planck’s constant $\hbar$ then appears as a scaling parameter whose dimension is that of an action. This fact is actually very much in accordance with Mackey’s [31] statement following which quantum mechanics is a refinement of classical mechanics. Quantum Mechanics can thus only emerge if one can give a physical meaning to this triad, justifying the need for Planck’s constant $\hbar$. We will not attempt to do this in the present work.

Here is a short description of what we are going to do:

• **The Schrödinger Equation.** According to conventional wisdom, Schrödinger’s equation cannot be derived from classical considerations; to witness Feynman’s declaration: “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. However, at the time when Feynman made this declaration (1965) it was already pretty well-known among specialists working in harmonic analysis, that the Schrödinger equation is implicit in the metaplectic representation of the symplectic group; in fact it was already known that Schrödinger’s equation could be rigorously derived for all Hamiltonians which are quadratic in the position and momentum variables. We will show that this equation can actually be derived for all Hamiltonian functions, not only for those of the classical type “kinetic energy plus potential”.

• **The Uncertainty Principle.** In Classical Mechanics (CM) it is tacitly assumed that the physical quantities that are measured have an exact value, that can be in principle...
determined. In practice, the numerical results obtained from our measurements will cluster around this objective exact value, and one then uses statistical methods to obtain approximations to the exact value. In Quantum Mechanics (QM) the situation is different; of course there are also measurements errors, but in the limiting case where perfect measurements is assumed, the results of identical experiments performed on identically prepared systems are generally not yielding identical results (this is well illustrated by the discussion in Peres [33], §4-3, of a spin experiment). This quantum indeterminacy is traditionally expressed in terms of the Uncertainty Principle of QM. We will show that the Heisenberg inequalities (or their refinement, the Robertson–Schrödinger inequalities) are—at least formally—perfectly classical.

- **The Heisenberg Group.** The theory of the Heisenberg group is of a slightly more abstract nature. Its “Schrödinger representation” is supposed to be a key to quantization, as explained in Refs. [23, 24]. We will see that the Heisenberg group and the associated Heisenberg–Weyl operators have a very simple interpretation in terms of the notion of phase of a Lagrangian manifold, using the properties of the Poincaré–Cartan invariant familiar from the theory of the Hamilton–Jacobi equation from classical mechanics.

**Acknowledgement**

This work has been financed by the Austrian Research Agency FWF (Projektnummer P 23902-N13). The author wishes to thank the organizers of the Heinz von Foerster Congress in Vienna.

2. Symplectic Mechanics

We gather in this preliminary section the tools from symplectic geometry and Hamiltonian mechanics we will need; for details and proofs see Refs. [1, 4, 13, 15, 18]; in [15] we have given a detailed study of the metaplectic group.

2.1. The symplectic group

A symplectic matrix is a real matrix $S$ of size $2n$ satisfying anyone of the equivalent relations $S^TJS = J$ or $SJS^T = J$; here $S^T$ is the transpose of $S$ and $J$ is the skew-symmetric matrix $egin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ where $0$ and $I$ are, respectively, the $n \times n$ zero and identity matrices. In particular a $2 \times 2$ real matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is symplectic if and only if its determinant $ad - bc$ is equal to one; but it is not true in dimension $n > 1$ that every unimodular matrix is symplectic. Using the identity $S^TJS = J$ (or $SJS^T = J$) one easily checks that the symplectic matrices form a group, which is denoted by $Sp(2n, \mathbb{R})$.

The symplectic group $Sp(2n, \mathbb{R})$ is generated by the matrix $J$ together with all matrices

$$M_L = \begin{pmatrix} L^{-1} & 0 \\ 0 & L^T \end{pmatrix}, \quad V_{-P} = \begin{pmatrix} I & 0 \\ P & I \end{pmatrix}$$

(1)

where $\det L \neq 0$ and $P = P^T$ ($L$ and $P$ having size $n$); the matrices $M_L$, and $V_{-P}$ are trivially symplectic. Since the determinants of $J$, $M_L$, and $V_{-P}$ are equal to one, this shows at the same time that the determinant of a symplectic matrix always is equal to one. This fact implies Liouville’s theorem on the conservation of phase space volume under Hamiltonian flows, see Refs. [4, 12].

2.2. The metaplectic representation

The symplectic group is connected and has covering spaces of all orders. Among all these, the two-fold covering group plays a very special role, because it can be faithfully represented by
a group of unitary operators acting on the square integrable functions. This operator group is called the metaplectic group; we will denote it by $\text{Mp}(2n, \mathbb{R})$. The easiest way to describe $\text{Mp}(2n, \mathbb{R})$ is to use generators. Since $\text{Mp}(2n, \mathbb{R})$ is a two-fold covering of $\text{Sp}(2n, \mathbb{R})$ to each of the generators $J, M_L, V_{-p}$ correspond two elements of $\text{Mp}(2n, \mathbb{R})$, differing by a sign. These operators $\hat{J}, \hat{M}_L, \hat{V}_{-p}$ are defined by

$$
\hat{J}_\psi(x) = \left( \frac{1}{2\pi i} \right)^{n/2} \int e^{ix \cdot y} \psi(y) d^n y
$$

$$
\hat{M}_L \psi(x) = \sqrt{\text{det} L} \psi(Lx)
$$

$$
\hat{V}_{-p} \psi(x) = e^{\frac{i}{\hbar} p \cdot x} \psi(x);
$$

the argument of $\text{det} L$ in formula (3) is chosen to be 0 or $2\pi$ if $\text{det} L > 0$, and $\pi$ or $3\pi$ if $\text{det} L < 0$. The covering projection $\Pi : \text{Mp}(2n, \mathbb{R}) \rightarrow \text{Sp}(2n, \mathbb{R})$ is determined by the equalities

$$
\Pi(\hat{J}) = J, \quad \Pi(\hat{M}_L) = M_L, \quad \Pi(\hat{V}_{-p}) = V_{-p}.
$$

However, one obtains another perfectly honest covering projection by using inner automorphisms of the base group $\text{Sp}(2n, \mathbb{R})$ (see e.g. de Gosson [13], §6.4.2, for a simple description of this conjugation relation). This means that given an arbitrary element $S_0$ of $\text{Sp}(2n, \mathbb{R})$ the mapping $\Pi_0$ defined by $\Pi_0(\hat{S}) = S_0 \Pi(\hat{S}) S_0^{-1}$ is also a bona fide covering projection. Notice that the definition of $\Pi_0$ can be rewritten in the form

$$
\Pi_0(\hat{S}) = \Pi(\hat{S_0} \hat{S} \hat{S_0}^{-1}), \quad \Pi_0(\hat{S}_0) = S_0
$$

since $\Pi$ is a group homomorphism. Choose now $S_0 = M_{1/\sqrt{\varepsilon}}$ (it is the symplectic matrix $M_L$ with $L = (1/\sqrt{\varepsilon})I$) and define the projection

$$
\Pi^\varepsilon(\hat{S}) = M_{1/\sqrt{\varepsilon}} \Pi(\hat{S}) M_{\sqrt{\varepsilon}} = \Pi(\hat{M}_{1/\sqrt{\varepsilon}} \hat{S} \hat{M}_{\sqrt{\varepsilon}}).
$$

One immediately verifies that if we use $\Pi^\varepsilon$ instead of $\Pi$, the projections formulas (5) become

$$
\Pi^\varepsilon(\hat{J}) = J, \quad \Pi^\varepsilon(\hat{M}_L) = M_L, \quad \Pi^\varepsilon(\hat{V}_{-p}) = V_{-p}.
$$

the metaplectic operators $\hat{J}^\varepsilon, \hat{M}_L^\varepsilon, \hat{V}_{-p}^\varepsilon$ being defined by

$$
\hat{J}^\varepsilon \psi(x) = \left( \frac{1}{2\pi i\varepsilon} \right)^{n/2} \int e^{\varepsilon x \cdot y} \psi(y) d^n y
$$

$$
\hat{M}_L^\varepsilon \psi(x) = \hat{M}_L \psi(x) = \sqrt{\text{det} L} \psi(Lx)
$$

$$
\hat{V}_{-p}^\varepsilon \psi(x) = \hat{V}_{-p} \psi(x) = e^{\frac{i}{\varepsilon} p \cdot x} \psi(x).
$$

In QM one chooses $\varepsilon = \hbar$.

