Quantum Harmonic Analysis of the Density Matrix

Maurice A. de Gosson

Faculty of Mathematics (NuHAG), University of Vienna, Vienna, Austria. E-mail: maurice.de.gosson@univie.ac.at

Editors: *Eliahu Cohen & Danko Georgiev*

Article history: Submitted on April 17, 2018; Accepted on July 5, 2018; Published on September 26, 2018.

We will study rigorously the notion of mixed states and their density matrices. We will also discuss the quantum-mechanical consequences of possible variations of Planck's constant h. This review has been written having in mind two readerships: mathematical physicists and quantum physicists. The mathematical rigor is maximal, but the language and notation we use throughout should be familiar to physicists.
Quanta 2018; 7: 74–110.

1 Preface

1.1 Quantum harmonic analysis

When writing a review on the density matrix you have several options. You can use the standard approach found in physics graduate textbooks, but then you cannot go very far, not only because the important mathematical tools are lacking, but also because the arguments are often flawed and incorrect: most of them strictly speaking only apply when the underlying Hilbert space is finite-dimensional (I am thinking here, among other things, about *trace taking* procedures for obtaining the averages of observables which are most of the time mathematically undefendable in infinite dimension in the absence of extra conditions on these observables). Then, at the other extreme, you can use the C^*-algebraic approach. This is certainly the most beautiful, elegant, and intellectually satisfying way to introduce the mixed states of quantum mechanics and their density matrices; after having listed some definitions and properties from the theory of operator algebras the Gelfand–Naimark–Segal construction then allows the passage from this abstract theory to the usual Hilbert-space picture of quantum mechanics. Unfortunately, this approach requires lots of preparatory work to become accessible, and it is often difficult for the uninitiated to grasp the physical meaning of the tools that are used. I have chosen here a third way, which has the advantage of being both mathematically rigorous and intuitive. It consists in using functional analysis together with tools from harmonic analysis on phase space, especially Weyl–Wigner–Moyal theory (which is one way to *quantize* classical observables). We will call this way of doing things the *quantum harmonic analysis* approach. One of its advantages is that it fully justifies from a mathematical viewpoint the introduction of the Wigner function of a density matrix, which otherwise appears as an *ad hoc* object pulled out of thin air. There is another major advantage to this approach: it highlights the sensitivity of the theory of density matrices on the choice of the value of Planck's constant, and this is precisely one of the main themes we want to address. We will mostly deal with continuous-variable systems with an infinite-dimensional Hilbert space described by observables with continuous eigenspectra.

1.2 Variability of Planck's constant

One topic we address in this review is the sensitivity of quantum states to possible variations of Planck's constant. Physically this is a very controversial question. Setting aside for a moment the debate on whether Planck's constant can vary or not, consider the following situation: we have an unknown quantum state, on which we perform a quorum of measurements in order to determine its density matrix $\widehat{\rho}$. Now, one should be aware of the fact that this density matrix will not be determined directly by these experiments; what one does is to measure by, say, a homodyne quantum tomography, certain properties of that system. Quantum homodyne tomography originates from the observation by Vogel and Risken [1] that the probability distributions determined by homodyne detection are just the Radon transforms of the Wigner function of the density matrix, and that the latter allows us to infer the density matrix using the Weyl correspondence. The method works as follows: suppose now that we have been able to determine a statistical function $\rho(x,p)$ of the position and momentum variables (called *quadratures* in this context), yielding the properties of the quantum system under investigation. If we now identify this function $\rho(x,p)$ with the Wigner distribution

$$\rho(x,p) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{-\frac{i}{\hbar}py}\langle x+\tfrac{1}{2}y|\widehat{\rho}|x-\tfrac{1}{2}y\rangle d^n y \qquad (1)$$

of the corresponding density matrix, this relation can be inverted and yields $\widehat{\rho}$; mathematically speaking $\widehat{\rho}$ is just, up to a factor, the Weyl operator corresponding to the *classical* observable $\rho(x,p)$, which can be written for instance as

$$\widehat{\rho}\psi(x) = \int \rho(x_0,p_0)e^{\frac{2i}{\hbar}p_0(x-x_0)}\psi(2x_0-x)d^n p_0 d^n x_0. \quad (2)$$

Now, two essential observations. The first is that when using this procedure we assume quite explicitly that we are using the Weyl–Wigner–Moyal formalism: we identify the function $\rho(x,p)$ with the Wigner function, and *quantize* it thereafter using the Weyl transform. This is very good, of course, but one should keep in mind that there are other possible representations in quantum mechanics; a physically very interesting one is for instance the Born–Jordan quantization scheme [2,3]. We will not investigate the implications of such a choice in this review, but we want here to emphasize another problem. Even if we place ourselves in the Weyl–Wigner–Moyal framework we are tacitly assuming that $\hbar = h/2\pi$ has a fixed value in time and space. If it happens that Planck's constant h has another value, h', at another location or at another time, formulas (1) and (2) would have to be replaced with the different expressions

$$\rho'(x,p) = \left(\tfrac{1}{2\pi\hbar'}\right)^n \int e^{-\frac{i}{\hbar'}py}\langle x+\tfrac{1}{2}y|\widehat{\rho}'|x-\tfrac{1}{2}y\rangle d^n y \quad (3)$$

and

$$\widehat{\rho}'\psi(x) = \int \rho(x_0,p_0)e^{\frac{2i}{\hbar'}p_0(x-x_0)}\psi(2x_0-x)d^n p_0 d^n x_0. \quad (4)$$

1.3 Notation

Let σ be the standard symplectic form on phase space $\mathbb{R}^{2n} \equiv \mathbb{R}^n_x \times \mathbb{R}^n_p$: by definition it is the mapping which to the pairs of vectors $z = (x,p)$ and $z' = (x',p')$ associates the number

$$\sigma(z,z') = \sum_j p_j x'_j - p'_j x_j = px' - p'x. \qquad (5)$$

The symplectic form can be conveniently written in matrix form as

$$\sigma(z,z') = (z')^T Jz \ , \ J = \begin{pmatrix} 0_{n\times n} & I_{n\times n} \\ -I_{n\times n} & 0_{n\times n}\end{pmatrix} \qquad (6)$$

(z and z' being here viewed as column vectors; J is the *standard symplectic matrix*).

The inner product of two vectors $\psi, \phi \in L^2(\mathbb{R}^n)$ is given by

$$\langle\psi|\phi\rangle = \int \psi^*(x)\phi(x)d^n x \qquad (7)$$

with $d^n x = dx_1 \cdots dx_n$. The associated norm is

$$\|\psi\| = \sqrt{\langle\psi|\psi\rangle}. \qquad (8)$$

In addition to the usual $\hbar$-dependent (unitary) Fourier transform

$$F\psi(p) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{-\frac{i}{\hbar}px}\psi(x)d^n x \qquad (9)$$

on $\mathbb{R}^n$ we will use the symplectic Fourier transform on $\mathbb{R}^{2n}$: it is the transformation F_σ which takes a square integrable function (or more generally a tempered distribution) a on phase space $\mathbb{R}^{2n}$ to the function (or tempered distribution)

$$F_\sigma a(z) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{-\frac{i}{\hbar}\sigma(z,z')}a(z')d^{2n}z'. \qquad (10)$$

That F_σ is just an elementary modification of the $\hbar$-Fourier transform on $\mathbb{R}^{2n}$ given by

$$Fa(z) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{-\frac{i}{\hbar}zz'}a(z')d^{2n}z' \qquad (11)$$

is easy to see: since $\sigma(z,z') = (z')^T Jz = Jz\cdot z'$ (J the standard symplectic matrix) we have $F_\sigma a(z) = Fa(Jz)$. The symplectic Fourier transform is that it is its own inverse: $F_\sigma^2 = F_\sigma F_\sigma$ is the identity (i.e. the symplectic Fourier transform is involutive) and it satisfies the modified Plancherel identity

$$\int a(z)b(z)d^{2n}z = \int F_\sigma a(z)F_\sigma b(-z)d^2 z. \qquad (12)$$

2 Introduction

We begin by recalling informally the basic definitions from the theory of quantum states; since the content of this section is common knowledge we do not provide any particular references (we are following here the terminology and the exposition in Peres [4] with some modifications). The formalism of density operators and matrices was introduced by John von Neumann [5] in 1927 and independently, by Lev Landau and Felix Bloch in 1927 and 1946 respectively. Ugo Fano was one of the first to put the theory of the density matrix in a rigorous form in his review paper [6].

2.1 Quantum states and observables

2.1.1 Pure and mixed states; maximal tests

A quantum system is said to be in a *pure* state if we have complete knowledge about that system, meaning we know exactly which state it is in. Pure states can be prepared using *maximal tests* [4, §2-3]: suppose we are dealing with a quantum system and let N be the maximum number of different outcomes that can be obtained in a test of that system. If such a test has exactly N different outcomes, it is called a maximal test. The quantum system under consideration is in a pure state if it is prepared in such a way that it certainly yields a predictable outcome in that maximal test, the outcomes in any other test having well-defined probabilities which do not depend on the procedure used for the preparation. A pure state can thus be identified by specifying the complete experiment that characterizes it uniquely (Fano [6]). One usually writes a pure state using Dirac's ket notation $|\psi\rangle$; for all practical purposes it is convenient to use the wavefunction ψ defined by $\psi(x) = \langle x|\psi\rangle$; it is a normalized element of a certain Hilbert space $\mathcal{H}$, which is usually identified in the case of continuous variables with $L^2(\mathbb{R}^n)$ (the square integrable functions). When doing this the state is identified with the linear span of the function ψ, that is the ray $\mathbb{C}\psi = \{\lambda\psi : \lambda \in \mathbb{C}\}$. It is very important to note that the pure state $|\psi\rangle$ can be identified with the orthogonal projection $\widehat{\rho}_\psi$ of $\mathcal{H}$ on the subspace $\mathbb{C}\psi$. This projection, which is of rank one, is denoted by $|\psi\rangle\langle\psi|$ in quantum mechanics; it is analytically given by the formula

$$\widehat{\rho}_\psi\phi = |\psi\rangle\langle\psi|\phi\rangle \tag{13}$$

where $\langle\psi|\phi\rangle$ is identified with the inner product in $\mathcal{H}$. Most tests are however not maximal, and most preparations do not produce pure states, so we only have partial knowledge of the quantum system under consideration. The information on such a system is less than a maximum, with reference to the lack of a complete experiment with a uniquely predetermined outcome. The state of the system is nevertheless fully identified by any data adequate to predict the statistical results of all conceivable observations on the system [6]. When this is the case we say that the system is in a *mixed state*. Mixed states are classical probabilistic mixtures of pure states; however, different distributions of pure states can generate physically indistinguishable mixed states (this possibility will be discussed later). A quantum mixed state can be viewed as the datum of a set of pairs $\{(\psi_j, \alpha_j)\}$ where ψ_j is a (normalized square integrable) pure state and α_j a classical probability; these probabilities sum up to one: $\sum_j \alpha_j = 1$. A *caveat*: one should not confuse the mixed state $\{(\psi_j, \alpha_j)\}$ with the superposition $\psi = \sum_j \alpha_j\psi_j$ which is a *pure* state!

2.1.2 The density matrix

From a mathematical point of view the quantum state can be advantageously described by a self-adjoint operator on a Hilbert space; this operator is called the density matrix of the quantum system; in the pure state case this operator is just the orthogonal projection (13) on a subspace of the Hilbert space, while the density matrix of a mixed state is

$$\widehat{\rho} = \sum_j \alpha_j\widehat{\rho}_j = \sum_j \alpha_j|\psi_j\rangle\langle\psi_j| \tag{14}$$

where we have set $\widehat{\rho}_j = \widehat{\rho}_{\psi_j}$. It should be kept in mind that the density matrix describes a preparation procedure for an ensemble of quantum systems whose statistical properties correspond to the given preparation procedure (this important aspect will be discussed in more detail below).

Theorem 1. *(i) The density operators on a Hilbert space $\mathcal{H}$ form a convex subset* Dens($\mathcal{H}$) *of the space $\mathcal{B}(\mathcal{H})$ of bounded operators on $\mathcal{H}$. (ii) The extreme points of this set are the rank-one projections, which correspond to the pure states on $\mathcal{H}$.*

Proof. (i) To say that the set of density matrices is convex means that if $\widehat{\rho}_1$ and $\widehat{\rho}_2$ are in Dens($\mathcal{H}$) then so is $\lambda\widehat{\rho}_1 + (1-\lambda)\widehat{\rho}_2$ for all real numbers λ such that $0 \leq \lambda \leq 1$. Let $\{(\psi_j, \alpha_j) : j \in K\}$ and $\{(\phi_j, \beta_j) : j \in L\}$ be two mixed states. Relabeling if necessary the indices we may assume that the sets K and L are disjoint: $K \cap L = \emptyset$. The corresponding density matrices are

$$\widehat{\rho}_1 = \sum_{k\in K} \alpha_k|\psi_k\rangle\langle\psi_k| \ , \ \widehat{\rho}_2 = \sum_{\ell\in L} \beta_\ell|\phi_\ell\rangle\langle\phi_\ell| \tag{15}$$

and hence we have by linearity

$$\lambda\widehat{\rho}_1 + (1-\lambda)\widehat{\rho}_2 = \sum_{j\in K\cup L} \gamma_j|\chi_j\rangle\langle\chi_j| \tag{16}$$

where $\chi_j = \psi_j$, $\gamma_j = \lambda\alpha_j$ if $j \in K$ and $\chi_j = \phi_j$, $\gamma_j = (1-\lambda)\beta_j$ if $j \in L$. That $\widehat{\rho} = \lambda\widehat{\rho_1} + (1-\lambda)\widehat{\rho_2}$ is a mixed state now follows from the equality

$$\sum_{j\in K\cup L} \gamma_j = \lambda \sum_{j\in K} \alpha_j + (1-\lambda)\sum_{j\in L} \beta_j = 1. \tag{17}$$

(ii) That a pure state really is *pure*, i.e. that it can never be represented as a mixed state is easily seen using the following algebraic argument: assume that $\widehat{\rho_\psi} = |\psi\rangle\langle\psi|$ can be rewritten as a sum $\sum_j \alpha_j |\psi_j\rangle\langle\psi_j|$ with $\alpha_j \geq 0$ and $\sum_j \alpha_j = 1$. In fact, discarding the terms with $\alpha_j = 0$ we may assume that $\alpha_j > 0$ for all indices j. Let now $(\mathbb{C}\psi)^\perp$ be the subspace of $\mathcal{H}$ orthogonal the ray $\mathbb{C}\psi$: it consists of all vectors ϕ in $\mathcal{H}$ such that $\langle\psi|\phi\rangle = 0$. For every $\phi \in (\mathbb{C}\psi)^\perp$ we have $\langle\phi|\widehat{\rho_\psi}|\phi\rangle = |\langle\psi|\phi\rangle|^2 = 0$ and hence also

$$\langle\phi|\widehat{\rho_\psi}|\phi\rangle = \sum_j \alpha_j |\langle\phi|\psi_j\rangle|^2 = 0; \tag{18}$$

since we are assuming that $\alpha_j > 0$ this equality implies that we must have $\langle\phi|\psi_j\rangle = 0$ for all $\phi \in S_\psi$ and every ψ_j, hence $\psi_j \in ((\mathbb{C}\psi)^\perp)^\perp$. The orthogonality relation for subspaces being reflexive we have $((\mathbb{C}\psi)^\perp)^\perp = \mathbb{C}\psi$, which means that each ψ_j belongs to the ray $\mathbb{C}\psi$, that is $\psi_j = \lambda\psi$ for some complex number λ with $|\lambda| = 1$; the vectors ψ_j thus define the state $\langle\psi|$. It follows from this argument that the pure state density matrices are the extreme points of Dens($\mathcal{H}$). This means that if $\widehat{\rho_\psi}$ is a pure state density matrix, then the relation $\widehat{\rho_\psi} = \lambda\widehat{\rho_1} + (1-\lambda)\widehat{\rho_2}$ with $\lambda \neq 0$ and $\lambda \neq 1$ implies $\widehat{\rho_\psi} = \widehat{\rho_1} = \widehat{\rho_2}$. That this is the case immediately follows from the argument above. □

A central problem the theory of mixed states is the *separability problem*. A density matrix $\widehat{\rho}$ on a Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ is called *separable* if it can be written as a sum

$$\widehat{\rho} = \sum_j \alpha_j \widehat{\rho}_j^A \otimes \widehat{\rho}_j^B \tag{19}$$

with $\alpha_j \geq 0$ and $\sum_j \alpha_j = 1$, the $\widehat{\rho}_j^A$ (resp. $\widehat{\rho}_j^B$) being density operators on $\mathcal{H}_A$ (resp. $\mathcal{H}_B$). If the state is not separable, it is called an entangled state. Peres [7] has given a necessary condition for separability using the notion of partial transpose and Horodecki *et al.* [8, 9] have shown that this condition is also sufficient in some cases; also see Leinaas *et al.* [10]. Separability implies that the Wigner distribution must go over to a Wigner distribution under the partial mirror reflection

$$(x_A, x_B, p_A, p_B) \longmapsto (x_A, x_B, p_A, -p_B) \tag{20}$$

that is, there must exist a Wigner distribution ρ' such that

$$\rho(x_A, x_B, p_A, -p_B) = \rho'(x_A, x_B, p_A, p_B). \tag{21}$$

Outside this partial result, the general case is open and lures prominent researchers in mathematics, quantum physics, and physical chemistry.