2.3. Hamiltonian flows

Let $H = H(z)$ (the “Hamiltonian”) be a smooth function of the variables $x, p$. Hamilton’s equations of motion $\dot{x} = -\nabla_p H(x, p)$, $\dot{p} = -\nabla_x H(x, p)$ can be written in compact form as

$$
\dot{z} = J \nabla_z H(z).
$$

The presence of the matrix $J$ in this formula indicates that the symplectic group might lurk behind Hamiltonian mechanics. Let us examine this “educated guess” somewhat more in detail.
We denote by \((f^H_t)\) the associated phase flow: \(f^H_t\) is the mapping which takes an initial point \(z_0 = (x_0, p_0)\) at time \(t = 0\) to the point \(z = (x, p)\) at time \(t\) along the solution curve to (12) through \(z_0\). That is, \(z(t) = f^H_t(z_0)\) is the solution of (12) with initial condition \(z(0) = z_0\).

The importance of the symplectic group in Hamiltonian mechanics comes from the following fundamental property of the phase flows \((f^H_t)\): each mapping \(f^H_t\) is a canonical transformation; see [1, 4, 12, 15, 18]. This means that for every \(z = (x, p)\) the Jacobian matrix of \(f^H_t\) calculated at \(z = (x, p)\) is symplectic:

\[
Df^H_t(z) = \frac{\partial (x(t), p(t))}{\partial (x_0, p_0)} \in \text{Sp}(2n, \mathbb{R}).
\]

In particular, when the Hamiltonian function is quadratic in the \(x\) and \(p\) variables, that is of the type

\[
H(z) = \frac{1}{2} M z \cdot z = \frac{1}{2} (x, p) M (x, p)^T,
\]

the flow determined by Hamilton’s equations is linear, and consists of the symplectic matrices \(S^H_t = e^{ttM}\). As we are going to see below, the study of linear flows is the first step towards a general derivation of Schrödinger’s equation.

3. **Schrödinger’s Equation**

We now proceed to show how the Schrödinger equation emerges from Hamiltonian mechanics. We assume that the Hamiltonians are time-independent, but everything in this Section can be extended to the time-dependent case at the price of some technical difficulties (see de Gosson [18], de Gosson and Hiley [19]).

3.1. **Quadratic Hamiltonians**

Consider first a quadratic Hamiltonian of the type (13). As we observed above, the phase flow can easily be calculated, and consists of symplectic matrices \(S^H_t\). We have of course \(S^H_t S^H_{t'} = S^H_{t+t'}\) so the flow is a one-parameter subgroup of \(\text{Sp}(2n, \mathbb{R})\). When \(t\) varies these matrices describe a curve in \(\text{Sp}(2n, \mathbb{R})\) which passes by the identity at the initial time \(t = 0\). Now, the unique path lifting theorem in the theory of covering groups says that this curve can be lifted to any of the covering groups of \(\text{Sp}(2n, \mathbb{R})\) and that this can be made in an unique way if one specifies a point through which the lifting should pass. Choosing the metaplectic group \(\text{Mp}(2n, \mathbb{R})\) as the covering group, with a projection \(\Pi^f : \text{Mp}(2n, \mathbb{R}) \rightarrow \text{Sp}(2n, \mathbb{R})\), and imposing to the lifting to pass through the identity of \(\text{Mp}(2n, \mathbb{R})\) at time \(t = 0\) we thus obtain a curve of metaplectic operators \(\hat{S}^H_t\) such that \(\Pi^f(\hat{S}^H_t) = S^H_t\); one moreover verifies that the group property \(S^H_t S^H_{t'} = S^H_{t+t'}\) carries over to these operators: \(\hat{S}^H_t \hat{S}^H_{t'} = \hat{S}^H_{t+t'}\). The operators \(\hat{S}^H_t\) are unitary and act on the square integrable functions on \(\mathbb{R}^n\). For such a function \(\psi_0\) we set \(\psi(x, t) = \hat{S}^H_t \psi_0\) and ask the question “what is the partial differential equation satisfied by \(\psi(x, t)\)?” The answer is: if the initial function \(\psi_0\) is at least twice continuously differentiable in the \(x\) variables, then

\[
\frac{\partial \psi}{\partial t}(x, t) = \hat{H}(x, -i\varepsilon \nabla_x) \psi(x, t)
\]

where \(\hat{H}(x, -i\varepsilon \nabla_x)\) is obtained from the Hamiltonian function (13) by the formal substitution \((x, p) \rightarrow (x, -i\hbar \nabla_x)\). Writing \(M = \begin{pmatrix} H_{xx} & H_{xp} \\ H_{px} & H_{pp} \end{pmatrix}\) this is

\[
\hat{H}(x, -i\varepsilon \nabla_x) = -\frac{\varepsilon^2}{2} H_{pp} x^2 - i\varepsilon H_{px} x \nabla_x + \frac{1}{2} H_{xx} x^2 - i \text{Tr}(H_{px})
\]
where $\text{Tr}(H_{px})$ (the trace) is the sum of the diagonal elements of $H_{px}$.

Thus, if we choose $\varepsilon = \hbar$ we obtain Schrödinger’s equation; but this particular choice of parameter has to be motivated by physical considerations once a physical meaning has been given to the solution $\psi$. What must be remembered is that the derivation of (14) heavily depends on the choice of projection $\Pi^\varepsilon$ of $\text{Mp}(2n, \mathbb{R})$ onto $\text{Sp}(2n, \mathbb{R})$: each choice of $\varepsilon$ is a priori equally good.

We have thus established the existence, for each $\varepsilon > 0$, of a one-to-one and onto correspondence

$$C_\varepsilon: (S^H_t) \mapsto (\hat{S}^H_t)$$

(16)

between phase flows arising from quadratic Hamiltonian functions and unitary evolution groups $\hat{S}^H_t = e^{-it\hat{H}/\varepsilon}$, $\hat{H} = \hat{H}(x, -i\varepsilon \nabla_x)$. We now make a fundamental observation. Assume that we make a symplectic change of variables in the Hamiltonian $H(z)$, i.e. that we replace $H(z)$ with $H(S^{-1}z)$ where $S$ is in $\text{Sp}(2n, \mathbb{R})$. Then, the phase flow $(f_t^{H_0S^{-1}})$ of $H(S^{-1}z)$ is obtained from the flow $(f_t^H)$ by conjugation with $S$:

$$f_t^{H_0S^{-1}} = S_t^H S^{-1}$$

(17)

(see e.g. Refs. [4, 15, 18, 12, 34]). In the case of quadratic Hamiltonians we have $S_t^{H_0S^{-1}} = SS_t^H S^{-1}$ and the group of unitary operators $S_t^H$ is replaced with $\hat{S}^H_t S^{-1}$. It follows that the correspondence $C_\varepsilon$ satisfies the covariance formula

$$C_\varepsilon(SS_t^H S^{-1}) = (\hat{S}SS_t^H \hat{S}^{-1}).$$

(18)

3.2. The general case

We begin by recalling two results from functional analysis. The first is Stone’s theorem [45] about one-parameter groups of operators (see Refs. [1, 35] for “modern” proofs):

**Stone’s Theorem:** (i) For every strongly continuous one-parameter group $(F_t)$ of unitary operators on a Hilbert space $L^2(\mathbb{R}^n)$ 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 $\mathcal{H}$. (ii) Conversely, if $A$ is a self-adjoint operator on $L^2(\mathbb{R}^n)$ then there exists a unique one-parameter unitary group $(F_t)$ whose infinitesimal generator is $A$, that is $F_t = e^{itA/\hbar}$.

The second result says that every continuous operator (in a sense we will be precise) has a distributional kernel:

**Schwartz’s kernel theorem:** Let $A$ be a linear operator $S(\mathbb{R}^n) \rightarrow S'(\mathbb{R}^n)$. We assume that $A$ is continuous in the sense that if $(\psi_k)$ is a sequence of functions such that $\psi_k \rightarrow \psi$ for $k \rightarrow \infty$ in $S(\mathbb{R}^n)$ then $A\psi_k \rightarrow A\psi$ in $S(\mathbb{R}^n)$. Then there exists a distribution $K_A$ (the kernel of $A$) such that

$$A\psi(x) = \int K_A(x,y)\psi(y)dy$$

(19)

for all functions $\psi$ in $S(\mathbb{R}^n)$ [the integral in the right hand-side of (19) is interpreted as a distributional bracket].

Let now $H$ be an arbitrary Hamiltonian function; we assume that the phase flow $(f_t^H)$ generated by the Hamilton equations for $H$ exist for all times $t$ (this is just a technical requirement, which can be alleviated; see de Gosson and Hiley [19]). We are going to show that the functor (16) can be extended in a unique way to arbitrary Hamiltonian flows if one makes the following requirement:
The infinitesimal generator \(\hat{L}\) thus show that the operators \(\hat{\omega}\) now show that definition (21) of the functor \(C\) so that

\[
C(\hat{f}_t^H) = (\hat{S}\hat{f}_t^H\hat{S}^{-1})
\]

should hold for each \(S\) in \(\text{Sp}(2n, \mathbb{R})\).

We have again \(\hat{f}_t^H S^{-1} = f_t H S^{-1}\) in view of formula (17), and \((\hat{f}_t^H)\) is a strongly continuous one-parameter group of unitary operators on \(L^2(\mathbb{R}^n)\) which has to be determined.

To the phase flow \((\hat{f}_t^H)\) (that is, equivalently, to the Hamiltonian function \(H\)) we associate the Weyl operator

\[
\hat{H} = \left(\frac{1}{2\pi i\varepsilon}\right)^{n/2} \int H_\sigma(z)\hat{T}_\varepsilon(z) d^{2n}z
\]

where \(\hat{T}_\varepsilon(z) = e^{-i\sigma(z)/\varepsilon}\) is the \(\varepsilon\)-Heisenberg–Weyl operator (cf. formula (38) and

\[
H_\sigma(z_0) = \left(\frac{1}{2\pi i}\right)^n \int e^{-i\sigma(z,z')} H(z') d^{2n}z'
\]

is the \(\varepsilon\)-symplectic Fourier transform of \(H\). Since \(H\) is real it follows from the standard theory of Weyl operators (see Refs. 15, 18, 29) that the operator \(\hat{H}\) is self-adjoint and densely defined on the Schwartz space of rapidly decreasing functions. In view of part (ii) of Stone’s theorem there exists a unique one-parameter group of unitary operators \(\hat{f}_t^H = e^{-it\hat{H}/\varepsilon}\), so we define

\[
C_\varepsilon(f_t^H) = (\hat{f}_t^H) = (e^{-it\hat{H}/\varepsilon}).
\]

To show that this is the desired extension of the functor \(C_\varepsilon\) we have to check that it satisfies the symplectic covariance property (20) and that its restriction to linear flows \((S_t^H)\) is given by (16). The symplectic covariance property is proven as follows: since \(\hat{f}_t^H S^{-1} = f_t H S^{-1}\) the one-parameter group \(e^{-it\hat{H} S^{-1}/\varepsilon}\) of Weyl operators we have \(H \circ S^{-1} = \hat{S} H \hat{S}^{-1}\) and hence

\[
e^{-it\hat{H} S^{-1}/\varepsilon} = e^{-it\hat{S} H \hat{S}^{-1}/\varepsilon} = \hat{S} e^{-it\hat{H}/\varepsilon} S^{-1}
\]

so that \(C_\varepsilon(f_t^H S^{-1}) = \hat{S} C_\varepsilon(f_t^H) \hat{S}^{-1}\) which is precisely formula (20) we set out to prove. Let us now show that definition (21) of the functor \(C_\varepsilon\) implies that we have \(C_\varepsilon(S_t^H) = (S_t^H)\). We must thus show that the operators \(\hat{f}_t^H = e^{-it\hat{H}/\varepsilon}\) are metaplectic when \(H\) is a quadratic Hamiltonian. The infinitesimal generator \(\hat{H}\) of \(\hat{f}_t^H\) satisfies the Schrödinger equation

\[
i\hbar \frac{d}{dt} \hat{f}_t^H = \hat{H} \hat{f}_t^H
\]

where \(\hat{H} \xrightarrow{\text{Weyl}} \hat{H}\). Now, the operator \(\hat{H}(x, -i\varepsilon \nabla_x)\) defined by formula (15) is precisely the Weyl quantization of the quadratic Hamiltonian function \(H\) hence the \(\hat{f}_t^H\) are precisely the metaplectic operators \(\hat{S}_t^H\) corresponding to the lift of the linear flow \((S_t^H)\) to \(\text{Mp}(2n, \mathbb{R})\).

We next have to show that, conversely, for every strongly-continuous one-parameter group \(\hat{f}_t\) of unitary operators on \(L^2(\mathbb{R}^n)\) there exists a Hamiltonian function \(H\) such that \(\hat{f}_t = C_\varepsilon(f_t^H)\). This amounts to show that there exists a Weyl operator \(\hat{H} \xrightarrow{\text{Weyl}} \hat{H}\) such that \(\hat{f}_t = e^{-it\hat{H}/\varepsilon}\). We begin by remarking that in view of part (i) of Stone’s theorem there exists a unique self-adjoint operator \(\hat{H}\) such that \(\hat{f}_t = e^{-it\hat{H}/\varepsilon}\) for all times \(t\), so all we have to do is to prove that \(\hat{H}\) is indeed a Weyl operator; the Weyl correspondence \(\hat{H} \xrightarrow{\text{Weyl}} H\) will then identify the Hamiltonian...
function such that $\tilde{F}_t = C_\varepsilon (f^H_t)$. The restriction of $\tilde{H}$ to the dense subspace $\mathcal{S}(\mathbb{R}^n)$ is continuous $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ because $\tilde{H}$ is essentially self-adjoint; it follows that Schwartz’s kernel theorem applies, and we can thus find $K$ such that

$$\hat{H} \psi(x) = \int K(x,y) \psi(y) d^n y.$$  \hfill (22)