2.1.3 The C^*-algebraic approach

The notion of quantum state and density matrices can be very concisely described using the language of C^*-algebras; this approach is useful when one wants to give a rigorous axiomatic description of the theory of quantum systems. Its inception goes back to early work by John von Neumann on operator algebras around 1930; the modern theory of C^*-algebras was developed one decade later by Gel'fand and Naimark. We recommend Rieffel's paper [11] for a survey of quantization using this approach. Let $\mathcal{A}$ be a non-commutative unital C^*-algebra (*i.e.* a unital complex Banach algebra equipped with an isometric involution $A \longmapsto A^*$ compatible with complex conjugation, and such that $\|A^*A\| = \|A\|^2$). If $A = A^*$ we call A an observable; the set of observables is denoted by $\mathcal{O}$. If in addition the spectrum of $A \in \mathcal{O}$ consists of numbers $\lambda \geq 0$ we say that A is positive ($A \geq 0$). A linear map $\rho : \mathcal{O} \longrightarrow \mathbb{C}$ is positive ($\rho \geq 0$) if $A \geq 0$ implies $\rho(A) \geq 0$; if in addition $\rho(1) = 1$ then S is called a *quantum state*; quantum states form the state space $\mathcal{S}(\mathcal{A})$ of $\mathcal{A}$; it is a convex cone whose extreme points are the *pure states* while the other elements of $\mathcal{S}(\mathcal{A})$ are *mixed states*. Now, the Gel'fand–Naimark theorem [12, 13] implies that for every unital C^*-algebra $\mathcal{A}$ there exists an isomorphism $\pi : A \longmapsto \widehat{A}$ of $\mathcal{A}$ onto a separable Hilbert space $\mathcal{H}$ such that two of the Dirac–von Neumann axioms [14, pp. 66–67] are satisfied: $\rho \in \mathcal{S}(\mathcal{A})$ if and only if there exists a positive $\widehat{\rho} \in \mathcal{B}(\mathcal{H})$ with $\mathrm{Tr}(\widehat{\rho}) = 1$ and $\rho(A) = \mathrm{Tr}(\widehat{\rho}\widehat{A})$ for every $A \in \mathcal{A}$; moreover the space of observables $\mathcal{O}$ is identified with the self-adjoint elements of $\mathcal{B}(\mathcal{H})$ and the number $\mathrm{Tr}(\widehat{\rho}\widehat{A})$ is then the expected value of $A \in \mathcal{O}$. The C^*-algebraic approach is very appealing because it helps in unifying classical and quantum states: when the system is classical, the algebra of observables becomes an abelian C^*-algebra and the states become ordinary probability measures.

2.2 Taking traces: heuristics

2.2.1 The average of an observable

Let $\widehat{A}$ be an observable, i.e. a Hermitian operator defined on some subspace $D_{\widehat{A}}$ of $\mathcal{H}$ (the domain of $\widehat{A}$); the average value of $\widehat{A}$ in the pure state $|\psi\rangle$ is by definition

$$\langle\widehat{A}\rangle_\psi = \langle\psi|\widehat{A}|\psi\rangle = \langle\psi|\widehat{A}\psi\rangle = \langle\widehat{A}\psi|\psi\rangle. \tag{22}$$

This relation is usually rewritten using the notion of trace Tr of an operator, defined as the sum of the eigenvalues

of that operator. In fact, one of the most used formulas in quantum mechanics is, no doubt,

$$\langle \widehat{A} \rangle_\psi = \mathrm{Tr}(\widehat{\rho}_\psi \widehat{A}). \tag{23}$$

To see this, we observe that since $\widehat{\rho}_\psi \widehat{A} = |\psi\rangle\langle\psi|\widehat{A}$ and $\widehat{\rho}_\psi^2 = \widehat{\rho}_\psi$ we have, by cyclicity,

$$\mathrm{Tr}(\widehat{\rho}_\psi \widehat{A}) = \mathrm{Tr}(\widehat{\rho}_\psi^2 \widehat{A}) = \mathrm{Tr}(\widehat{\rho}_\psi \widehat{A} \widehat{\rho}_\psi) \tag{24}$$

that is, in *bra-ket notation*, and using the homogeneity of the trace,

$$\begin{aligned} \mathrm{Tr}(\widehat{\rho}_\psi \widehat{A}) &= \mathrm{Tr}(|\psi\rangle\langle\psi|\widehat{A}|\psi\rangle\langle\psi|) \\ &= \langle\psi|\widehat{A}|\psi\rangle \, \mathrm{Tr}(|\psi\rangle\langle\psi|) \\ &= \langle\psi|\widehat{A}|\psi\rangle \end{aligned}$$

since $\mathrm{Tr}(|\psi\rangle\langle\psi|) = \mathrm{Tr}(\widehat{\rho}_\psi) = 1$ (the only eigenvalues of $\widehat{\rho}_\psi$ are 1 and 0). Formula (23) is immediately generalized by linearity of the trace to arbitrary mixed states, and one finds, that the average value of the observable $\widehat{A}$ in a mixed state with density operator (14) is

$$\langle \widehat{A} \rangle_{\widehat{\rho}} = \mathrm{Tr}(\widehat{\rho} \widehat{A}). \tag{25}$$

It turns out that several mixtures can lead to the same density matrix. Since formula (25) only involves the density operator itself, such equivalent mixtures cannot be distinguished by measurement of observables alone because all observable results can be predicted from the density matrix, without needing to know the ensemble that was used to construct it.

2.2.2 Precautions

So far, so good. However, the manipulations made above can—at best—be justified when the Hilbert space $\mathcal{H}$ is finite dimensional because in this case the notion of trace class operators (and their properties) involves elementary calculations of finitely-dimensional matrices. This is almost always implicitly assumed in textbooks, where the reader is invited to study *ad nauseam* the same elementary examples in low dimensions. Unfortunately the definitions above, and the proof of the formulas (23)–(25) are *incorrect* in the general case of an infinite-dimensional $\mathcal{H}$, and formulas like (25) are not justifiable unless one makes very restrictive hypotheses on the observable $\widehat{A}$ because of convergence problems for the involved integrals. These issues are seldom questioned and treated sloppily in most of the physical literature (in this context it is very instructive to read the excellent paper [15] by Daubechies which warns of such misconceptions). The situation is very much similar to what happens when physicists manipulate the Feynman path integral (I am not talking here about the rigorous theory of the Feynman integral, as found for instance in Albeverio *et al.* [16] or Nicola [17]). It is a heuristic object which has, in the form used most of the time in physics, no mathematical justification, but which easily allows one to get new intuitions and insights (hence its usefulness). Of course, such formal manipulations are okay when one is only interested in qualitative statements, but they are certainly not okay when one wants to perform exact calculations (numerical, or theoretical).

Here is another popular (and questionable!) way of calculating traces. Let $\widehat{\rho}$ be a density matrix on $L^2(\mathbb{R}^n)$; assume that the kernel of $\widehat{\rho}$ is a function $K(x, y)$:

$$\widehat{\rho}\psi(x) = \int K(x, y)\psi(y) d^n y. \tag{26}$$

It is customary (especially in the physical literature) to calculate the trace of $\widehat{\rho}$ by integrating the kernel along the diagonal, that is by using the formula

$$\mathrm{Tr}(\widehat{A}) = \int K(x, x) d^n x \tag{27}$$

which is obviously an extension to the infinite dimensional case of the usual definition of the trace of a matrix as the sum of its diagonal elements. Needless to say, this formula does not follow directly from the definition of a trace class operator! In fact, even when the integral in (27) is absolutely convergent, this formula has no reason to be true in general (the kernel of an operator is defined uniquely only up to a set of measure zero, and in (27) we are integrating along the diagonal, which precisely has measure zero!). Brislawn [18] discusses formulas of the type (27) in the context of traceable Hilbert–Schmidt operators and gives sufficient conditions for such formulas to hold using Mercer's theorem and its variants. As shown in Barry Simon's little monograph [19] on trace ideals, the right condition for formula (27) to hold in the case $n = 1$ is that K must be of positive type, which means that we have

$$\sum_{1 \leq j,k \leq N} \lambda_j \overline{\lambda_k} K(x_j, x_k) \geq 0 \tag{28}$$

for all integers N, all $x_j \in \mathbb{R}$ and all $\lambda_j \in \mathbb{C}$ (in particular we must thus have $K \geq 0$). On the positive side, Simon [19] notes that if a density matrix has kernel K satisfying $\int |K(x, x)| d^n x < \infty$ then we are *almost sure* that formula (27) holds. Of course this vague statement is not a charter allowing carefree calculations! Needless to say, the *derivations* above are formal and one should be extremely cautious when using the *formulas* thus obtained. Shubin [20, §27], discusses a (not so easy to use explicitly) step-by-step procedure for checking such identities. We will come back to these essential points in Section 4.4.

2.3 Quantum Tomography

Reconstructing a quantum state (or even a classical state) is an extremely important problem, and it is, generally speaking, a difficult one. We will discuss here very briefly this topic; we will mainly be giving references to past works; such a list is of course incomplete due to the large number of contributions, which is steadily growing. We will mainly focus on quantum tomography, which is a technique for characterizing a state of a quantum system by subjecting it to a large number of quantum measurements, each time preparing the system anew.

2.3.1 Estimating the density matrix: generalities

Experiments performed in a laboratory on quantum systems do not lead directly and precisely to a determination of the density matrix. The problem can be formulated as a statistical one: can we estimate the density matrix using repeated measurements on quantum systems that are identically prepared? The following strategy is used: after having obtained measurements on these identical quantum systems we can make a statistical inference about the probability distribution of the measurements, and thus indirectly about the density matrix of the quantum system [21]. This procedure is called *quantum state tomography*, and is practically implemented using a set of measurements of a *quorum* of observables, *i.e.*, of a minimal *complete* set of non-commuting observables. See D'Ariano and his collaborators [22, 23] for various strategies, and also Vogel and Risken [1] or Leonhardt and Paul [24]. Bužek and his collaborators [25] show how to reconstruct the Wigner function using the Jaynes principle of Maximum Entropy. In that context, also see the recent paper Thekkadath *et al*. [26] where a scheme that can be used to directly measure individual density matrix elements is given; Lvovsky and Raymer [27] give a very interesting review in the continuous-variable case. The volume [28] edited by Paris and Reháček contains a selection of texts by leading experts presenting various aspects of quantum state estimation. Mancini *et al*. [29] apply symplectic geometry to describe the dynamics of a quantum system in terms of equations for a purely classical probability distribution; also see Man'ko and Man'ko [30]. Ibort *et al*. [31] introduce what they call the tomographic picture of quantum mechanics.

Although the Wigner function cannot be measured directly as a probability density, all its marginal distributions can. Once we know all the marginal distributions associated with different quadratures, we can reconstruct the Wigner function. Historically, the problem can be traced back to Wolfgang Pauli's question in his 1958 book:

> The mathematical problem as to whether, for given probability densities $W(p)$ and $W(x)$, the wavefunction $\psi(\dots)$ is always uniquely determined, has still not been investigated in all its generality. [32]

It turns out that it is not possible to determine a state by knowing only the configuration space and the momentum space marginal distributions. The answer is negative; here is an elementary example (Esposito *et al*. [33]): consider the two normalized wavefunctions

$$\begin{aligned}\psi_1(x) &= \left(\frac{2}{\pi}\operatorname{Re}\alpha\right)^{1/4} e^{-\alpha x^2}\\ \psi_2(x) &= \left(\frac{2}{\pi}\operatorname{Re}\alpha\right)^{1/4} e^{-\alpha^* x^2}\end{aligned}$$

where α is a complex number whose real part is positive: $\operatorname{Re}\alpha > 0$ and imaginary part nonzero: $\operatorname{Im}\alpha \neq 0$. These functions become

$$\begin{aligned}\phi_1(p) &= \left(\frac{2}{\pi}\operatorname{Re}\alpha\right)^{1/4} \frac{1}{\sqrt{2\alpha}} e^{-\alpha p^2}\\ \phi_2(p) &= \left(\frac{2}{\pi}\operatorname{Re}\alpha\right)^{1/4} \frac{1}{\sqrt{2\alpha^*}} e^{-\alpha p^2}\end{aligned}$$

in the momentum representation, so that the associated probability distributions are the same

$$\begin{aligned}|\psi_1(x)|^2 = |\psi_2(x)|^2 &= \left(\frac{2}{\pi}\operatorname{Re}\alpha\right)^{1/2} e^{-2(\operatorname{Re}\alpha)x^2}\\ |\phi_1(p)|^2 = |\phi_2(p)|^2 &= \left(\frac{2}{\pi}\operatorname{Re}\alpha\right)^{1/2} \frac{1}{2|\alpha|} e^{-2(\operatorname{Re}\alpha)x^2}.\end{aligned}$$

However the states $|\psi_1\rangle$ and $|\psi_2\rangle$ are different, because

$$|\langle\psi_1|\psi_2\rangle|^2 = \frac{\operatorname{Re}\alpha}{|\alpha|} \neq 1 \tag{29}$$

so that we cannot have $\psi_1 = c\psi_2$ for some complex constant c such that $|c| = 1$.

2.3.2 The Radon transform

Let us begin with a short discussion of the classical case followed by a motivation for the quantum case; we are following the very clear exposition [33, §12.5] by Esposito *et al*. Let $\rho = \rho(x, p)$ be some function defined on the phase space $\mathbb{R}^2$ (it could be a classical probability, or a quantum quasiprobability). The Radon transform of this function is defined formally in physics texts by

$$\mathcal{R}\rho(X, \mu, \nu) = \iint \rho(x, p)\delta(X - \mu x - \nu p)dpdx. \tag{30}$$

The function $\mathcal{R}\rho(X,\mu,\nu)$ is called a *tomogram*. The meaning of this integral is that one integrates the function ρ along the line ℓ with equation $X-\mu x-\nu p=0$ where the variable X has an arbitrary value. Now, the crucial point is that if we know fully the function $R\rho$—which depends on the three real variables X,μ,ν—then one can reconstruct the function ρ by using the so-called *inverse Radon transform*:

$$\rho(x,p)=\left(\tfrac{1}{2\pi}\right)^2\iiint \mathcal{R}\rho(X,\mu,\nu)e^{i(X-\mu x-\nu p)}dXd\mu d\nu. \tag{31}$$

Let us *verify* this formula, just for the fun of making formal calculations. Denoting by A the triple integral above we have

$$\begin{aligned}
A&=\int\rho(x',p')\left\{\iiint\left[\int e^{i(X-\mu x-\nu p)}\delta(X-\mu x'-\nu p')dX\right]d\mu d\nu\right\}dp'dx'\\
&=\int\rho(x',p')\left\{\iiint\left[e^{i(\mu(x-x)+\nu(p-p'))}\int\delta(X-\mu x'-\nu p')dX\right]d\mu d\nu\right\}dp'dx'\\
&=\int\rho(x',p')\left\{\iiint e^{i(\mu(x-x)+\nu(p-p'))}d\mu d\nu\right\}dp'dx'=(2\pi)^2\int\rho(x',p')\delta(x-x')\delta(p-p')dp'dx'
\end{aligned}$$

and hence $A=(2\pi)^2\rho(x,p)$, so we are done.

In the quantum case one proceeds as follows: noting that $\delta(X-\mu x-\nu p)$ can be rewritten as the Fourier integral

$$\delta(X-\mu x-\nu p)=\frac{1}{2\pi\hbar}\int_{-\infty}^{\infty}e^{\frac{i}{\hbar}k(X-\mu x-\nu p)}dk \tag{32}$$

and hence (30) as

$$\mathcal{R}\rho(X,\mu,\nu)=\frac{1}{2\pi\hbar}\iint\rho(x,p)\left(\int_{-\infty}^{\infty}e^{\frac{i}{\hbar}k(X-\mu x-\nu p)}dk\right)dpdx \tag{33}$$

one *defines* the quantum Radon transform by analogy with (30) by the formula

$$\mathcal{R}\rho(X,\mu,\nu)=\frac{1}{2\pi\hbar}\operatorname{Tr}\left(\widehat{\rho}\int_{-\infty}^{\infty}e^{\frac{i}{\hbar}k(\widehat{X}-\mu\widehat{x}-\nu\widehat{p})}dk\right). \tag{34}$$

Doing this one has of course to give a precise meaning to the exponential $e^{\frac{i}{\hbar}k(\widehat{X}-\mu\widehat{x}-\nu\widehat{p})}$, and one then *inverts* the formula above, which yields in analogy with (31)

$$\widehat{\rho}=\frac{1}{2\pi\hbar}\int\mathcal{R}\rho(X,\mu,\nu)e^{\frac{i}{\hbar}(\widehat{X}-\mu\widehat{x}-\nu\widehat{p})}dXd\mu d\nu. \tag{35}$$

Of course these manipulations are heuristic and completely formal and do not have any mathematical sense unless one defines rigorously the involved operators.

2.3.3 Quantum holography

Quantum holography is the process of discovering features of an unknown quantum state from the knowledge of its phase. It has important practical applications in communication systems such as optical channels, precision measurement devices such as atomic clocks, and quantum computation. The starting point is the observation that a pure quantum state $\psi=Re^{i\varphi/\hbar}$ is described by two real quantities, its modulus R, and its phase φ. The full recognition of the importance of the phase can be traced back to M. Berry [34], and to Y. Aharonov and D. Bohm [35]. Berry showed that the total phase of a quantum state is composed of two essentially distinct parts: a *dynamical phase*, related to the system dynamics, and a *geometrical phase*, which mirrors the geometry of the underlying Hilbert space. This situation is quite similar to the revolution brought by optical holography, which makes possible the full recording of the light-field information by the knowledge of the field's phase, and for which Dennis Gabor was awarded the Nobel prize in 1971 [36].

We will consider the following problem: let $(\widehat{U}_t)$ be a quantum isotopy (i.e. a C^1 path $t\longmapsto\widehat{U}_t$ of unitary operators on $L^2(\mathbb{R}^n)$ such that $\widehat{U}_0=I_{\rm d}$) and $\widehat{\rho}$ a mixed quantum state. We want to calculate in a rigorous way the *Pancharatnam–Sjöqvist relative phase shift* [37, 38]

$$\varphi(t)=\operatorname{Arg}\operatorname{Tr}(\widehat{U}_t\widehat{\rho}) \tag{36}$$

for all t such that $\operatorname{Tr}(\widehat{U}_t\widehat{\rho})\neq 0$; when $\widehat{\rho}$ represents a pure state ψ then (36) reduces to the formula

$$\varphi(t)=\operatorname{Arg}(\widehat{U}_t\psi|\psi)_{L^2}, \tag{37}$$

which was originally proposed by Pancharatnam [37] in 1956. This question is extremely important because, as Nicacio shows in [39], one can reconstruct a Gaussian state $\widehat{\rho}$ from the knowledge of the phase $\varphi(t)$ (this is a variant of the *phase retrieval problem* considered in time-frequency analysis and quantum optics).

We have proved with F. Nicacio in [40], that when $(\widehat{U}_t)$ consists of metaplectic operators $\widehat{S}_t$ then if that $S_t = \pi^{\mathrm{Mp}}(\widehat{S}_t)$ has no eigenvalue equal to one the Wigner distribution ρ is Gaussian distribution

$$\rho(z) = (\pi\hbar)^{-n}\sqrt{\det\Sigma}e^{-\frac{1}{\hbar}\Sigma z\cdot z} \tag{38}$$

with covariance matrix Σ then

$$\mathrm{Tr}(\widehat{S}_t\widehat{\rho}) = \frac{i^{\nu(\widehat{S}_t)}}{\sqrt{|\det(S_t - I)|}}\det{}^{-1/2}(\tfrac{1}{2}F^{-1} + iM(S_t^T)). \tag{39}$$

so that the relative phase shift is given by the formula

$$\phi(t) = \frac{\pi}{2}\nu(\widehat{S}_t) + \mathrm{Arg}\det{}^{-1/2}(\tfrac{1}{2}F^{-1} + iM(S_t^T)). \tag{40}$$

In the two formulas above $\nu(\widehat{S}_t)$ is the Conley–Zehnder index of the path $t' \longmapsto \widehat{S}_{t'}$, $0 \leq t' \leq t$ (it is a variant of the Maslov index [41–43]); see Section 5). The proofs of these formulas relies on the fact that $\widehat{S}_t\widehat{\rho}$ is of trace class and that we have

$$\mathrm{Tr}(\widehat{S}_t\widehat{\rho}) = \left(\tfrac{1}{2\pi\hbar}\right)^n\int s_{t,\sigma}(z)\rho_\sigma(z)d^{2n}z \tag{41}$$

where s_t is the Weyl symbol of $\widehat{S}_t$; the subscript σ in $s_{t,\sigma}$ and ρ_σ stands for the symplectic Fourier transforms (10) of s_t and ρ.

3 Mathematical Theory of the Density Matrix

To give a precise definition of the trace formulas above, we need to work a little bit and use some operator theory (the theory of Hilbert–Schmidt and trace class operators suffices at this point). We will use several times the generalized Bessel equality

$$\sum_j\langle\psi|\psi_j\rangle\langle\phi|\psi_j\rangle^* = \langle\psi|\phi\rangle \tag{42}$$

valid for every orthonormal basis of a Hilbert space $\mathcal{H}$. It is an immediate consequence of the equality $\sum_j|\psi_j\rangle\langle\psi_j| = I$.

3.1 Trace class operators

3.1.1 General algebraic definition

Let us now define the notion of density operator in terms of the so important notion of trace class operators. The starting point is to notice that formula (14) implies that a density operator has, to begin with, three basic properties:

(1) It is a *bounded operator* on $\mathcal{H}$; in particular $\widehat{\rho}$ is well-defined for all $\psi \in \mathcal{H}$;

(2) It is a *self-adjoint operator*: $\widehat{\rho}^\dagger = \widehat{\rho}$;

(3) It is a *positive operator*: $\langle\psi|\widehat{\rho}|\psi\rangle \geq 0$ for all $\psi \in \mathcal{H}$.

The boundedness of $\widehat{\rho}$ follows from the observation that

$$\|\widehat{\rho}\psi\| \leq \sum_j\alpha_j\|\widehat{\rho}_j\psi\| \leq \sum_j\alpha_j\|\psi\| = \|\psi\| \tag{43}$$

where $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$; that $\|\widehat{\rho}_j\psi\| \leq \|\psi\|$ follows from the Cauchy–Schwarz inequality since $\widehat{\rho}_j\psi = |\psi_j\rangle\langle\psi_j|\psi\rangle$. The self-adjointness of $\widehat{\rho}$ is clear, since it is a linear combination of the self-adjoint operators $\widehat{\rho}_j = |\psi_j\rangle\langle\psi_j|$. Finally, the positivity of $\widehat{\rho}$ is clear since we have

$$\langle\psi|\widehat{\rho}|\psi\rangle = \sum_j\alpha_j\langle\psi|\psi_j\rangle\langle\psi_j|\psi\rangle = \sum_j\alpha_j|\langle\psi|\psi_j\rangle|^2 \geq 0 \tag{44}$$

(notice that (3) $\Longrightarrow$ (2) when $\mathcal{H}$ is a complex Hilbert space). There remains the trace issue. As we mentioned above, physicists define the trace of an operator $\widehat{\rho}$ as the sum of its diagonal elements, and this definition only makes sense without further restrictions when $\dim\mathcal{H} < \infty$. The way to do things correctly consists in using the mathematical definition of trace class operators; the latter is very general (and hence very useful) because it involves arbitrary bases of $\mathcal{H}$. The definition goes as follows: a bounded operator $\widehat{\rho}$ on a Hilbert space $\mathcal{H}$ (we do not assume self-adjointness or positivity at this point) is of *trace class* if there exist two orthonormal bases (ϕ_j) and (χ_j) of $\mathcal{H}$ (indexed by the same set) such that

$$\sum_j|\langle\widehat{\rho}\phi_j|\chi_j\rangle| < \infty. \tag{45}$$

Notice that this definition immediately implies that $\widehat{\rho}$ is of trace class if and only if its adjoint $\widehat{\rho}^\dagger$ is. Now, a crucial property is that if condition (45) holds for one pair of orthonormal bases, then it also holds for *all* pairs of orthonormal bases, and that if (ψ_j) and (ϕ_j) are two such pairs then

$$\sum_i\langle\psi_j|\widehat{\rho}\psi_j\rangle = \sum_i\langle\phi_j|\widehat{\rho}\phi_j\rangle \tag{46}$$

each series being absolutely convergent. The proof of this result is not difficult; it consists in expanding each base using the vectors of the other; we refer to de Gosson [44, Chapter 12] and Shubin [20, Appendix 3] for complete proofs. This being done, we *define* the trace of $\widehat{\rho}$ by the formula

$$\mathrm{Tr}(\widehat{\rho}) = \sum_k\langle\psi_k|\widehat{\rho}\psi_k\rangle \tag{47}$$

where (ψ_j) is any orthonormal basis of $\mathcal{H}$; that the result does not depend on the choice of such a basis follows from the identity (46). We leave to the reader to verify that as a consequence of this definition the sum of two

trace class operators is again a trace class operator, and that the trace of their sum is the sum of the traces. Also, $\mathrm{Tr}(\lambda\widehat{\rho}) = \lambda\,\mathrm{Tr}(\widehat{\rho})$ for every complex number λ, hence density matrices form a convex cone. Also notice that it immediately follows from definition (47) that

$$\mathrm{Tr}(\widehat{\rho}^{\dagger}) = \mathrm{Tr}(\widehat{\rho})^{*} \tag{48}$$

hence the trace is real if $\widehat{\rho}$ is self-adjoint.

3.1.2 Invariance under unitary conjugation

The following invariance of the trace under unitary conjugation is well known:

Theorem 2. *Let $\widehat{\rho}$ be a density matrix on the Hilbert space $\mathcal{H}$ and $\widehat{U}$ a unitary operator on $\mathcal{H}$. Then $\widehat{U}^{\dagger}\widehat{\rho}\widehat{U}$ is also a positive trace class operator and we have*

$$\mathrm{Tr}(\widehat{U}^{\dagger}\widehat{\rho}\widehat{U}) = \mathrm{Tr}(\widehat{A}). \tag{49}$$

Proof. It is clear that $\widehat{U}^{\dagger}\widehat{\rho}\widehat{U}$ is a positive, bounded, and self-adjoint operator. The operator $\widehat{\rho}$ is of trace class if and only if $\sum_j \langle\psi_j|\widehat{\rho}|\psi_j\rangle < \infty$ for one (and hence every) orthonormal basis $(\psi_j)_j$ of $\mathcal{H}$. Since $\langle\psi_j|\widehat{U}^{\dagger}\widehat{\rho}\widehat{U}|\psi_j\rangle = \langle\widehat{U}\psi_j|\widehat{\rho}\widehat{U}\psi_j\rangle$ and the basis $(\widehat{U}\psi_j)_j$ also is orthonormal, the operator $\widehat{U}^{\dagger}\widehat{\rho}\widehat{U}$ is of trace class. The trace formula (49) follows from the orthonormal basis independence of the trace formula (47). □

We have not yet related our definition of trace to that used in the previous section. Let us show that these definitions do coincide. Let

$$\widehat{\rho} = \sum_j \alpha_j |\psi_j\rangle\langle\psi_j| \tag{50}$$

be a density matrix in the sense of (14); the vectors ψ_j being normalized we have

$$\langle\psi_k|\widehat{\rho}|\psi_k\rangle = \sum_j \alpha_j |\langle\psi_j|\psi_k\rangle|^2 = \alpha_k \tag{51}$$

and hence, using definition (47) of the trace,

$$\mathrm{Tr}(\widehat{\rho}) = \sum_k \alpha_k = 1. \tag{52}$$

3.1.3 The spectral theorem

An especially useful expansion of a density operator is obtained using elementary functional analysis (the spectral resolution theorem). Recall that each eigenvalue of a selfadjoint compact operator (except possibly zero) has finite multiplicity (see any book on elementary functional analysis, e.g. Blanchard and Brüning [45]).

Theorem 3. *A bounded linear self-adjoint operator $\widehat{\rho}$ on a complex Hilbert space $\mathcal{H}$ is of trace class if and only if there exists a sequence of real numbers $\lambda_j \geq 0$ and an orthonormal basis (ψ_j) of $\mathcal{H}$ such that for all $\psi \in \mathcal{H}$*

$$\widehat{\rho}\psi = \sum_j \lambda_j \langle\psi_j|\psi\rangle\psi_j \tag{53}$$

that is $\widehat{\rho} = \sum_j \lambda_j \widehat{\rho}_j$ where $\widehat{\rho}_j$ is the orthogonal projection on the ray $\mathbb{C}\psi_j$. In particular, $\widehat{\rho}$ is a density matrix if and only if $\lambda_j \geq 0$ for every j and $\sum_j \lambda_j = 1$; the vector ψ_j is the eigenvector corresponding to the eigenvalue λ_j.