Let now $\tau$ be an arbitrary real number and define the “$\tau$-symbol” $H_\tau$ of $\hat{H}$ by the Fourier integral

$$H_\tau(z) = \int e^{-\frac{i}{\varepsilon}py} K(x + \tau y, x - (1 - \tau)y) d^n y.$$ \hfill (23)

One shows (de Gosson [18], Shubin [44]) that the operator $\hat{H}$ can be expressed in terms of $H_\tau$ by the formula

$$\hat{H} \psi(x) = \left( \frac{1}{2\pi \varepsilon} \right)^n \int \int e^{-\frac{i}{\varepsilon}py} H_\tau((1 - \tau)x + \tau y, p) \psi(y) d^n y d^n p.$$ \hfill (24)

which is the “pseudo-differential representation of $\hat{H}$ in terms of its $\tau$-symbol”. We next observe that if the operator is multiplication by a function $a(x)$ then we must have $K(x,y) = a(x) \delta(x-y)$ and hence, by (23),

$$H_\tau(z) = a(x) \int e^{-\frac{i}{\varepsilon}py} \delta(y) d^n y = a(x).$$

This property, together with the symplectic covariance requirement (20) implies (see Refs. [43, 50]), that this is only possible if we choose $\tau = \frac{1}{2}$, in which case formula (24) becomes

$$\hat{H} \psi(x) = \left( \frac{1}{2\pi \varepsilon} \right)^n \int \int e^{-\frac{i}{\varepsilon}py} H_{1/2}(\frac{1}{2}(x + y), p) \psi(y) d^n y d^n p.$$ \hfill (24)

But this is exactly the integral representation of the Weyl operator $\hat{H} \xleftarrow{\text{Weyl}} H = H_{1/2}$, and we are done.

### 3.3. The Groenewold and Van Hove “no-go” theorem

It is sometimes claimed that one cannot “quantize classically” arbitrary non-quadratic observables; to sustain this claim an old (and famous) series of results going back to Groenewold [21] and van Hove [26, 27] is invoked. But this objection cannot be retained in the present case: a careful mathematical reading of the Groenewold-van Hove obstruction results reveals that what is actually proven is the impossibility to quantize observables in such a way that Dirac’s [9] prescription that Poisson brackets should correspond to commutators is preserved. And, indeed, such a quantization is a chimera. This state of affairs is very lucidly analysed by Giulini’s [11] who reminds us that Dirac [9] himself had noted that

...The strong analogy between quantum Poisson brackets [i.e. commutators] and classical Poisson brackets leads us to make the assumption that the quantum Poisson brackets, or at any rate the simpler ones of them, have the same values as the corresponding classical Poisson brackets.

The italics (added) show that Dirac himself was very aware of the fact that it was not reasonable to expect that the Poisson bracket/commutator correspondence should hold for all pairs of observables $H$ and $K$. The Groenewold–van Hove obstructions thus have nothing to do with the impossibility of a derivation of Schrödinger’s equation. They only show that one has to go beyond the metaplectic representation if one wants to quantize non-quadratic Hamiltonians. This is precisely what we have done by extending the correspondence $C_\varepsilon$ to arbitrary Hamiltonian flows.
4. The Uncertainty Principle

We are following in this section the discussion in de Gosson [16, 17].

4.1. Uncertainty inequalities

Contrarily to what is often believed Heisenberg’s uncertainty principle \( \Delta p_j \Delta x_j \geq \frac{1}{2} \hbar \) is not a statement about the accuracy of our measurement instruments; its derivation assumes on the contrary perfect instruments. The correct interpretation of Heisenberg’s inequalities is the following (see e.g. Peres [33], p.93): if the same preparation procedure is repeated a large number of times, and is followed by either by a measurement of \( x_j \), or by a measurement of \( p_j \), the results obtained have standard deviations \( \Delta x_j \) and \( \Delta p_j \) satisfying the Heisenberg inequalities. The same interpretation is of course true for the stronger Robertson–Schrödinger [36, 40] inequalities

\[
(\Delta p_j)^2 (\Delta x_j)^2 \geq \Delta(x_j, p_j)^2 + \frac{1}{4} \hbar^2
\]

(25)

to which the Heisenberg inequalities reduce if one neglects the covariances \( \Delta(x_j, p_j)^2 \); they are complemented by the trivial inequalities

\[
\Delta p_j \Delta x_k \geq 0 \text{ if } j \neq k, \quad \Delta p_j \Delta p_k \geq 0, \quad \Delta x_j \Delta x_k \geq 0.
\]

(26)

where \( \theta_{jk}^2 = \theta_{kj}^2 \) and \( \eta_{jk}^2 = \eta_{kj}^2 \) (see Dias et al. [8] and the references therein). In classical statistical mechanics the situation is somewhat different: due to the inherent inaccuracy of the measurement apparatus there are uncertainties for all pairs of variables, conjugate or not. This is actually reminiscent of what happens in noncommutative mechanics (NCQM) where the second and third inequalities (26) are replaced with

\[
\Delta p_j^2 \Delta p_k^2 \geq \Delta(p_j, p_k)^2 + \frac{1}{4} \theta_{jk}^2, \quad \Delta x_j^2 \Delta x_k^2 \geq \Delta(x_j, x_k)^2 + \frac{1}{4} \eta_{jk}^2.
\]

(27)

4.2. The Minimum Volume Ellipsoid method

Performing position and momentum measurements on a large number \( K \) of identical copies of a system of particles, we get a cloud \( S = \{z_1, z_2, ..., z_N\}, N = nK \), of points in \( \mathbb{R}^{2n} \). An efficient method for studying that cloud consists in using the minimum volume ellipsoid (MVE) method (Rousseeuw [37], Van Aelst and Rousseeuw [42]). Geometrically speaking this method is an application of John’s theorem [28] which says that a convex set is contained in a unique ellipsoid with smallest volume. It works as follows: consider a subset \( \{z_{i1}, z_{i2}, ..., z_{ik}\} \) of \( S \) of points in general position. This condition is sufficient and necessary for any ellipsoid containing these points to have positive volume. The points \( z_{i1}, z_{i2}, ..., z_{ik} \) determine a polyhedron in \( \mathbb{R}^{2n} \); we denote by \( K \) the convex hull of that polyhedron (it is the smallest subset of \( \mathbb{R}^{2n} \) containing \( \{z_{i1}, z_{i2}, ..., z_{ik}\} \)). John’s theorem ensures us that there exists a unique ellipsoid \( J \) in \( \mathbb{R}^{2n} \) containing \( K \) and having minimum volume among all the ellipsoids having this property. Repeating this process for all subsets of the cloud \( S \) having \( k \) elements in general position we get a family of ellipsoids; by definition the MVE is the one with the smallest volume. To determine that ellipsoid one proceeds as follows: choose an integer \( k \) between \( \lfloor N/2 \rfloor + 1 \) and \( N \) (\( \lfloor N/2 \rfloor \) the integer part of \( N/2 \)); this constant \( k \) determines the robustness of the resulting estimator; a common choice is to take \( k = \lfloor (N + 2n + 1)/2 \rfloor \). The first step of the procedure is to minimize the determinant of the matrices \( M \) subject to the condition

\[
\# \{ j : (z_j - \bar{z})^T M^{-1} (z_j - \bar{z}) \leq m^2 \} \geq k
\]

(28)

where the minimization is over all \( \bar{z} \in \mathbb{R}^{2n} \) and all positive definite symmetric matrices \( M \) of size \( 2n \). Here \( m \) is a fixed constant, chosen so that the MVE estimator is a consistent estimator for of the covariance matrix for data coming from a multivariate normal distribution, that is
\[ m = \sqrt{\chi_{2n, \alpha}^2}, \quad \alpha = k/N \] where \( \chi_{2n, \alpha}^2 \) is a chi-square distribution with \( 2n \) degrees of freedom (see Lopuhaä and Rousseeuw [30]). Once the pair \((M, z)\) is determined, the minimum volume ellipsoid (MVE) is the set of all \( z \) in \( \mathbb{R}^{2n} \) such that
\[ (z - \bar{z})^T M^{-1} (z - \bar{z}) \leq m^2. \tag{29} \]

The second step consists in associating to the MVE \( J \) a covariance matrix. For this one has to choose an adequate value \( m_0 \) for \( m \); denoting the corresponding matrix \( M \) by \( \Sigma \) the MVE is the ellipsoid
\[ C : (z - \bar{z})^T \Sigma^{-1} (z - \bar{z}) \leq m_0^2 \tag{30} \]
and \( \Sigma \) is then precisely the covariance matrix. We will write that matrix in the form
\[ \Sigma = \begin{pmatrix} \Delta(x, x) & \Delta(x, p) \\ \Delta(p, x) & \Delta(p, p) \end{pmatrix} \tag{31} \]
where \( \Delta(x, x) = ((\Delta(x_i, x_j))_{1 \leq i,j \leq n} \) and so on; writing as is customary \( \Delta(x_i, x_i) = \Delta x_i^2 \) and \( \Delta(p_i, p_i) = \Delta p_i^2 \) we have \( \Delta(x_i, x_j) = \Delta(x_i, x_i), \Delta(p_i, p_j) = \Delta(p_j, p_i), \Delta(x_i, p_j) = \Delta(p_j, x_i) \) (observe that \( \Sigma \) is symmetric).

### 4.3. A condition on the covariance matrix

Let \( A = (a_{jk})_{1 \leq j,k \leq n} \) and \( C = (c_{jk})_{1 \leq j,k \leq n} \) be two real antisymmetric matrices, and \( B = (b_{jk})_{1 \leq j,k \leq n} \) a real symmetric matrix. To \( A, B, C \) we associate the \( 2n \times 2n \) antisymmetric matrix
\[ \Omega = \begin{pmatrix} A & B \\ -B & C \end{pmatrix} \]
which is the most general form an antisymmetric of size \( 2n \) can have. Assuming \( \Omega \) invertible, the bilinear form \( \omega \) on \( \mathbb{R}^{2n} \) defined by \( \omega(z, z') = -(z')^T \Omega^{-1} z \) is a symplectic form on \( \mathbb{R}^{2n} \). Notice that when \( A = C = 0 \) and \( B = I \) we have \( \Omega = J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \). The general case is reduced to the standard case by noting that there exists a matrix \( F \) such that \( \Omega = F^T J F \); equivalently \( F \) is a linear canonical transformation \( (\mathbb{R}^{2n}, \sigma) \rightarrow (\mathbb{R}^{2n}, \omega) \).

Let now \( \Sigma \) be the covariance matrix (31). The matrix
\[ \Sigma + i\Omega = \begin{pmatrix} \Delta(x, x) + iA & \Delta(x, p) + iB \\ \Delta(p, x) - iB & \Delta(p, p) + iC \end{pmatrix} \tag{32} \]
is Hermitian since \( \Sigma \) is symmetric and \( (i\Omega)^* = i\Omega \), hence the eigenvalues of \( \Sigma + i\Omega \) are real. Assuming in addition that these eigenvalues are nonnegative, \( \Sigma + i\Omega \) is semi-definite positive, which we write \( \Sigma + i\Omega \geq 0 \). It follows that
\[ \Delta(x, x) + iA \geq 0, \quad \Delta(p, p) + iC \geq 0. \tag{33} \]

When \( n = 1 \) the covariance matrix is just
\[ \Sigma = \begin{pmatrix} \Delta x^2 & \Delta(x, p) \\ \Delta(p, x) & \Delta p^2 \end{pmatrix} \]
and the antisymmetric matrices \( \Theta \) and \( N \) are zero so that \( \Omega = aJ = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \). The condition \( \Sigma + i\Omega = \Sigma + iaJ \geq 0 \) is in this case equivalent to
\[ \begin{pmatrix} \Delta x^2 + ia & \Delta(x, p) \\ \Delta(p, x) & \Delta p^2 - ia \end{pmatrix} \geq 0 \]
that is, to the single inequality
\[ \Delta x^2 \Delta p^2 \geq \Delta (x, p)^2 + a^2 \]
which formally reduces to the Robertson–Schrödinger inequality (25) if one chooses \( \varepsilon = \hbar/2 \).

In higher dimensions one shows, using Sylvester’s criterion which says that a Hermitian matrix is positive semidefinite if and only if all of its principal minors are nonnegative, that the condition \( \Sigma + i \Omega \geq 0 \) implies that
\begin{align*}
\Delta x_j^2 \Delta x_k^2 &\geq \Delta (x_j, x_k)^2 + a_{jk}^2 \quad (34) \\
\Delta p_j^2 \Delta p_k^2 &\geq \Delta (p_j, p_k)^2 + c_{jk}^2 \quad (35) \\
\Delta x_j^2 \Delta p_k^2 &\geq \Delta (x_j, p_k)^2 + b_{jk}^2. \quad (36)
\end{align*}

In particular, if \( \Omega = \varepsilon J, \varepsilon > 0 \), these conditions again reduce to the Robertson–Schrödinger type inequalities
\[ \Delta x_j^2 \Delta p_k^2 \geq \Delta (x_j, p_k)^2 + \varepsilon^2. \]

In [16, 17] we have given an interpretation of the uncertainty principle (both quantum, and classical) in terms of the topological notion of symplectic capacity (also see the review paper by de Gosson and Luef [20]). This approach seems promising since it allows to recast the Robertson–Schrödinger inequalities in a topological form, invariant even under non-linear canonical transformations. This is an extension of the fact that the Robertson–Schrödinger inequalities are invariant under linear symplectic transformations.

5. Heisenberg Group and Operators

In this Section we show that the Heisenberg group is essentially a classical object associated to the notion of phase of a Lagrangian manifold (see de Gosson [14, 15]).