Proof. This is a classical result from the theory of compact self-adjoint operators on a Hilbert space; see any introductory book on functional analysis, for instance Blanchard and Brüning [45]. That ψ_j is an eigenvector corresponding to λ_j is clear: since $\langle\psi_k|\psi_j\rangle = \delta_{kj}$ we have

$$\widehat{\rho}\psi_j = \sum_k \lambda_k \langle\psi_k|\psi_j\rangle\psi_k = \lambda_j\psi_j. \tag{54}$$

□

An immediate consequence of Theorem 3 is that if $\widehat{\rho}$ is a density matrix, then $\mathrm{Tr}(\widehat{\rho}^2) \leq 1$ with equality if and only if $\widehat{\rho}$ represents a pure state: we have

$$\widehat{\rho}^2 = \left(\sum\nolimits_j \lambda_j\widehat{\rho}_j\right)^2 = \sum\nolimits_{j,k} \lambda_j\lambda_k\widehat{\rho}_j\widehat{\rho}_k = \sum\nolimits_j \lambda_j^2\widehat{\rho}_j \tag{55}$$

the second equality because $\widehat{\rho}_j\widehat{\rho}_k = 0$ if $j \neq k$ since ψ_j and ψ_k are then orthogonal, and $\widehat{\rho}_j^2 = \widehat{\rho}_j$. Since $\lambda_j^2 \leq \lambda_j \leq 1$ we have

$$\mathrm{Tr}(\widehat{\rho}^2) = \sum\nolimits_j \lambda_j^2 \leq 1. \tag{56}$$

The equality $\sum_j \lambda_j^2 = 1$ can only occur if all the coefficients λ_j are equal to zero, except one which is equal to one. Thus $\mathrm{Tr}(\widehat{\rho}^2) = 1$ if and only if $\widehat{\rho}$ represents a pure state. The number $\mathrm{Tr}(\widehat{\rho}^2)$ is therefore called the *purity* of the quantum state represented by the density matrix $\widehat{\rho}$; it is often denoted by $\mu(\widehat{\rho})$. Another way of measuring the purity of a state is to use the von Neumann entropy $S(\widehat{\rho})$. By definition:

$$S(\widehat{\rho}) = -\sum_j \lambda_j \ln\lambda_j \tag{57}$$

(with the convention $0\ln 0 = 0$). One often uses the suggestive notation

$$S(\widehat{\rho}) = -\,\mathrm{Tr}(\widehat{\rho}\ln\widehat{\rho}) \tag{58}$$

but one should then not forget that the right-hand side of this equality is defined by (57), and not the other way around! Notice that the von Neumann entropy $S(\widehat{\rho})$ is zero if and only if $\widehat{\rho}$ is a pure state. This is because the equality $\mathrm{Tr}(\widehat{\rho}\ln\widehat{\rho}) = 0$ is equivalent to $\lambda_j = 1$ for all $j = 1, 2, \ldots$.

3.1.4 Functional properties

Perhaps the most directly useful property of trace class operators (and hence of density matrices) is Theorem 4 below; it says that if we compose a trace class operator with any bounded operator we obtain again a trace class operator.

We denote by $\mathcal{L}_1(\mathcal{H})$ the set of all trace class operators on the Hilbert space $\mathcal{H}$ and by $\mathcal{B}(\mathcal{H})$ the algebra of bounded linear operators on $\mathcal{H}$. The following important result justifies the trace formula (25) for bounded observables:

Theorem 4. *The set $\mathcal{L}_1(\mathcal{H})$ of all trace class operators on $\mathcal{H}$ is both a vector subspace of $\mathcal{B}(\mathcal{H})$ and a two-sided ideal in $\mathcal{B}(\mathcal{H})$: if $\widehat{\rho} \in \mathcal{L}_1(\mathcal{H})$ and $\widehat{A} \in \mathcal{B}(\mathcal{H})$ then $\widehat{\rho}\widehat{A} \in \mathcal{L}_1(\mathcal{H})$ and $\widehat{A}\widehat{\rho} \in \mathcal{L}_1(\mathcal{H})$ and we have*

$$\mathrm{Tr}(\widehat{\rho}\widehat{A}) = \mathrm{Tr}(\widehat{A}\widehat{\rho}). \tag{59}$$

Formula (59) applies in particular when $\widehat{\rho}$ is a density matrix and $\widehat{A}$ a bounded quantum observable on $\mathcal{H}$.

Proof. That $\mathcal{L}_1(\mathcal{H})$ is a vector space is clear using formula (45): if $\sum_j |\langle \widehat{\rho}_1\phi_j|\chi_j\rangle| < \infty$ and $\sum_j |\langle \widehat{\rho}_2\phi_j|\chi_j\rangle| < \infty$ then we have, using the triangle inequality,

$$\sum_j |\langle (\widehat{\rho}_1+\widehat{\rho}_2)\phi_j|\chi_j\rangle| \leq \sum_j |\langle \widehat{\rho}_1\phi_j|\chi_j\rangle| + \sum_j |\langle \widehat{\rho}_2\phi_j|\chi_j\rangle| < \infty \tag{60}$$

so that $\widehat{\rho}_1 + \widehat{\rho}_2 \in \mathcal{L}_1(\mathcal{H})$; that $\lambda\widehat{\rho} \in \mathcal{L}_1(\mathcal{H})$ if $\widehat{\rho} \in \mathcal{L}_1(\mathcal{H})$ and $\lambda \in \mathbb{C}$ is clear. Let us show that $\widehat{A}\widehat{\rho} \in \mathcal{L}_1(\mathcal{H})$ if $\widehat{\rho} \in \mathcal{L}_1(\mathcal{H})$ and $\widehat{A}$ is a bounded operator on $\mathcal{H}$. Recall that the boundedness of $\widehat{A}$ is equivalent to the existence of a number $C \geq 0$ such that $\|\widehat{A}\psi\| \leq C\|\psi\|$ for all $\psi \in \mathcal{H}$. Let now (ψ_j) and (ϕ_j) be two orthonormal bases of $\mathcal{H}$; writing $\langle \widehat{A}\widehat{\rho}\psi_j|\phi_j\rangle = \langle \widehat{\rho}\psi_j|\widehat{A}^\dagger\phi_j\rangle$ and applying Bessel's equality (42) to $\langle \widehat{\rho}\psi_j|\widehat{A}^\dagger\phi_j\rangle$ we get

$$\langle \widehat{A}\widehat{\rho}\psi_j|\phi_j\rangle = \sum_k \langle \widehat{\rho}\psi_j|\phi_k\rangle\langle \widehat{A}^\dagger\phi_j|\psi_k\rangle^*; \tag{61}$$

using the Cauchy–Schwarz inequality we have, since $\|\widehat{A}^\dagger\phi_j\| \leq C$,

$$|\langle \widehat{A}^\dagger\phi_j|\psi_k\rangle^*| \leq \|\widehat{A}^\dagger\phi_j\|\,\|\psi_k\| \leq C \tag{62}$$

and hence

$$|\langle \widehat{A}\widehat{\rho}\psi_j|\phi_j\rangle| \leq \sum_k |\langle \widehat{\rho}\psi_j|\phi_k\rangle\langle \widehat{A}^\dagger\phi_j|\psi_k\rangle^*| \leq C\sum_k |\langle \widehat{\rho}\psi_j|\phi_k\rangle|. \tag{63}$$

Summing this inequality with respect to the index j yields, since $\widehat{\rho}$ is of trace class

$$|\sum_j \langle \widehat{A}\widehat{\rho}\psi_j|\phi_j\rangle| \leq C\sum_{j,k} |\langle \widehat{\rho}\psi_j|\phi_k\rangle| < \infty \tag{64}$$

hence $\widehat{A}\widehat{\rho}$ is of trace class as claimed. That $\widehat{\rho}\widehat{A}$ also is of trace class is immediate noting that we can write $\widehat{\rho}\widehat{A} = (\widehat{A}^\dagger\widehat{\rho}^\dagger)^\dagger$. There remains to prove the trace equality (59). Choosing $(\psi_j) = (\phi_j)$ the Bessel equality (61) becomes

$$\langle \widehat{A}\widehat{\rho}\psi_j|\psi_j\rangle = \sum_k \langle \widehat{\rho}\psi_j|\psi_k\rangle\langle \widehat{A}^\dagger\psi_j|\psi_k\rangle^* \tag{65}$$

and, similarly,

$$\begin{aligned}\langle \widehat{\rho}\widehat{A}\psi_j|\psi_j\rangle &= \sum_k \langle \widehat{A}\psi_j|\psi_k\rangle\langle \widehat{\rho}^\dagger\psi_j|\psi_k\rangle^* \\ &= \sum_k \langle \widehat{\rho}\psi_k|\psi_j\rangle\langle \widehat{A}^\dagger\psi_k|\psi_j\rangle^*.\end{aligned} \tag{66}$$

Summing this equality over j we get

$$\mathrm{Tr}(\widehat{\rho}\widehat{A}) = \sum_j \langle \widehat{A}\widehat{\rho}\psi_j|\psi_j\rangle = \sum_{j,k} \langle \widehat{\rho}\psi_j|\psi_k\rangle\langle \widehat{A}^\dagger\psi_j|\psi_k\rangle^* \tag{67}$$

$$\mathrm{Tr}(\widehat{A}\widehat{\rho}) = \sum_j \langle \widehat{\rho}\widehat{A}\psi_j|\psi_j\rangle = \sum_{j,k} \langle \widehat{\rho}\psi_k|\psi_j\rangle\langle \widehat{A}^\dagger\psi_k|\psi_j\rangle^* \tag{68}$$

hence $\mathrm{Tr}(\widehat{\rho}\widehat{A}) = \mathrm{Tr}(\widehat{A}\widehat{\rho})$ since the sums of both right-hand sides are identical. □

3.2 Hilbert–Schmidt operators

We are following here de Gosson [44, §12.1] and Shubin [20, Appendix 3].

3.2.1 The trace norm

An operator $\widehat{A}$ on a Hilbert space $\mathcal{H}$ is called a *Hilbert–Schmidt operator* if there exists an orthonormal basis (ψ_j) of $\mathcal{H}$ such that

$$\sum_j \langle \widehat{A}\psi_j|\widehat{A}\psi_j\rangle = \sum_j \|\widehat{A}\psi_j\|^2 < \infty. \tag{69}$$

In particular such an operator is bounded on $\mathcal{H}$. As in the case of trace class operators one shows that if this property holds for one orthonormal basis then it holds for all, and that the number $\sum_j \|\widehat{A}\psi_j\|^2$ is independent of the choice of such a basis: Let in fact (ϕ_j) be a second orthonormal basis, and write $\widehat{A}\psi_j = \sum_k \langle \phi_j|\widehat{A}\psi_j\rangle\phi_k$. Then, using the Bessel equality (42)

$$\sum_j \|\widehat{A}\psi_j\|^2 = \sum_{j,k} |\langle \phi_j|\widehat{A}\psi_j\rangle|^2 = \sum_{j,k} |\langle \widehat{A}^\dagger\phi_j|\psi_j|^2 \tag{70}$$

that is, using again (42),

$$\sum_j \|\widehat{A}\psi_j\|^2 = \sum_k \|\widehat{A}^\dagger\phi_k\|^2 < \infty. \tag{71}$$

Observe that this equality shows that if we take $(\psi_j)_j = (\phi_j)_j$ then $\sum_k \|\widehat{A}^\dagger\psi_k\|^2_{\mathcal{H}} < \infty$ hence the adjoint $\widehat{A}^\dagger$ of a

Hilbert–Schmidt operator is also a Hilbert–Schmidt operator; we may thus replace $\widehat{A}$ by $\widehat{A}^\dagger$ in the inequality above, which yields $\sum_k \|\widehat{A}^\dagger\phi_k\|^2 < \infty$ as claimed.

Hilbert–Schmidt operators form a vector space $\mathcal{L}_2(\mathcal{H})$ and

$$\|\widehat{A}\|_{\rm HS} = \left(\sum_j \|\widehat{A}\psi_j\|^2\right)^{1/2} \tag{72}$$

defines a norm on this space ; this norm is associated with the scalar product

$$\langle\widehat{A}|\widehat{B}\rangle_{\rm HS} = {\rm Tr}(\widehat{A}^\dagger\widehat{B}) = \sum_j \langle\widehat{A}\psi_j|\widehat{B}\psi_j\rangle. \tag{73}$$

If $\widehat{A}$ and $\widehat{B}$ are Hilbert–Schmidt operators then $\lambda\widehat{A}$ is trivially a Hilbert–Schmidt operator and $\|\lambda\widehat{A}\|_{\rm HS} = |\lambda|\,\|\widehat{A}\|_{\rm HS}$ for every $\lambda \in \mathbb{C}$; on the other hand, by the triangle inequality,

$$\sum_j \|(\widehat{A}+\widehat{B})\psi_j\|^2 \leq \sum_j \|\widehat{A}\psi_j\|^2 + \sum_j \|\widehat{B}\psi_j\|^2 < \infty \tag{74}$$

for every orthonormal basis (ψ_j) hence $\widehat{A}+\widehat{B}$ is also a Hilbert–Schmidt operator and we have

$$\|\widehat{A}+\widehat{B}\|_{\rm HS}^2 \leq \|\widehat{A}\|_{\rm HS}^2 + \|\widehat{B}\|_{\rm HS}^2 \leq (\|\widehat{A}\|_{\rm HS} + \|\widehat{B}\|_{\rm HS})^2 \tag{75}$$

and hence

$$\|\widehat{A}+\widehat{B}\|_{\rm HS} \leq \|\widehat{A}\|_{\rm HS} + \|\widehat{B}\|_{\rm HS}. \tag{76}$$

Finally, $\|\widehat{A}\|_{\rm HS} = 0$ is equivalent to $\widehat{A}\psi_j = 0$ for every index j that is to $\widehat{A} = 0$.

The space $\mathcal{L}_2(\mathcal{H})$ is complete for that norm, and hence a Banach space (it is actually even a Hilbert space when equipped with the scalar product (73). In addition $\mathcal{L}_2(\mathcal{H})$ is closed under multiplication (and hence an algebra).

It turns out that the space $\mathcal{L}_2(\mathcal{H})$ is a two-sided deal in the algebra of $\mathcal{B}(\mathcal{H})$ of bounded operators: if $\widehat{A} \in \mathcal{L}_2(\mathcal{H})$ and $\widehat{B} \in \mathcal{B}(\mathcal{H})$ then $\widehat{AB} \in \mathcal{L}_2(\mathcal{H})$ and $\widehat{BA} \in \mathcal{L}_2(\mathcal{H})$. Let us show that $\widehat{BA} \in \mathcal{L}_2(\mathcal{H})$. We have, denoting by $\|\widehat{B}\|$ the operator norm of $\widehat{B}$,

$$\|\widehat{BA}\|_{\rm HS}^2 = \sum_j \|\widehat{BA}\psi_j\|^2 \leq \|\widehat{B}\|\left(\sum_j \|\widehat{A}\psi_j\|^2\right) < \infty. \tag{77}$$

Applying the same argument to $\widehat{AB} = (\widehat{B}^\dagger\widehat{A}^\dagger)^\dagger$ shows that we have $\widehat{AB} \in \mathcal{L}_2(\mathcal{H})$ as well.