5.1. The Heisenberg group

The Heisenberg group \( \mathbb{H}_n \) is the space \( \mathbb{R}^{2n+1} \) equipped with the product
\[ (z_1, t_1) \circ (z_0, t_0) = (z_1 + z_0, t_0 + t_1 + \frac{i}{2} \sigma(z_0, z_1)). \quad (37) \]

It is intimately related to the Heisenberg–Weyl operators
\[ \hat{T}(z_0) = e^{-\frac{i}{\hbar} \sigma(z, z_0)}, \quad \sigma(z, z_0) = p_0 x - x_0 (-i\hbar \nabla_x) \quad (38) \]
whose explicit action on a function \( \psi(x) \) is given by
\[ \hat{T}(z_0) \psi(x) = e^{\frac{i}{\hbar} (p_0 x - \frac{1}{2} p_0 x_0)} \psi(x - x_0) \quad (39) \]
(see e.g. de Gosson [15, 18], Littlejohn [29]). The Heisenberg–Weyl operators do not commute:
\[ \hat{T}(z_1) \hat{T}(z_0) = e^{-\frac{i}{\hbar} \sigma(z_0, z_1)} \hat{T}(z_0) \hat{T}(z_1) \quad (40) \]
as immediately follows from the property
\[ \hat{T}(z_0 + z_1) = e^{-\frac{i}{\hbar} \sigma(z_0, z_1)} \hat{T}(z_0) \hat{T}(z_1) \quad (41) \]
which leads to the definition of the Heisenberg group (see Refs. [23, 24]). This traditional construction gives the impression that some strange algebraic structure which is supposed to be the quintessence of QM abruptly emerges! We will see that this is not the case, by showing that formula (39) has a perfectly classical interpretation in terms of the notion of phase.
5.2. The phase of a Lagrangian manifold

A venerable topic in CM is the Hamilton–Jacobi equation

\[ \frac{\partial \varphi}{\partial t} + H(x, \nabla_x \varphi) = 0, \quad \varphi(x, 0) = \varphi_0(x); \]  \hspace{1cm} (42)

for instance, if \( H \) is of the classical type kinetic energy plus potential it takes the more familiar form (see Refs. [4, 12]):

\[ \frac{\partial \varphi}{\partial t} + \sum_{j=1}^{n} \frac{1}{2m} (\nabla_x \phi)^2 + U(x) = 0, \quad \varphi(x, 0) = \varphi_0(x). \]

The Hamilton–Jacobi equation allows, among other things, to solve explicitly Hamilton’s equations for a wide class of systems. The solution (which we assume exists for \( t \) in some open interval surrounding the initial value \( t = 0 \)), is constructed as follows. Let us denote by \( V_0 \) the graph of the equation \( p = \nabla_x \varphi_0(x) \) (i.e. the set of all points \((x, \nabla_x \varphi(x))\) where \( x \) takes its values in the domain of \( \varphi \)). \( V_0 \) is a \( n \)-dimensional submanifold of phase space, in fact a Lagrangian manifold. This means that the symplectic product of any two tangent vectors \( Z, Z' \) at a same point \( z \) of \( V_0 \) is equal to zero. In [14] we have called the function \( \varphi_0 \) the phase of the manifold \( V_0 \); it is uniquely determined up to a constant. Lagrangian manifolds are the fundamental geometric objects of classical mechanics in its Hamiltonian formulation (see Refs [1, 4, 47, 48]; for instance, the invariant tori associated with completely integrable systems are Lagrangian manifolds (albeit not of the type \( V_0 \) above).

Consider now the time-dependent flow \((f_t^H)\) determined by the Hamilton equations for \( H \), and set \( V_t = f_t^H V_0 \). If \(|t|\) is sufficiently small, the manifold \( V_t \) will again be a graph, that is, there exists a function \( \varphi(x, t) \) (the phase of \( V_t \)) such that the equation \( p = \nabla_x \varphi(x, t) \) represents \( V_t \) (if \( t \) is too large, bending of the original manifold \( V_0 \) may lead to the appearance of several points on \( V_t \) having same projection on \( x \)-space). The phase \( \varphi = \varphi(x, t) \) (which is precisely the solution of the Hamilton–Jacobi equation \((42)\)) is explicitly calculated as follows: since \( V_t \) is a graph, for each value \( x \) in the domain of \( \varphi \) there exits a unique \( p \) such that \( p = \nabla_x \varphi(x, t) \), and the phase space point \((x, p)\) is obtained from a unique point \((x(0), p(0))\) in \( V_0 \) by the flow \((f_t^H)\). Denoting by \( \Gamma_t \) the arc of phase space trajectory joining \((x(0), p(0))\) in \( V_0 \) to \((x, p)\) in \( V_t \) the phase \( \varphi \) of \( V_t \) satisfies

\[ \varphi(x, t) = \varphi_0(x(0)) + \int_{\Gamma_t} \alpha_H \]  \hspace{1cm} (43)

where

\[ \alpha_H = pdx - Hdt \]  \hspace{1cm} (44)

is the Poincaré–Cartan invariant (see Refs. [1, 4, 15]). We now remark that since both \( V_0 \) and \( V_t \) are graphs, the datum of \( x(0) \) uniquely determines \( x(t) \) and hence the formula \((x, p) = f_t^H(x(0), p(0))\) can be inverted, expressing \( x(0) \) as a function of \( x(t) \); we will write this as

\[ x(0) = (f_t^H)^{-1}(x) \]  \hspace{1cm} (45)

hence the solution \( \varphi(x, t) \) is given by

\[ \varphi(x, t) = \varphi_0((f_t^H)^{-1}(x)) + \int_{\Gamma_t} \alpha_H. \]  \hspace{1cm} (46)

It is customary in theoretical statistical mechanics to identify a physical system with a pair \( (\rho, V) \) where \( V \) is a Lagrangian manifold and \( \rho \) a measure (probability density) carried by that manifold. At time \( t = 0 \) we thus have a pair \( (\rho_0, V_0) \) where the Lagrangian manifold \( V_0 \) is the
the graph of \( p = \nabla_x \varphi_0(x) \). Since \( V_0 \) is uniquely determined by its phase \( \varphi_0 \) up to an inessential additive constant, we can without any loss of generality identify that system with the pair \((\rho_0, \varphi_0)\). Under Hamiltonian evolution \((\rho_0, \varphi_0)\) will become \((\rho, \varphi)\) where \( \rho \) is obtained by “push-forward” of \( \rho_0 \), that is \( \rho(z) = \rho_0((f^z_0)^{-1}z) \) and \( \varphi = \varphi(x, t) \) is given by formula (46). We now remark that since the variable \( z = (x, p) \) in \( \rho(z) \) is constrained by the condition \( p = \nabla_x \varphi \) the measure \( \rho \) is actually only a function of \( x \) so we may write \( \rho_0(z) = \rho_{0,X}(x) \) and

\[
\rho(z, t) = \rho_X(x, t) = \rho_{0,X}[(f^t_0)^{-1}x].
\]

We next remark that the datum of the pair of real functions \((\rho_X, \varphi)\) defined on \( \mathbb{R}^n \) is mathematically equivalent to the datum of the complex function

\[
\mu(x) = \rho_X(x)e^{i\phi(x)}
\]

and the transformation formulas (45), (47) can then collectively be written in the compact form:

\[
\mu(x, t) = e^{i\int_{t_1}^t \alpha_H}\mu_0[(f^t_0)^{-1}x].
\]

5.3. Displacement operators

Let us now calculate \((\rho, \varphi)\) in the particular case where \( H \) is the Hamiltonian

\[
H^{z_0}(z) = \sigma(z, z_0) = px_0 - p_0x;
\]

here \( z_0 = (x_0, p_0) \) will play the role of a translation vector. The function \( H^{z_0} \) is called the “displacement Hamiltonian”, because the solutions of the corresponding Hamilton equations

\[
\dot{x} = x(t) = x(0) + tx_0, \quad \dot{p} = p(t) = x(0) + tp_0
\]

so the flow \((f^t_0)\) determined by \( H^{z_0} \) is just the phase space translation with vector \( t(z_0) = (tx_0, tp_0) \). Let us calculate \( \varphi(x, t) \). The arc of trajectory \( \Gamma_t \) is here the line segment joining \((x(0), p(0))\) to \((x(0) + tx_0, p(0) + tp_0)\) and the Hamiltonian is \( \sigma(z(0) + t(z_0), z_0) = \sigma(z(0), z_0) \) along \( \Gamma_t \) so that

\[
\alpha_H = [(p(0) + tp_0)x_0 - \sigma(z(0), z_0)] dt
\]