3.2.2 Hilbert–Schmidt and trace class

An essential property is that every trace class operator is the product of two Hilbert–Schmidt operators (and hence itself a Hilbert–Schmidt operator). Let us glorify this important statement as a theorem:

Theorem 5. *(i) A bounded operator $\widehat{A}$ on $\mathcal{H}$ is of trace class if and only if it is the product of two Hilbert–Schmidt operators: $\mathcal{L}_1(\mathcal{H}) = (\mathcal{L}_2(\mathcal{H}))^2$. (ii) A trace class operator $\widehat{A}$ on $\mathcal{H}$ is itself a Hilbert–Schmidt operator: $\mathcal{L}_1(\mathcal{H}) \subset \mathcal{L}_2(\mathcal{H})$.*

Proof. In what follows (ψ_j) is an orthonormal basis in $\mathcal{H}$. (i) Assume that $\widehat{A} = \widehat{B}\widehat{C}$ where $\widehat{B}$ and $\widehat{C}$ are both Hilbert–Schmidt operators. We have, using respectively the triangle and the Cauchy–Schwarz inequalities,

$$|\sum_j \langle\psi_j|\widehat{A}\psi_j\rangle| \leq \sum_j |\langle\widehat{B}^\dagger\psi_j|\widehat{C}\psi_j\rangle| \leq \sum_j \|\widehat{B}^\dagger\psi_j\|\,\|\widehat{C}\psi_j\|; \tag{78}$$

in view of the trivial inequality

$$\sum_j \|\widehat{B}^\dagger\psi_j\|\,\|\widehat{C}\psi_j\| \leq \frac{1}{2}\left(\sum_j \|\widehat{B}^\dagger\psi_j\|^2 + \|\widehat{C}\psi_j\|^2\right) \tag{79}$$

we get, since $\widehat{B}$ and $\widehat{C}$ are both Hilbert–Schmidt operators,

$$|\sum_j \langle\psi_j|\widehat{A}\psi_j\rangle| \leq \frac{1}{2}\left(\sum_j \|\widehat{B}^\dagger\psi_j\|^2 + \|\widehat{C}\psi_j\|^2\right) < \infty \tag{80}$$

proving that $\widehat{A}$ is indeed of trace class. Assume, conversely, that $\widehat{A} \in \mathcal{L}_1(\mathcal{H})$. In view of the polar decomposition theorem there exists a unitary operator $\widehat{U}$ on $\mathcal{H}$ such that $\widehat{A} = \widehat{U}(\widehat{A}^\dagger\widehat{A})^{1/2}$. Setting $\widehat{B} = \widehat{U}(\widehat{A}^\dagger\widehat{A})^{1/4}$ and $\widehat{C} = (\widehat{A}^\dagger\widehat{A})^{1/4}$ we have $\widehat{A} = \widehat{B}\widehat{C}$; let us show that $\widehat{C}$ and $\widehat{B}$ are Hilbert–Schmidt operators. We have

$$\begin{aligned}\sum_j |\langle\widehat{C}\psi_j|\widehat{C}\psi_j\rangle| &= \sum_j |\langle\widehat{C}^\dagger\widehat{C}\psi_j|\psi_j\rangle| \\ &= \sum_j |\langle(\widehat{A}^\dagger\widehat{A})^{1/2}\psi_j|\psi_j\rangle| < \infty\end{aligned}$$

because $(\widehat{A}^\dagger\widehat{A})^{1/2} = \widehat{U}^\dagger\widehat{A}$ is of trace class (Theorem 4), hence $\widehat{C} \in \mathcal{L}_2(\mathcal{H})$. It follows that $\widehat{B} = \widehat{U}\widehat{C} \in \mathcal{L}_2(\mathcal{H})$ as well. (ii) We have seen that every trace class operator is a product $\widehat{A} = \widehat{B}\widehat{C}$ of two Hilbert–Schmidt operators. In view of the algebra property (i) of $\mathcal{L}_2(\mathcal{H})$ the operator $\widehat{A}$ is itself Hilbert–Schmidt operator. □

3.2.3 The case of $L^2(\mathbb{R}^n)$

Let us now specialize to the case where $\mathcal{H} = L^2(\mathbb{R}^n)$. In this case Hilbert–Schmidt operators are exactly those operators that have a square integrable kernel; as a consequence a density matrix also has square integrable kernel.

Theorem 6. *Let $\widehat{A}$ be a bounded operator on $L^2(\mathbb{R}^n)$. (i) It is a Hilbert–Schmidt operator if and only if there exists a function $K \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$ such that*

$$\widehat{A}\psi(x) = \int K(x,y)\psi(y)d^n y. \tag{81}$$

(ii) Every trace class operator (and hence every density matrix) on $L^2(\mathbb{R}^n)$ can be represented by (81) with kernel $K \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$.

Proof. (i) *The condition is necessary*. Let (ψ_j) be an orthonormal basis in $L^2(\mathbb{R}^n)$. Let $\widehat{A} \in \mathcal{L}_2(L^2(\mathbb{R}^n))$; we have $\widehat{A}\psi_i = \sum_j \langle \psi_j|\widehat{A}\psi_i\rangle \psi_j$ and hence

$$\widehat{A}\psi = \sum_i \langle \psi_i|\psi\rangle \widehat{A}\psi_i = \sum_{i,j} \langle \psi_i|\psi\rangle \langle \psi_j|\widehat{A}\psi_i\rangle \psi_j \tag{82}$$

which we can rewrite, using the definition

$$\langle \psi_i|\psi\rangle = \int \psi_i^*(y)\psi(y)d^n y \tag{83}$$

of the inner product as

$$\begin{aligned}\widehat{A}\psi(x) &= \sum_{i,j} \langle \psi_i|\psi\rangle \langle \psi_j|\widehat{A}\psi_i\rangle \psi_j(x) \\ &= \sum_{i,j} \langle \psi_i|\psi\rangle \langle \psi_j|\widehat{A}\psi_i\rangle \psi_j(x) \\ &= \sum_{i,j} \langle \psi_j|\widehat{A}\psi_i\rangle \int \psi_j(x)\psi_i^*(y)\psi(y)d^n y.\end{aligned}$$

This is now (81) with

$$K(x,y) = \sum_{i,j} \langle \psi_j|\widehat{A}\psi_i\rangle \psi_j(x)\psi_i^*(y). \tag{84}$$

Let us show that $K \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$. Remarking that the tensor products $(\psi_j \otimes \psi_i^*)$ form an orthonormal basis in $L^2(\mathbb{R}^n \times \mathbb{R}^n)$ we have

$$\int |K(x,y)|^2 d^n x d^n y \leq \sum_{i,j} |\langle \psi_j|\widehat{A}\psi_i\rangle|^2; \tag{85}$$

applying the Bessel equality (42) to $|\langle \psi_j|\widehat{A}\psi_i\rangle|^2$ we get

$$\sum_{i,j} |\langle \psi_j|\widehat{A}\psi_i\rangle|^2 = \sum_i |\langle \widehat{A}\psi_i|\widehat{A}\psi_i\rangle|^2 < \infty \tag{86}$$

since $\widehat{A}$ is a Hilbert–Schmidt operator. It follows that $K \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$ as claimed. *The condition is sufficient*. Assume that the kernel K of $\widehat{A} \in \mathcal{B}(L^2(\mathbb{R}^n))$ belongs to $L^2(\mathbb{R}^n \times \mathbb{R}^n)$. Since $(\psi_j \otimes \psi_i^*)$ is an orthonormal basis in $L^2(\mathbb{R}^n \times \mathbb{R}^n)$ we can find numbers c_{ij} such that $\sum_{i,j} |c_{ij}|^2 < \infty$ and

$$K(x,y) = \sum_{i,j} c_{ij}\psi_j(x) \otimes \psi_i^*(y). \tag{87}$$

Define now the operator $\widehat{A}$ by the equality (81); we have

$$\begin{aligned}\widehat{A}\psi(x) &= \sum_{i,j} c_{ij}\psi_j(x) \int \psi_i^*(y)\psi(y)d^n y \\ &= \sum_{i,j} c_{ij}\langle \psi_i|\psi\rangle \psi_j(x)\end{aligned}$$

and hence, since the basis $(\psi_j)_j$ is orthonormal,

$$\widehat{A}\psi_k(x) = \sum_{i,j} c_{ij}\langle \psi_i|\psi_k\rangle \psi_j(x) = \sum_j c_{kj}\psi_j(x) \tag{88}$$

so that

$$\textstyle\sum_k \|\widehat{A}\psi_k\|^2 = \sum_{j,k} |c_{kj}|^2 < \infty \tag{89}$$

and $\widehat{A}$ is thus a Hilbert–Schmidt operator. (ii) In view of property the algebra property (see (i) in Theorem 5) a trace class operator is *a fortiori* a Hilbert–Schmidt operator. The claim follows in view of the statement (i). □

4 The Phase Space Picture

From now on we assume that the Hilbert space $\mathcal{H}$ is $L^2(\mathbb{R}^n)$, the space of complex-valued square integrable functions on $\mathbb{R}^n$ (we are thus dealing with quantum systems with n degrees of freedom).

4.1 The Weyl correspondence

4.1.1 Weyl operators and symbols

Let us first explain what we mean by a Weyl symbol. Recall that a function $K(x,y)$ defined on $\mathbb{R}^n \times \mathbb{R}^n$ is the kernel of an operator $\widehat{A}$ if we have

$$\widehat{A}\psi(x) = \int K(x,y)\psi(y)d^n y \tag{90}$$

for all $\psi \in L^2(\mathbb{R}^n)$. A deep theorem from functional analysis (*Schwartz's kernel theorem*) tells us that every continuous linear operator from spaces of test functions to the tempered distributions (In this context, we use the word *distribution* in the sense of L. Schwartz's *generalized functions*) can be represented in this way, the integral in (90) being possibly replaced by a distributional bracket. By definition, the Weyl symbol of the operator $\widehat{A}$ is the function

$$a(x,p) = \int e^{-\frac{i}{\hbar}py} K(x + \tfrac{1}{2}y, x - \tfrac{1}{2}y)d^n y; \tag{91}$$

this formula is easily inverted using an inverse Fourier transform in p, yielding the expression of the kernel in terms of the symbol:

$$K(x,p) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{\frac{i}{\hbar}p(x-y)} a(\tfrac{1}{2}(x+y), p)d^n p. \tag{92}$$

These two formulas uniquely define the kernel and the symbol in terms of each other, and imply the *Weyl correspondence (or transform)*, which expresses unambiguously the operator $\widehat{A}$ in terms of the symbol a:

$$\widehat{A}\psi(x) = \left(\tfrac{1}{2\pi\hbar}\right)^n \iint e^{\frac{i}{\hbar}p(x-y)} a(\tfrac{1}{2}(x+y), p)\psi(y)d^n y d^n p; \tag{93}$$

one often writes $\widehat{A} = \mathrm{Op_W}(a)$, and this notation is unambiguous because the symbol of $\widehat{A}$ is uniquely determined by (91). Formula (93) can be rewritten in several different ways; one common expression is

$$\widehat{A}\psi(x) = \left(\tfrac{1}{\pi\hbar}\right)^n \iint a(x_0,p_0)\widehat{\Pi}(x_0,p_0)\psi(x)d^nx_0d^np_0 \tag{94}$$

where $\widehat{\Pi}(x_0,p_0)$ is the reflection (or parity) operator [42, 44, 46]

$$\widehat{\Pi}(x_0,p_0)\psi(x) = e^{\frac{2i}{\hbar}p_0(x-x_0)}\psi(2x_0-x). \tag{95}$$

The usefulness of the Weyl correspondence in quantum mechanics comes from the fact that it associates to real symbols self-adjoint operators. In fact, more generally:

$$\widehat{A} = \mathrm{Op^W}(a) \implies \widehat{A}^\dagger = \mathrm{Op^W}(a^*). \tag{96}$$

We refer to de Gosson [42, 44, 46] for detailed discussions of the Weyl correspondence from the point of view outlined above; Littlejohn's well-cited paper [47] contains a very nice review of the topic with applications to semiclassical approximations.

4.1.2 Twisted products and convolutions

Composing Weyl operators leads in a natural way to the notion of *twisted product*, which is essential in the theory of deformation quantization: assume that the two operators $\widehat{A} = \mathrm{Op_W}(a)$ and $\widehat{B} = \mathrm{Op_W}(b)$ can be composed, and set $\widehat{A}\widehat{B} = \widehat{C} = \mathrm{Op_W}(c)$. Then the Weyl symbol c is given by the expression $c = a\times b$ where

$$(a\times b)(z) = \left(\tfrac{1}{4\pi\hbar}\right)^{2n}\iint e^{\frac{i}{2\hbar}\sigma(z',z'')}a(z+\tfrac{1}{2}z')b(z-\tfrac{1}{2}z'')d^{2n}z'd^{2n}z''. \tag{97}$$

Setting $u = z+\frac{1}{2}z'$ and $v = z-\frac{1}{2}z''$ we have we have $d^{2n}z'd^{2n}z'' = 4^{2n}d^{2n}ud^{2n}v$ so this formula can be rewritten

$$(a\times b)(z) = \left(\tfrac{1}{\pi\hbar}\right)^n\iint e^{\frac{2i}{\hbar}\partial\sigma(u,z,v)}a(u)b(v)d^{2n}ud^{2n}v \tag{98}$$

where $\partial\sigma(u,z,v)$ is the coboundary of $\sigma(u,v)$ viewed as a 1-chain:

$$\partial\sigma(u,z,v) = \sigma(u,z)-\sigma(u,v)+\sigma(z,v). \tag{99}$$

For detailed proofs of these properties see de Gosson [44, Chapter 10]. The function

$$c = a\times b \tag{100}$$

is called the twisted product of a and b (It was defined by J. von Neumann following the work of Weyl). Thus, by definition,

$$\mathrm{Op_W}(a\times b) = \mathrm{Op_W}(a)\,\mathrm{Op_W}(b). \tag{101}$$

To the twisted product is associated the *twisted convolution* $a\#b$ of two symbols; it is defined by

$$a\#b = F_\sigma(F_\sigma a\times F_\sigma b) \tag{102}$$

or, equivalently, by

$$F_\sigma(a\#b) = F_\sigma a\times F_\sigma b \tag{103}$$

where F_σ is the symplectic Fourier transform. The equivalence of both definitions is due to the fact that the symplectic Fourier transform is its own inverse. Explicitly (de Gosson, [46, § 11.1]):

$$a\#b(z) = \left(\tfrac{1}{2\pi\hbar}\right)^n\int e^{\frac{i}{2\hbar}\sigma(z,z')}a_\sigma(z-z')b_\sigma(z')d^2z'; \tag{104}$$

an equivalent statement is to say that the twisted symbol c_σ of the product $\widehat{C} = \widehat{A}\widehat{B}$ is given by

$$c_\sigma(z) = \left(\tfrac{1}{2\pi\hbar}\right)^n\int e^{\frac{i}{2\hbar}\sigma(z,z')}a_\sigma(z-z')b_\sigma(z')d^{2n}z'. \tag{105}$$

4.2 The Wigner function of a density matrix

The essential point to understand now is that the Wigner function $\rho = W_{\widehat{\rho}}(x,p)$ we are going to define below is (up to an unimportant constant factor) the *Weyl symbol* of the operator $\widehat{\rho}$: the Wigner function is thus a *dequantization* of $\widehat{\rho}$, that is a phase space function obtained from this operator.(It is perhaps a little bit daring to speak about *dequantization* in this context since the Wigner functon is not really a classical object.) Also notice that it is the first time Planck's constant appears in a quite explicit way; we could have *a priori* replaced $\hbar$ with any other real parameter η: this change wouldn't have consequence for the involved mathematics (but it would of course change the physics!). We will come back to this essential point later, but

4.2.1 Definition of the Wigner function of a density matrix

To a density matrix $\widehat{\rho}$ on $L^2(\mathbb{R}^n)$ one associates in standard texts its Wigner function (also called Wigner distribution). It is the function $W_{\widehat{\rho}}$ of the variables $x = (x_1,\dots,x_n)$ and of the conjugate momenta $p = (p_1,\dots,p_n)$ usually

defined in physics texts by

$$W_{\widehat{\rho}}(x,p) = \left(\tfrac{1}{\pi\hbar}\right)^n \int e^{-\frac{2i}{\hbar}px'} \langle x + x' | \widehat{\rho} | x - x' \rangle d^n x' \quad (106)$$

where $|x\rangle$ is an eigenstate of the operator $\widehat{x} = (\widehat{x}_1, \ldots, \widehat{x}_n)$ (where $\widehat{x}_j$ = multiplication by x_j). Performing the change of variables $x \longmapsto y = 2x'$ we can rewrite this definition in the equivalent form

$$W_{\widehat{\rho}}(x,p) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{-\frac{i}{\hbar}py} \langle x + \tfrac{1}{2}y | \widehat{\rho} | x - \tfrac{1}{2}y \rangle d^n y; \quad (107)$$

this has some practical advantages when one uses the Wigner–Weyl–Moyal formalism. In spite of their formal elegance, formulas (106), (107) are at first sight somewhat obscure and need to be clarified, especially if one wants to work analytically with them. Assume first that $\widehat{\rho}$ represents a pure state: $\widehat{\rho} = \widehat{\rho}_\psi = |\psi\rangle\langle\psi|$ where $\psi \in L^2(\mathbb{R}^n)$ is normalized. We get, using the relations $\psi(x) = \langle x|\psi\rangle$ and $\psi^*(x) = \langle\psi|x\rangle$,

$$\begin{aligned} \langle x + x' | \widehat{\rho} | x - x' \rangle &= \langle x + x' | \psi \rangle \langle \psi | x - x' \rangle \\ &= \psi(x + x')\psi^*(x - x') \end{aligned}$$

and hence $W_{\widehat{\rho}_\psi}(x,p) = W\psi(x,p)$ where

$$W\psi(x,p) = \left(\tfrac{1}{\pi\hbar}\right)^n \int e^{-\frac{2i}{\hbar}px'} \psi(x + x')\psi^*(x - x') d^n x' \quad (108)$$

is the usual Wigner function (or Wigner distribution, or Wigner transform) of $\psi \in L^2(\mathbb{R}^n)$ (see Wigner [48], Hillery *et al.* [49]; for the mathematical theory [42, 44, 46]); equivalently

$$W\psi(x,p) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{-\frac{i}{\hbar}py} \psi(x + \tfrac{1}{2}y)\psi^*(x - \tfrac{1}{2}y) d^n y. \quad (109)$$

In the general case, where $\widehat{\rho} = \sum_j \alpha_j |\psi_j\rangle\langle\psi_j|$ is a convex sum of operators of the type above one immediately gets, by linearity, the expression

$$W_{\widehat{\rho}}(x,p) = \sum_j \alpha_j W\psi_j(x,p). \quad (110)$$

A very important result we will prove later on (Theorem 9), but use immediately, is the following:

> The Wigner function of a mixed state is square integrable: $W_{\widehat{\rho}} \in L^2(\mathbb{R}^{2n})$.

One also often uses the cross-Wigner transform of a pair of square integrable functions. It is given by

$$W(\psi,\phi)(x,p) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{-\frac{i}{\hbar}py} \psi(x + \tfrac{1}{2}y)\phi^*(x - \tfrac{1}{2}y) d^n y. \quad (111)$$

It naturally appears as an interference term when calculating the Wigner function of a sum; in fact, using definition (109) one immediately checks that

$$W(\psi + \phi) = W\psi + W\phi + 2\,\mathrm{Re}\, W(\psi,\phi). \quad (112)$$

Notice that $W(\psi,\phi)$ is in general a complex number and that

$$W(\psi,\phi)^* = W(\phi,\psi). \quad (113)$$

The cross-Wigner function has many applications; in particular it allows to reformulate the notion of weak-value as an interference between the past and the future in the time-symmetric approach to quantum mechanics; see Section 4.2.3 below.

4.2.2 The Weyl symbol of a density matrix

In the case of density matrices we have:

Theorem 7. *Let $\widehat{\rho}$ be a density matrix on $L^2(\mathbb{R}^n)$:*

$$\widehat{\rho} = \sum_j \alpha_j |\psi_j\rangle\langle\psi_j| \ \text{ with } \alpha_j \geq 0 \text{ and } \sum_j \alpha_j = 1. \quad (114)$$

The Weyl symbol of $\widehat{\rho}$ is $a = (2\pi\hbar)^n\rho$ where

$$\rho(x,p) = W_{\widehat{\rho}}(x,p) = \sum_j \alpha_j W\psi_j(x,p) \quad (115)$$

is the Wigner function of $\widehat{\rho}$.

Proof. The action of the projection $\widehat{\rho}_j = |\psi_j\rangle\langle\psi_j|$ on a vector $\psi \in L^2(\mathbb{R}^n)$ is given by

$$\widehat{\rho}_j\psi(x) = \langle\psi_j|\psi\rangle\psi_j(x) = \int \psi_j^*(y)\psi(y)\psi_j(x) d^n y \quad (116)$$

hence the kernel of $\widehat{\rho}_j$ is the function

$$K_j(x,y) = \psi_j(x)\psi_j^*(y). \quad (117)$$

It follows, using formula (91), that the Weyl symbol of $\widehat{\rho}_j$ is

$$\begin{aligned} a_j(x,p) &= \int e^{-\frac{i}{\hbar}py} \psi_j(x + \tfrac{1}{2}y)\psi_j^*(x - \tfrac{1}{2}y) d^n y \\ &= (2\pi\hbar)^n W\psi_j(x,p). \end{aligned}$$

Formula (115) follows by linearity. □

4.2.3 Statistical interpretation of the Wigner function

The importance of the Wigner function of a density matrix comes from the fact that we can use it as a substitute for an ordinary probability density for calculating averages (it is precisely for this purpose Wigner introduced his

eponymous transform in [48]). For all $\psi \in L^2(\mathbb{R}^n)$ such that $\psi \in L^1(\mathbb{R}^n)$ and $F\psi \in L^1(\mathbb{R}^n)$ (The conditions $\psi \in L^1(\mathbb{R}^n)$ and $F\psi \in L^1(\mathbb{R}^n)$ are necessary to ensure the convergence of the x and p integrals. The condition $\psi \in L^2(\mathbb{R}^n)$ is not sufficient for the marginal properties to hold (see e.g. Daubechies [15])) the marginal properties are

$$\int W\psi(x,p)d^np = |\psi(x)|^2 \tag{118}$$

$$\int W\psi(x,p)d^nx = |F\psi(p)|^2 \tag{119}$$

and hence, in particular,

$$\iint W\psi(x,p)d^npd^nx = 1 \text{ if } ||\psi|| = 1. \tag{120}$$

In the second equality (119)

$$F\psi(p) = \left(\tfrac{1}{2\pi\hbar}\right)^{n/2} \int e^{-\frac{i}{\hbar}px}\psi(x)d^nx \tag{121}$$

is the $\hbar$-Fourier transform of ψ. One should be aware of the fact that while $W\psi$ is always real (and hence so is $\rho = W_{\widehat{\rho}}$) as can be easily checked by taking the complex conjugates of both sides of the equality (109), it takes negative values for all ψ which are not Gaussian functions. (This is the celebrated *Hudson theorem* [50]; also see Janssen [51] for the multidimensional case.) A *caveat*: this result is only true for the Wigner function $W\psi$ of a single function ψ; the case of a general distribution $\rho = \sum_j \alpha_j W\psi_j$ is much subtler, and will be discussed later.

Let us introduce the following terminology: we call an observable $\widehat{A}$ a *good observable* for the density matrix $\widehat{\rho}$ if its Weyl symbol a (i.e. the corresponding classical observable) satisfies $a\rho \in L^1(\mathbb{R}^{2n})$, that is

$$\int |a(z)\rho(z)|d^{2n}z < \infty \tag{122}$$

(ρ the Wigner function of $\widehat{\rho}$; we are using the shorthand notation $z = (x,p)$, $d^{2n}z = d^nxd^np$). We assume in addition that a is real so that $\widehat{A}$ is Hermitian. Notice that *goodness* is guaranteed if the symbol a is square integrable, because the Cauchy–Schwarz inequality then implies that

$$\left(\int |a(z)\rho(z)|d^{2n}z\right)^2 \leq \int \rho(z)^2 d^{2n}z \int |a(z)|^2 d^{2n}z < \infty \tag{123}$$

since ρ is square integrable (as mentioned above, see Theorem 9).

Theorem 8. *Let $\widehat{\rho}$ be a density matrix on $L^2(\mathbb{R}^n)$ and ρ its Wigner function. The average value of every good observable $\widehat{A}$ with respect to $\widehat{\rho}$ is then finite and given by the formula*

$$\langle\widehat{A}\rangle_{\widehat{\rho}} = \int a(z)\rho(z)d^{2n}z. \tag{124}$$

Proof. By linearity it suffices to consider the case where $\widehat{\rho} = |\psi\rangle\langle\psi|$ so that $\rho = W\psi$; this reduces the proof of formula (124) to that of the simpler equality

$$\langle\widehat{A}\rangle_{\psi} = \int a(x,p)W\psi(x,p)d^nxd^np. \tag{125}$$

Replacing in the equality above $W\psi(x,p)$ by its integral expression (109) yields

$$\langle\widehat{A}\rangle_{\psi} = \left(\tfrac{1}{2\pi\hbar}\right)^n \iint a(x,p)\left(\int e^{-\frac{i}{\hbar}py}\psi(x+\tfrac{1}{2}y)\psi^*(x-\tfrac{1}{2}y)d^ny\right)d^nxd^np. \tag{126}$$

Since we assume that the *goodness* assumption (122) is satisfied, we can use Fubini's theorem and rewrite this equality as a double integral:

$$\langle\widehat{A}\rangle_{\psi} = \left(\tfrac{1}{2\pi\hbar}\right)^n \iint a(x,p)e^{-\frac{i}{\hbar}py}\psi(x+\tfrac{1}{2}y)\psi^*(x-\tfrac{1}{2}y)d^nyd^nxd^np. \tag{127}$$

Let us perform the change of variables $x' = x + \frac{1}{2}y$ and $y' = x - \frac{1}{2}y$; we have $x = \frac{1}{2}(x'+y')$ and $y = x' - y'$ and hence, using definition (93) of the Weyl operator $\widehat{A}$,

$$\begin{aligned}\langle\widehat{A}\rangle_{\psi} &= \left(\tfrac{1}{2\pi\hbar}\right)^n \iint e^{-\frac{i}{\hbar}p(x'-y')}a(\tfrac{1}{2}(x'+y'),p)\psi(x')\psi^*(y')d^ny'd^nx'd^np \\ &= \left(\tfrac{1}{2\pi\hbar}\right)^n \int\left(\iint e^{-\frac{i}{\hbar}p(x'-y')}a(\tfrac{1}{2}(x'+y'),p)\psi^*(y')d^ny'd^np\right)\psi(x')d^nx'\end{aligned}$$

and hence

$$\langle \widehat{A} \rangle_\psi = \int \widehat{A}\psi^*(x')\psi(x')d^n x' = \langle \widehat{A}\psi|\psi\rangle \tag{128}$$

which we set out to prove. □

We remark that the identity (125) can be extended to the cross-Wigner function (111); in fact, adapting the proof of (125) one sees that if ψ and ϕ are square integrable, then

$$\langle \psi|\widehat{A}\phi\rangle = \int a(x,p)W(\psi,\phi)(x,p)d^n x d^n p. \tag{129}$$

A related notion is that of weak value, at the origin of time symmetric quantum mechanics. In time symmetric quantum mechanics, the state of a system is represented by a two-state vector $\langle\phi|\ |\psi\rangle$ where the state $\langle\phi|$ evolves backwards from the future and the state $|\psi\rangle$ evolves forwards from the past. By definition, if $\langle\phi|\psi\rangle \neq 0$, the complex number

$$\langle \widehat{A} \rangle_{\phi,\psi} = \frac{\langle\phi|\widehat{A}|\psi\rangle}{\langle\phi|\psi\rangle} \tag{130}$$

is the *weak value* of $\widehat{A}$. One proves (de Gosson and de Gosson [52, 53]) using (129) that

$$\langle \widehat{A} \rangle_{\phi,\psi} = \frac{1}{\langle\phi|\psi\rangle}\iint a(x,p)W(\psi,\phi)(x,p)d^n p d^n x \tag{131}$$

This relation allows to build time symmetric quantum mechanics entirely within the Weyl–Wigner formalism.

4.3 The displacement operator and the ambiguity function

In this subsection we review a few properties of the Wigner function which are perhaps not all so well-known in quantum mechanics; these properties are important because they give an insight into some of the subtleties of the Weyl–Wigner–Moyal transform. We also define a related transform, the ambiguity function.

4.3.1 Redefinition of the Wigner function

Recall that the reflection operator (95) is explicitly given by the formula

$$\widehat{\Pi}(x_0,p_0)\psi(x) = e^{\frac{2i}{\hbar}p_0(x-x_0)}\psi(2x_0-x). \tag{132}$$

It can be used to define the Wigner function in a very concise way. In fact: for every $\psi \in L^2(\mathbb{R}^n)$ we have

$$W\psi(x_0,p_0) = \left(\tfrac{1}{\pi\hbar}\right)^n \langle\psi|\widehat{\Pi}(x_0,p_0)\psi\rangle. \tag{133}$$

This is easy to verify: we have, by definition (95) of $\widehat{\Pi}(x_0,p_0)$,

$$\langle\psi|\widehat{\Pi}(x_0,p_0)\psi\rangle = \int e^{\frac{2i}{\hbar}p_0(x-x_0)}\psi(2x_0-x)\psi^*(x)d^n x; \tag{134}$$

setting $y = 2(x_0-x)$ we have $x = x_0-\frac{1}{2}y$, $2x_0-x = x_0+\frac{1}{2}y$, and $d^n x = 2^{-n}d^n y$ hence

$$\langle\psi|\widehat{\Pi}(x_0,p_0)\psi\rangle = 2^{-n}\int e^{-\frac{i}{\hbar}p_0 y}\psi(x_0+\tfrac{1}{2}y)\psi^*(x_0-\tfrac{1}{2}y)d^n y \tag{135}$$

which proves (133), taking definition (109) of the Wigner function into account. Formula (133) shows quite explicitly that, up to the factor $(\pi\hbar)^{-n}$, the Wigner function is the probability amplitude for the state $|\psi\rangle$ to be in the state $|\widehat{\Pi}(x_0,p_0)\psi\rangle$; this was actually already observed by Grossmann [54] and Royer [55] in the mid 1970s.