A straightforward calculation then shows that we have

\[
\int_{\Gamma_t} pdx - H^{z_0}dt = tp_0x(0) + \frac{1}{2}t^2p_0x_0
\]

and hence, applying formula (46), the phase \( \varphi(x, t) \) is given by

\[
\varphi(x, t) = \varphi_0(x(0)) + tp_0x(0) + \frac{1}{2}t^2p_0x_0
\]

that is, using the relation \( x = x(0) + tx_0 \),

\[
\varphi(x, t) = \varphi_0(x - tx_0) + tp_0x - \frac{1}{2}t^2p_0x_0.
\]

The measure \( \rho_t(z) = \rho_0((f^z_0)^{-1}z) \) is also easily calculated: since \( f^z_0 \) is the translation \( z \rightarrow z + tz_0 \) we immediately get

\[
\rho_t(z) = \rho_0(z - tz_0).
\]

Let us now choose \( t = 1 \) and set \( T(z_0) = f^z_0 \): it is the phase space translation with vector \( z_0 = (x_0, p_0) \). Under the action of \( T(z_0) \) the complex function \( \mu_0(x) = \rho_{0,X}(x)e^{i\phi_0(x)} \) becomes a new function

\[
\tilde{T}(z_0)\mu_0(x) = \mu_0(x - x_0)e^{i\int_{t_1}^t \alpha_H}
\]
where \( \Gamma_1 = \Gamma_{t=1} \). In view of Eqn. (53) we have
\[
\int_{\Gamma_1} \alpha_H = p_0 x - \frac{1}{2} p_0 x_0
\]  
and hence
\[
\tilde{T}(z_0) \mu_0(x) = \mu_0(x - x_0) e^{i(p_0 x - \frac{1}{2} p_0 x_0)}
\]  
and this is exactly formula (39) giving the action of the Heisenberg–Weyl operator with \( \hbar = 1 \) on the function \( \psi = \mu_0 \)!

From this we immediately deduce without any calculation at all that the operators \( \tilde{T}(z_0) \) satisfy the relations
\[
\tilde{T}(z_1) \tilde{T}(z_0) = e^{-i \sigma(z_0,z_1)} \tilde{T}(z_0) \tilde{T}(z_1)
\]  
and
\[
\tilde{T}(z_0 + z_1) = e^{-\frac{i}{2} \sigma(z_0,z_1)} \tilde{T}(z_0) \tilde{T}(z_1).
\]

Let us however check, for the sake of completeness, the relation (57) by a calculation of classical phases (the relation (58) will immediately follow). Let \( z_1 \) be a second fixed point in phase space; we set out to compare the states
\[
\tilde{T}(z_0 + z_1)(\rho_0, \varphi_0) \quad \text{and} \quad \tilde{T}(z_0) \tilde{T}(z_1)(\rho_0, \varphi_0).
\]

Both Lagrangian manifolds \( \tilde{T}(z_0 + z_1)V_0 \) and \( \tilde{T}(z_0) \tilde{T}(z_1)V_0 \) are identical, being both obtained by a translation with vector \( z_0 + z_1 \), but these two states are not the same! First, replacing \( (x_0, p_0) \) in Eqn. (55) with \( (x_0 + x_1, p_0 + p_1) \) the phase of \( T(z_0 + z_1)V_0 \) is
\[
\varphi(z_0 + z_1)(x) = \varphi_0(x) + (p_0 + p_1)[x + \frac{1}{2}(x_0 + x_1)].
\]

The phase \( \varphi(z_1,z_0)(x) \) of \( \tilde{T}(z_1) \tilde{T}(z_0)V_0 \) is obtained by adding to the expression (55) of \( \varphi(z_0)(x) \) the integral of the Poincaré form \( p \mathrm{d}x - \sigma(z, z_1) \mathrm{d}t \) along the line segment \( \Gamma_1' \) joining \( (x + x_0, p + p_0) \) to \( (x + x_0 + x_1, p + p_0 + p_1) \); this is
\[
\int_{\Gamma_1'} p \mathrm{d}x - \sigma(z, z_1) \mathrm{d}t = p_1(x + x_0) + \frac{1}{2} p_1 x_1
\]
and hence
\[
\varphi(z_1,z_0)(x) = \varphi_0(x) + p_0 x + p_1(x + x_0) + \frac{1}{2} (p_0 x_0 + p_1 x_1)
\]
A straightforward calculation shows that the phases given by Eqns. (59) and (60) differ by the quantity
\[
\varphi(z_1,z_0)(x) - \varphi(z_0 + z_1)(x) = -\frac{1}{2} \sigma(z_0, z_1)
\]
(it is, up to the sign, the area of the triangle with sides the vectors \( z_0 \), and \( z_1 \); see Fig. 3 in Littlejohn [29]); this immediately implies formula (57).

The important thing to remember is that everything in our construction is entirely classical, without any reference to the canonical commutation relations which are historically at the heart of the definition of the Heisenberg–Weyl operators and of the Heisenberg group.
6. Discussion and (Unconcluding) Remarks
Here are a few facts that could be used to illustrate the topics developed in this paper. It is usually assumed that the quantum wave-particle duality can have no counterpart in classical physics. It is also said that the probabilistic aspects of quantum dynamics are intrinsic and to have no possible relation with underlying unresolved dynamical phenomena. However, Yves Couder and his collaborators (Couder et al. [7]) have recently investigated the properties of a droplet bouncing on a vibrated bath. In this system, a propagative wave-particle association appears when the droplet couples to the surface wave it excites. Surprisingly, both a form of uncertainty and a form of quantization are observed. In these experiments you see textbook quantum mechanics, but everything is perfectly classical! It would be interesting to interpret Couder’s results using our constructs. We mention that in a recent work [5] Bartlett et al. study the reconstruction of a version of QM from classical considerations; this work where the notion of epistemic restriction plays a key role is certainly closely related to our ideas. In [46] Gerand ’t Hooft contends that “...Quantum mechanics is ‘emergent’ if a statistical treatment of large scale phenomena in a locally deterministic theory requires the use of quantum operators”. The fact that Schrödinger’s equation is (mathematically speaking) a classical equation contradicts in some sense this statement, since we are able to associate operators to classical observables, and these operators have a perfectly classical meaning unless one decides, a posteriori, to give their characteristics (e.g., their eigenvalues) a precise physical meaning.

To conclude, let us cite Stephen Adler [2]: “...quantum theory is not a complete, final theory, but is in fact an emergent phenomenon arising from a deeper level of dynamics...” (also see Bateson [6] who makes similar claims).