4.3.2 The Moyal identity

An important equality satisfied by the Wigner function is *Moyal's identity* (It is sometimes also called the *orthogonality relation* for the Wigner function)

$$\int W\psi(z)W\phi(z)d^{2n}z = \left(\tfrac{1}{2\pi\hbar}\right)^n |\langle\psi|\phi\rangle|^2 \tag{136}$$

(see [44, 46] for a proof); it is valid for all square-integrable functions ψ and ϕ on $\mathbb{R}^n$. In particular:

$$\int W\psi(z)^2 d^{2n}z = \left(\tfrac{1}{2\pi\hbar}\right)^n \|\psi\|^4. \tag{137}$$

This formula implies the following interesting fact which is not immediately obvious: consider the spectral decomposition (53) of a density operator in Theorem 3:

$$\widehat{\rho}\psi = \sum_j \lambda_j\langle\psi_j|\psi\rangle\psi_j \tag{138}$$

here the λ_j are the eigenvalues of $\widehat{\rho}$ and the corresponding eigenvectors ψ_j form an orthonormal basis of $\mathcal{H}$. When $\mathcal{H} = L^2(\mathbb{R}^n)$ the corresponding Wigner function is therefore

$$\widehat{\rho}\psi = \sum_j \lambda_j W\psi_j. \tag{139}$$

It follows from Moyal's identity (136) that the $W\psi_j$ form an orthonormal system of vectors in the Hilbert space $L^2(\mathbb{R}^{2n})$ (but not a basis as is easily seen by *dimension count*).

As a consequence of the Moyal identity we prove the fact, mentioned in Section 4.2.1, that the Wigner function of a density matrix is square integrable.

Theorem 9. *Let $\{(\psi_j,\alpha_j)\}$ be a mixed state ($\psi_j \in L^2(\mathbb{R}^n)$, $\alpha_j \geq 0$, $\sum_j \alpha_j = 1$). The Wigner function $\rho = W_{\widehat{\rho}}$ is square integrable: $\rho \in L^2(\mathbb{R}^{2n})$.*

Proof. Since $L^2(\mathbb{R}^{2n})$ is a vector space it is sufficient to consider the pure case, that is to prove that $W\psi \in L^2(\mathbb{R}^{2n})$ if $\psi \in L^2(\mathbb{R}^{2n})$. But this immediately follows from Moyal's identity (137). □

The Moyal identity can be extended to the cross-Wigner function (111); recall that for $\psi, \phi \in L^2(\mathbb{R}^n)$ it is defined by

$$W(\psi,\phi)(x,p) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{-\frac{i}{\hbar}py}\psi(x+\tfrac{1}{2}y)\phi^*(x-\tfrac{1}{2}y)d^n y. \tag{140}$$

In fact, for all $\psi, \psi', \phi, \phi' \in L^2(\mathbb{R}^n)$ we have

$$\int W(\psi,\psi')^*(z)W(\phi,\phi')(z)d^{2n}z = \left(\tfrac{1}{2\pi\hbar}\right)^n \langle\psi|\phi\rangle\langle\psi'|\phi'\rangle^* \tag{141}$$

(see for instance de Gosson [44, 46]). Denoting by $\langle\langle\cdot|\cdot\rangle\rangle$ the inner product on $L^2(\mathbb{R}^{2n})$ this identity can be written in the form

$$\langle\langle W(\psi,\psi')|W(\phi,\phi')\rangle\rangle = \left(\tfrac{1}{2\pi\hbar}\right)^n \langle\psi|\phi\rangle\langle\psi'|\phi'\rangle^* \tag{142}$$

In particular

$$\int |W(\psi,\psi')(z)|^2 d^{2n}z = \left(\tfrac{1}{2\pi\hbar}\right)^n \|\psi\|^2\,\|\psi'\|^2. \tag{143}$$

An important remark: one can prove [44, 46], using this generalized Moyal identity, that if vectors ψ_j form an orthonormal basis of $L^2(\mathbb{R}^n)$ then the vectors $(2\pi\hbar)^{n/2}W(\psi_j,\psi_k)$ form an orthonormal basis of the space $L^2(\mathbb{R}^{2n})$ (that these vectors are orthonormal is clear from (142)).

4.3.3 The ambiguity function

A transform closely related to the Wigner function and well-known from signal analysis (especially radar theory) is the *ambiguity function* $\mathrm{Amb}\,\psi$ (it is also called the *auto-correlation function*). It can be introduced in several equivalent ways; we begin by defining it explicitly by a formula: for $\psi \in L^2(\mathbb{R}^n)$

$$\mathrm{Amb}\,\psi(x,p) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{-\frac{i}{\hbar}py}\psi(y+\tfrac{1}{2}x)\psi^*(y-\tfrac{1}{2}x)d^n y. \tag{144}$$

Comparing with the definition

$$W\psi(x,p) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{-\frac{i}{\hbar}py}\psi(x+\tfrac{1}{2}y)\psi^*(x-\tfrac{1}{2}y)d^n y \tag{145}$$

of the Wigner function one cannot help being surprised by the similarity of both definitions. In fact, it is easy to show by performing an elementary change of variables that if ψ is an even function (that is $\psi(-x) = \psi(x)$ for all $x \in \mathbb{R}^n$) then $W\psi$ and $\mathrm{Amb}\,\psi$ are related by

$$\mathrm{Amb}\,\psi(x,p) = 2^{-n}W\psi(\tfrac{1}{2}x,\tfrac{1}{2}p). \tag{146}$$

There are two complementary *natural* ways to define the ambiguity function. The first is to observe that the Wigner function and the ambiguity function are symplectic Fourier transforms of each other:

$$\mathrm{Amb}\,\psi = F_\sigma W\psi \quad\text{and}\quad W\psi = F_\sigma\,\mathrm{Amb}\,\psi; \tag{147}$$

they are of course equivalent since $F_\sigma^{-1} = F_\sigma$. For a proof, see de Gosson [44, 46]. There is still another way to define the ambiguity function. Let $\widehat{D}(z_0) = \widehat{D}(x_0,p_0)$ be the *Weyl displacement operator* (it is also called the Glauber–Sudarshan displacement operator, or the Heisenberg operator, or the Heisenberg–Weyl operator). It is defined by

$$\widehat{D}(z_0)\psi(x) = e^{\frac{i}{\hbar}(p_0x-\frac{1}{2}p_0x_0)}\psi(x-x_0). \tag{148}$$

This operator is the time-one propagator for the Schrödinger equation associated with the classical translation Hamiltonian $\sigma(z,z_0) = x_0p - p_0x$ (see the discussions in de Gosson [2, 42, 44] and Littlejohn [47]); this observation motivates the notation

$$\widehat{D}(z_0) = e^{-\frac{i}{\hbar}\sigma(\widehat{z},z_0)} = e^{-\frac{i}{\hbar}(x_0\widehat{p}-p_0\widehat{x})} \tag{149}$$

often found in the literature. We are using here the coordinate expression of the displacement operator; we leave it to the reader as an exercise to check that $\widehat{D}(z_0)$ coincides with the operator

$$D(\alpha) = \exp\left[\frac{i}{\hbar}(\alpha a^\dagger - \alpha^* a)\right] \tag{150}$$

commonly used in quantum optics (a and $a^\dagger$ are the annihilation and creation operators; see Potoček [56] for a discussion of these notational issues). The displacement operator is related to the reflection operator $\widehat{\Pi}(z_0)$ by the simple formula

$$\widehat{\Pi}(z_0) = \widehat{D}(z_0)\Pi\widehat{D}(z_0)^\dagger \tag{151}$$

where Π is the parity operator $\Pi\psi(x) = \psi(-x)$. That the operators $\widehat{D}(z_0)$ correspond to translations in phase space quantum mechanics is illustrated by the following important relation satisfied by the Wigner transform:

$$W(\widehat{D}(z_0)\psi)(z) = W\psi(z-z_0) \tag{152}$$

(it is easily proven by a direct computation, see de Gosson [42, 44, 46], Littlejohn [47]). Using the displacement operator, the ambiguity function is given by

$$\mathrm{Amb}\,\psi(z_0) = \left(\tfrac{1}{2\pi\hbar}\right)^n \langle\widehat{D}(z_0)\psi|\psi\rangle; \tag{153}$$

one verifies by a direct calculation using (148) that one recovers the first analytical definition (144).

The displacement operators play a very important role, not only in quantum mechanics, but also in related disciplines such as harmonic analysis, signal theory, and time-frequency analysis. They can be viewed as a representation of the canonical commutation relations (the Schrödinger representation of the Heisenberg group); this is related to the fact these operators satisfy

$$\widehat{D}(z_0)\widehat{D}(z_1) = e^{\frac{i}{\hbar}\sigma(z_0,z_1)}\widehat{D}(z_1)\widehat{D}(z_0) \tag{154}$$

and also

$$\widehat{D}(z_0+z_1) = e^{-\frac{i}{2\hbar}\sigma(z_0,z_1)}\widehat{D}(z_0)\widehat{D}(z_1). \tag{155}$$

The second formula shows that the displacement operators form a projective representation of the phase space translation group. In addition to being used to define the ambiguity function, the displacement operators allow one to define Weyl operators in terms of their *twisted symbol* (sometimes also called *covariant symbol*), which is by definition the symplectic Fourier transform

$$a_\sigma(z) = F_\sigma a(z) \tag{156}$$

of the ordinary symbol a. Let in fact $\widehat{A} = \mathrm{Op_W}(a)$, that is

$$\widehat{A} = \left(\tfrac{1}{\pi\hbar}\right)^n \int a(z_0)\widehat{\Pi}(z_0)d^{2n}z_0 \tag{157}$$

(formula 94). Using the displacement operator $\widehat{D}(z_0)$ in place of the reflection operator $\widehat{\Pi}(z_0)$ we have

$$\widehat{A} = \left(\tfrac{1}{2\pi\hbar}\right)^n \int a_\sigma(z_0)\widehat{D}(z_0)d^{2n}z_0 \tag{158}$$

(see [42,44,46,47]). This formula has many applications; it is essential in the study of the positivity properties of trace class operators as we will see in a moment. Notice that formula (158) is Weyl's original definition [57] in disguise: making the change of variables $z_0 \longmapsto -Jz_0$ in this formula one gets, noting that $a_\sigma(-Jz_0) = Fa(z_0)$ and $\widehat{D}(-Jz_0) = e^{-\frac{i}{\hbar}(x_0\widehat{x}+p_0\widehat{p})}e^{-\frac{i}{\hbar}(x_0\widehat{p}-p_0\widehat{x})}$

$$\widehat{A} = \left(\tfrac{1}{2\pi\hbar}\right)^n \iint Fa(x,p)e^{\frac{i}{\hbar}(x\widehat{x}+p\widehat{p})}d^npd^nx \tag{159}$$

which is the formula originally proposed by Weyl [57], in analogy with the Fourier inversion formula (see the discussion in de Gosson [2,46]).

4.4 Calculating traces: rigorous theory

4.4.1 Kernels and symbols

It is tempting to redefine the trace of a Weyl operator $\widehat{A} = \mathrm{Op_W}(a)$ by the formula

$$\mathrm{Tr}(\widehat{A}) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int a(z)d^{2n}z. \tag{160}$$

But doing this one should not forget that even if the operator $\widehat{A}$ is of trace class, formula (160) need not give the actual trace. First, the integral in the right-hand side might not be convergent; secondly even if it is we have to prove that it really yields the right result. We will discuss the validity of formula (160) and of other similar formulas below, but let us first prove some intermediary results.

We begin by discussing the Weyl symbols of Hilbert–Schmidt and trace class operators.

Theorem 10. *Let $\widehat{A} = \mathrm{Op^W}(a)$ be a Hilbert–Schmidt operator. Then $a \in L^2(\mathbb{R}^{2n})$ and we have*

$$\int |a(z)|^2 d^{2n}z = (2\pi\hbar)^{n/2} \iint K(x,y)d^nxd^ny. \tag{161}$$

Conversely, every Weyl operator with symbol $a \in L^2(\mathbb{R}^{2n})$ is a Hilbert–Schmidt operator.

Proof. In view of Theorem 6 we have $K \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$. Let us prove formula (161) when $K \in \mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n)$; it will then hold by continuity for arbitrary $K \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$ in view of the density of $\mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n)$ in $L^2(\mathbb{R}^n \times \mathbb{R}^n)$. In view of formula (91) relating the kernel and the symbol of a Weyl operator we have

$$\int |a(x,p)|^2 dp = (2\pi\hbar)^n \int |K(x+\tfrac{1}{2}y, x-\tfrac{1}{2}y)|^2 d^ny. \tag{162}$$

Integrating this equality with respect to the x variables we get, using Fubini's theorem

$$\begin{aligned}\int |a(z)|^2 d^{2n}z &= (2\pi\hbar)^n \int \left(\int |K(x+\tfrac{1}{2}y, x-\tfrac{1}{2}y)|^2 d^ny\right) d^nx \\ &= (2\pi\hbar)^n \iint |K(x+\tfrac{1}{2}y, x-\tfrac{1}{2}y)|^2 d^nxd^ny.\end{aligned}$$

Set now $x' = x + \frac{1}{2}y$ and $y' = x - \frac{1}{2}y$; we have $d^nx'd^ny' = d^nxd^ny$ and hence

$$\int |a(z)|^2 d^{2n}z = (2\pi\hbar)^n \iint |K(x',y')|^2 d^nx'd^ny' \tag{163}$$

which we set out to prove. The converse is obvious since the condition $a \in L^2(\mathbb{R}^{2n})$ is equivalent to $K \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$ in view of the inequality above. □

4.4.2 Rigorous results

We now specialize our discussion to the case $\mathcal{H} = L^2(\mathbb{R}^n)$. Recall that we showed in Section 3.2 that the formula

$$\langle \widehat{A}|\widehat{B}\rangle_{\mathrm{HS}} = \mathrm{Tr}(\widehat{A}^\dagger\widehat{B}) \tag{164}$$

defines an inner product on the ideal $\mathcal{L}_2(\mathcal{H})$ of Hilbert–Schmidt operators in $\mathcal{H}$ the associated *trace norm* being defined by

$$\|\widehat{A}\|_{\mathrm{HS}} = \mathrm{Tr}(\widehat{A}^\dagger\widehat{A})^{1/2} \tag{165}$$

Let us state and prove the following result:

Theorem 11. *Let $\widehat{A} = \mathrm{Op}^{\mathrm{W}}(a)$ and $\widehat{B} = \mathrm{Op}^{\mathrm{W}}(a)$ be Hilbert–Schmidt operators: $\widehat{A}, \widehat{B} \in \mathcal{L}_2(\mathcal{H})$. (i) The trace class operator $\widehat{A}\widehat{B}$ has trace*

$$\mathrm{Tr}(\widehat{A}\widehat{B}) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int a(z)b(z)d^{2n}z; \tag{166}$$

(ii) The Hilbert–Schmidt inner product is given by the convergent integral

$$\langle \widehat{A}|\widehat{B}\rangle_{\mathrm{HS}} = \left(\tfrac{1}{2\pi\hbar}\right)^n \int a^*(z)b(z)d^{2n}z \tag{167}$$

and hence $\|\widehat{A}\|^2_{\mathrm{HS}} = \mathrm{Tr}(\widehat{A}^\dagger\widehat{A})$ is given by

$$\|\widehat{A}\|^2_{\mathrm{HS}} = \left(\tfrac{1}{2\pi\hbar}\right)^n \int |a(z)|^2 d^{2n}z. \tag{168}$$