References
[1] Abraham R and Marsden J E 1978 Foundations of Mechanics. The Benjamin/Cummings Publishing Company, 2nd edition
[2] Adler S I 2004 Quantum Theory as an Emergent Phenomenon. The Statistical Mechanics of Matrix Models as the Precursor of Quantum Field Theory. Cambridge University Press
[3] Anderson P W 4 August 1972 More Is Different. Science, Vol. 177 No. 4047 393–396
[4] Arnold V I 1989 Mathematical Methods of Classical Mechanics. Graduate Texts in Mathematics, 2nd edition, Springer-Verlag
[5] Bartlett S D , Rudolph T and Spekkens R W Reconstruction of Gaussian quantum mechanics from Liouville mechanics with an epistemic restriction. arXiv:1111.5057
[6] Bateson R D 2010 Causal space-Time. Emergent Physics from Causality. CreateSpace ; extended version to be published by World Scientific.
[7] Couder Y, Boudaoud A, Protière S, Moukhtarb and Fort J E Walking droplets: a form of wave-particle duality at macroscopic level? , DOI: 10.1051/epn/2010101
[8] Dias N C, de Gosson M, Luef F and Prata J N 2010 A deformation quantization theory for noncommutative quantum mechanics, J. Math. Phys. 51 072101
[9] Dirac P A M 1999 The Principles of Quantum Mechanics. Oxford Science Publications, 4th revised edition
[10] Fort E, Eddi A , Boudaoud A, Moukhtar J and Couder Y 2010 Path-memory induced quantization of classical orbits. PNAS, 107(41), 17515–17520
[11] Giulini D 2003 That Strange Procedure Called Quantisation. Lect. Notes Phys. 361 17–40
[12] Goldstein H 1959 Classical Mechanics. Addison–Wesley (1950), 2nd edition, (1980), 3d edition (2002)
[13] de Gosson M 2001 The Principles of Newtonian and Quantum Mechanics: The need for Planck’s constant, h. With a foreword by Basil Hiley. Imperial College Press
[14] de Gosson M 2004 On the notion of phase in mechanics, J. Phys. A: Math. Gen., 37(29), 7297–7314
[15] de Gosson M 2006 Symplectic Geometry and Quantum Mechanics. Birkhäuser, Basel, series “Operator Theory: Advances and Applications” (subseries: “Advances in Partial Differential Equations”), Vol. 166
[16] de Gosson M 2009 The Symplectic Camel and the Uncertainty Principle: The Tip of an Iceberg? Found. Phys. 39 191–214
[17] de Gosson M 2010 On the Use of Minimum Volume Ellipsoids and Symplectic Capacities for Studying Classical Uncertainties for Joint Position-Momentum Measurements. J. Stat. Mech. P11005
[18] de Gosson M 2011 Symplectic Methods in Harmonic Analysis and in Mathematical Physics. Birkhäuser, Basel
[19] de Gosson M and Hiley B J 2011 Imprints of the Quantum World in Classical Mechanics. *Found. Phys.* **41**, 1415–1436
[20] de Gosson M and Luef F 2009 Symplectic Capacities and the Geometry of Uncertainty: the Irruption of Symplectic Topology in Classical and Quantum Mechanics. *Physics Reports* **484**, 131–179
[21] Groenewold H J 1946 On the principles of elementary quantum mechanics. *Physica* **12**, 405–460
[22] Gromov M 1985 Pseudoholomorphic curves in symplectic manifolds. *Invent. Math.* **82**, 307–47 (1985)
[23] Guillemin V and Sternberg S 1978 Geometric Asymptotics. *Math. Surveys Monographs* **14**, *Amer. Math. Soc.*, Providence R.I.
[24] Guillemin V and Sternberg S 1984 Symplectic Techniques in Physics. Cambridge University Press, Cambridge, Mass.
[25] Hall M J W and Reginatto M 2002 Schrödinger equation from an exact uncertainty principle. *J. Phys. A: Math. Gen.* **35**, 3289
[26] van Hove L 1951 Sur le Problème des Relations entre les Transformations Unitaires de la Mécanique Quantique et les Transformations Canoniques de la Mécanique Classique. *Mém. Acad. Roy. Belg.* **26**(6)
[27] van Hove L 1952 Sur certaines représentations unitaires d’un groupe fini de transformations. *Mém. Acad. Roy. Belg. Classe des Sci.* **26**(1)
[28] John F 1948 Extremum problems with inequalities as subsidiary conditions, Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, Interscience Publishers, Inc., New York, N.Y. 187–204
[29] Littlejohn R G 1986 The semiclassical evolution of wave packets. *Physics Reports* **138**(4–5), 193–291
[30] Lopuhäa H P and Rousseeuw P J 1991 Breakdown points of affine equivariant estimators of multivariate location and covariance matrices, *The Annals of Statistics* **19**(1), 229–248
[31] Mackey GW 1998 The Relationship Between Classical and Quantum Mechanics. In Contemporary Mathematics **214**, *Amer. Math. Soc.*, Providence, RI
[32] Nelson E 1966 Derivation of the Schrödinger Equation from Newtonian Mechanics. *Phys. Rev. A*, **150**(4):6, 1079–1085
[33] Peres A 1993 Quantum Theory: Concepts and Methods. Kluwer Academic Publishers
[34] Polterovich L 2001 The Geometry of the Group of Symplectic Diffeomorphisms. Lectures in Mathematics, Birkhäuser
[35] Reed M., Simon B 1972 Methods of Modern Mathematical Physics. Academic Press, New York
[36] Robertson H P 1929 The uncertainty principle. *Phys. Rev.* **26**(2), 15–18
[37] Rousseeuw P J and Leroy A M, 1987 Robust regression and outlier detection, John Wiley & Sons, New York
[38] Scheeres D J, Hsiao F Y, Park R S, Villac B F, and Maruskin J M 2006 Fundamental Limits on Spacecraft Orbit Uncertainty and Distribution Propagation, *Journal of the Astronautical Sciences* **54**, 505–523
[39] Schrödinger E 1930 Zum Heisenbergschen Unschärfeprinzip, Berliner Berichte 296–303 [English translation: Angelow, A., Batoni, M.C.: About Heisenberg Uncertainty Relation. *Bulg. Journal of Physics*, **26**, nos.5/6, 193–203 (1999), and http://arxiv.org/abs/quant-ph/9903100]
[40] Schrödinger E 1926 Quantisierung als Eigenwertproblem, *Ann. der Physik*, **384**, 361–376
[41] Van Aelst S and Rousseeuw P J, 2009 Minimum volume ellipsoid, Wiley Interdisciplinary Reviews: *Computational Statistics* **1**(1) 71–82
[42] Stein EM 1993 Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals, Princeton University Press
[43] Shubin M A 1978 Pseudodifferential Operators and Spectral Theory, Springer-Verlag, (1987) [original Russian edition in Nauka, Moskva ]
[44] Stone M H 1930 Linear transformations in Hilbert space, III: operational methods and group theory. *Proc. Nat. Acad. Sci.* U.S.A., 172–175
[45] ‘t Hooft G 2007 Emergent Quantum Mechanics and Emergent Symmetries. AIP Conf. Proc. **957**, 154-163 ; http://arxiv.org/abs/0707.4568
[46] Tulczyjew W 1976 *C.R. Acad. Sci.* Paris **283**, 15–18
[47] Tulczyjew W M 1977 The Legendre transformation. *Ann. Inst. H. Poincaré, Sect. A.*, **27**(1), 101–114
[48] Williamson J 1936 On the algebraic problem concerning the normal forms of linear dynamical systems, *Amer. J. of Math.* **58**, 141–163.
[49] Wong M W 1998 Weyl Transforms. Springer
[50] Woodhouse N M 1991 *Geometric Quantization*, second edition. Oxford Science Publications.