Proof. (i) We first observe that in view of Theorem 10 we have $a \in L^2(\mathbb{R}^n)$ and $b \in L^2(\mathbb{R}^n)$ hence the integrals in (166) and (167) are absolutely convergent. Let (ψ_j) be an orthonormal basis of $L^2(\mathbb{R}^n)$; by definition of the trace we have

$$\mathrm{Tr}(\widehat{A}\widehat{B}) = \sum_j \langle \psi_j|\widehat{A}\widehat{B}\psi_j\rangle = \sum_j \langle \widehat{A}^\dagger\psi_j|\widehat{B}\psi_j\rangle. \tag{169}$$

Expanding $\widehat{B}\psi_j$ and $\widehat{A}^\dagger\psi_j$ in the basis (ψ_j) we get

$$\widehat{B}\psi_j = \sum_k \langle \psi_k|\widehat{B}\psi_j\rangle\psi_k \ , \ \widehat{A}^\dagger\psi_j = \sum_\ell \langle \widehat{A}\psi_\ell|\psi_j\rangle\psi_\ell \tag{170}$$

and hence, using the Bessel equality (42),

$$\langle \widehat{A}^\dagger\psi_j|\widehat{B}\psi_j\rangle = \sum_k \langle \widehat{A}^\dagger\psi_j|\psi_k\rangle\langle \widehat{B}\psi_j|\psi_k\rangle^* \tag{171}$$

In view of formula (129), we have

$$\langle \widehat{A}^\dagger\psi_j|\psi_k\rangle = \int a(z)W(\psi_j,\psi_k)(z)d^{2n}z \tag{172}$$

$$\langle \widehat{B}\psi_j|\psi_k\rangle = \int b^*(z)W(\psi_j,\psi_k)(z)d^{2n}z \tag{173}$$

where $W(\psi_j,\psi_k)$ is the cross-Wigner transform (111) of ψ_j,ψ_k; denoting by $\langle\langle\cdot|\cdot\rangle\rangle$ the inner product on $L^2(\mathbb{R}^{2n})$ these equalities can be rewritten

$$\langle \widehat{A}^\dagger\psi_j|\psi_k\rangle = \langle\langle a^*|W(\psi_j,\psi_k)\rangle\rangle \tag{174}$$

$$\langle \widehat{B}\psi_j|\psi_k\rangle = \langle\langle b|W(\psi_j,\psi_k)\rangle\rangle \tag{175}$$

and hence it follows from the extended Moyal identity (141) that

$$\mathrm{Tr}(\widehat{A}\widehat{B}) = \sum_{j,k} \langle\langle a^*|W(\psi_j,\psi_k)\rangle\rangle\langle\langle b|W(\psi_j,\psi_k)\rangle\rangle^*. \tag{176}$$

Since (ψ_j) is an orthonormal basis the vectors $(2\pi\hbar)^{n/2}W(\psi_j,\psi_k)$ also form an orthonormal basis (see the remark following formula (143)), hence the Bessel identity (42) allows us to write the equality above as

$$\mathrm{Tr}(\widehat{A}\widehat{B}) = \left(\tfrac{1}{2\pi\hbar}\right)^n \langle\langle a^*|b\rangle\rangle \tag{177}$$

which is formula (166). (ii) It immediately follows from formula (166) using (164) and (165), recalling that if $\widehat{A} = \mathrm{Op}^{\mathrm{W}}(a)$ then $\widehat{A}^\dagger = \mathrm{Op}^{\mathrm{W}}(a^*)$. □

Part (i) of the result above allows us—at last!—to express the trace of a Weyl operator in terms of its symbol provided that the latter is absolutely integrable:

Corollary 12. *Let $\widehat{A} = \mathrm{Op}^{\mathrm{W}}(a)$ be a trace class operator. If in addition we have $a \in L^1(\mathbb{R}^n)$ then*

$$\mathrm{Tr}(\widehat{A}) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int a(z)d^{2n}z. \tag{178}$$

Proof. (Cf. Du and Wong [58, Theorem 2.4].) It is equivalent to prove that the symplectic Fourier transform $a_\sigma = F_\sigma a$ satisfies

$$\mathrm{Tr}(\widehat{A}) = a_\sigma(0). \tag{179}$$

Writing $\widehat{A} = \widehat{B}\widehat{C}$ where $\widehat{B}$ and $\widehat{C}$ are Hilbert–Schmidt operators we have

$$\mathrm{Tr}(\widehat{A}) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int b(z)c(z)d^{2n}z \tag{180}$$

hence it suffices to show that

$$a_\sigma(0) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int b(z)c(z)d^{2n}z. \tag{181}$$

We have, in view of formula (105) giving the twisted symbol of the product of two Weyl operators

$$a_\sigma(z) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{\frac{i}{2\hbar}\sigma(z,z')} b_\sigma(z-z')c_\sigma(z')d^{2n}z' \tag{182}$$

and hence, using the Plancherel identity (12) for F_σ,

$$\begin{aligned} a_\sigma(0) &= \left(\tfrac{1}{2\pi\hbar}\right)^n \int b_\sigma(-z')c_\sigma(z')d^{2n}z' \\ &= \left(\tfrac{1}{2\pi\hbar}\right)^n \int b(z)c(z)d^{2n}z \end{aligned} \tag{183}$$

which proves the formula (179). □

The following result is very much in the spirit of the C^*-algebraic approach outlined in Section 2.1.3 (cf. Gracia-Bondía and Várilly [59, 60]). Let us denote the Weyl transform by Op^{W}: it thus associates to every symbol $a \in \mathcal{S}'(\mathbb{R}^n)$ the Weyl operator $\widehat{A} = \mathrm{Op}^{\mathrm{W}}(a)$.

Theorem 13. *The restriction of Op^{W} to the Hilbert space $L^2(\mathbb{R}^n)$ is an isomorphism of $L^2(\mathbb{R}^n)$ onto the algebra $\mathcal{L}_2(L^2(\mathbb{R}^n))$ of Hilbert–Schmidt operators on $L^2(\mathbb{R}^n)$.*

Proof. In view of Theorem 10, a bounded operator $\widehat{A}$ on $L^2(\mathbb{R}^n)$ is Hilbert–Schmidt if and only if its Weyl symbol a is in $L^2(\mathbb{R}^n)$. The Weyl correspondence being linear and one-to-one the statement follows. □

5 Metaplectic Group and Symplectic Covariance

For a complete study of the metaplectic group in quantum mechanics see our book [61]; on a slightly more general and technical level see [42,44]. An excellent introduction to the symplectic group is given in Arvind *et al*. [62], also see García-Bullé *et al*. [63].

5.1 The metaplectic representation

5.1.1 The generators of $\mathrm{Sp}(n)$

Recall that the symplectic form on phase space $\mathbb{R}^{2n}$ can be defined by $\sigma(z,z') = (z')^T Jz$ where

$$J = \begin{pmatrix} 0_{n\times n} & I_{n\times n} \\ -I_{n\times n} & 0_{n\times n} \end{pmatrix} \tag{184}$$

is the standard symplectic matrix (we use the notation $z = (x,p)$, $z' = (x',p')$). By definition, the symplectic group $\mathrm{Sp}(n)$ consists of all real $2n\times 2n$ matrices S such that $\sigma(Sz,Sz') = \sigma(z,z')$ for all vectors z,z'; such a matrix is called a *symplectic matrix*. Rewriting this condition as $(Sz')^T JSz = (z')^T Jz$ we thus have $S \in \mathrm{Sp}(n)$ if and only if $S^T JS = J$. It is an easy exercise to show that if S is symplectic then S^{-1} and S^T are symplectic as well, hence this defining relation is equivalent to $SJS^T = J$. The symplectic group plays an essential role in classical mechanics in its Hamiltonian formulation; its role in quantum mechanics is no less important, in association with its double covering, the metaplectic group $\mathrm{Mp}(n)$ which we briefly describe now.

There are several ways to introduce the metaplectic group. We begin by giving a definition using the notion of free symplectic matrix and its generating function (we are following here our presentation in [42,44]). Let

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{185}$$

be a symplectic matrix, where the *blocks* A,B,C,D are $n\times n$ matrices. It is easy to show that the relations $SJS^T = S^T JS = J$ are equivalent to the two groups of conditions

$$A^T C,\ B^T D \text{ are symmetric, and } A^T D - C^T B = I \tag{186}$$

$$AB^T,\ CD^T \text{ are symmetric, and } AD^T - BC^T = I \tag{187}$$

One says that the block-matrix (185) is a *free symplectic matrix* if B is invertible, i.e. $\det B \neq 0$. To a free symplectic matrix is associated a generating function: it is the quadratic form

$$\mathcal{A}(x,x') = \frac{1}{2}DB^{-1}x\cdot x - B^{-1}x\cdot x' + \frac{1}{2}B^{-1}Ax'\cdot x'. \tag{188}$$

The terminology comes from the fact that the knowledge of $\mathcal{A}(x,x')$ uniquely determines the free symplectic matrix S: we have

$$\begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} x' \\ p' \end{pmatrix} \Longleftrightarrow \begin{cases} p = \nabla_x \mathcal{A}(x,x') \\ p' = -\nabla_{x'}\mathcal{A}(x,x') \end{cases} \tag{189}$$

as can be verified by a direct calculation. The interest of the notion of free symplectic matrix comes from the fact that such matrices generate the symplectic group $\mathrm{Sp}(n)$. More precisely every $S \in \mathrm{Sp}(n)$ can be written as a product $S = S_{\mathcal{A}}S_{\mathcal{A}'}$ (we place the corresponding generating functions $\mathcal{A}$ and $\mathcal{A}'$ as subscripts).

Defining, for symmetric P and invertible L, the symplectic matrices V_{-P} and M_L by

$$V_{-P} = \begin{pmatrix} I & 0 \\ P & I \end{pmatrix}\ ,\ M_L = \begin{pmatrix} L^{-1} & 0 \\ 0 & L^T \end{pmatrix} \tag{190}$$

a straightforward calculations shows that the free symplectic matrix $S_{\mathcal{A}}$ can be factored as

$$S_{\mathcal{A}} = V_{-DB^{-1}}M_{B^{-1}}JV_{-B^{-1}A}. \tag{191}$$

This implies that the symplectic group $\mathrm{Sp}(n)$ is generated by the set of all matrices V_{-P} and M_L together with J. It is easy to deduce from this that the determinant of a symplectic matrix always is equal to one since we obviously have $\det V_{-P} = \det M_L = \det J$.

5.1.2 Generalized Fourier transforms

Now, to every free symplectic matrix $S_{\mathcal{A}}$ we associate two operators $\widehat{S}_{\mathcal{A},m}$ by the formula

$$\widehat{S}_{\mathcal{A},m}\psi(x) = \left(\tfrac{1}{2\pi\hbar}\right)^{\frac{n}{2}} i^{m-\frac{n}{2}}\sqrt{|\det B^{-1}|}\int e^{\frac{i}{\hbar}\mathcal{A}(x,x')}\psi(x')d^n x' \tag{192}$$

where m corresponds to a choice of argument for $\det B^{-1}$: $m = 0 \bmod 2$ if $\det B^{-1} > 0$ and $m = 1 \bmod 2$ if $\det B^{-1} < 0$. It is not difficult to prove that the generalized Fourier transforms $\widehat{S}_{\mathcal{A},m}$ are unitary operators on $L^2(\mathbb{R}^n)$. These operators generate a group: the metaplectic group $\mathrm{Mp}(n)$. One shows that, as for the symplectic group, every $\widehat{S} \in \mathrm{Mp}(n)$ can be written (non uniquely) as a product $\widehat{S}_{\mathcal{A},m}\widehat{S}_{\mathcal{A}',m'}$. This group is a double covering of $\mathrm{Sp}(n)$, the covering projection being simply defined by

$$\pi_{\mathrm{Mp}} : \mathrm{Mp}(n) \longrightarrow \mathrm{Sp}(n)\ ,\ \pi_{\mathrm{Mp}}(\widehat{S}_{\mathcal{A},m}) = S_{\mathcal{A}}. \tag{193}$$

Here are three examples of free symplectic matrices and of their metaplectic counterparts; these can be used to give an alternative definition of the metaplectic group:

- *The standard symplectic matrix* J has as generating function $\mathcal{A}(x,x') = -x\cdot x'$ and hence the two corresponding metaplectic operators are $\pm\widehat{J}$ with

$$\widehat{J}\psi(x) = \left(\tfrac{1}{2\pi\hbar i}\right)^{n/2}\int e^{-\frac{i}{\hbar}x\cdot x'}\psi(x')d^n x'; \tag{194}$$

observe that $\widehat{J} = i^{-n/2}F$ where F is the usual Fourier transform;

- *The symplectic shear* $V_{-P} = \begin{pmatrix} I & 0 \\ P & I \end{pmatrix}$ $(P = P^T)$ is not free, but

$$U_{-P} = JV_{-P}J^{-1} = \begin{pmatrix} I & -P \\ 0 & I \end{pmatrix} \tag{195}$$

is if $\det P \neq 0$. In this case we have

$$\mathcal{A}(x, x') = -\frac{1}{2}P^{-1}x \cdot x + P^{-1}x \cdot x' \tag{196}$$

and the corresponding metaplectic operators are hence $\pm\widehat{U}_{-P}$ with

$$\widehat{U}_{-P} = \left(\tfrac{1}{2\pi\hbar}\right)^{\frac{n}{2}} i^{\frac{n}{2}} |\det P|^{-1} e^{-\frac{i}{2\hbar}P^{-1}x\cdot x} \times \int e^{\frac{i}{\hbar}P^{-1}x\cdot x'}\psi(x')d^n x'. \tag{197}$$

- *The symplectic rescaling matrix* $M_L = \begin{pmatrix} L^{-1} & 0 \\ 0 & L^T \end{pmatrix}$ is not free but the product

$$M_L J = \begin{pmatrix} 0 & L^{-1} \\ L^T & 0 \end{pmatrix} \tag{198}$$

is and has $\mathcal{A}(x, x') = Lx \cdot x'$ as generating function; the corresponding metaplectic operator are $\widehat{M}_{L,m}\widehat{J}$ where

$$\widehat{M}_{L,m}\psi(x) = i^m \sqrt{|\det L|}\psi(Lx) \tag{199}$$

the integer m (the *Maslov index*) corresponding to a choice of $\arg\det L$.

It turns out that an easy calculation shows that, similarly to the factorization (191) of free symplectic matrices, the quadratic Fourier transform (192) can be written

$$\widehat{S}_{\mathcal{A},m} = \widehat{V}_{-B^{-1}A}\widehat{M}_{B^{-1},m}\widehat{J}\widehat{V}_{-DB^{-1}} \tag{200}$$

where $\widehat{M}_{L,m}$ and $\widehat{J}$ are defined as above and

$$\widehat{V}_{-P}\psi(x) = e^{\frac{i}{2\hbar}Px^2}\psi(x) \tag{201}$$

when $P = P^T$. It follows that the elementary operators $\widehat{V}_{-P}$, $\widehat{M}_{L,m}$ and $\widehat{J}$ generate $\mathrm{Mp}(n)$ (these operators are used in many texts to define the metaplectic group; our approach using (192) has some advantages since among other things it makes immediately clear that metaplectic operators are generalized Fourier transforms).

The factorization $\widehat{S} = \widehat{S}_{\mathcal{A},m}\widehat{S}_{\mathcal{A}',m'}$ of a metaplectic operator is by no means unique; for instance we can write the identity operator I as $\widehat{S}_{\mathcal{A},m}\widehat{S}^{-1}_{\mathcal{A},m} = \widehat{S}_{\mathcal{A},m}\widehat{S}_{\mathcal{A}^*,m^*}$ for every quadratic Fourier transform $\widehat{S}_{\mathcal{A},m}$. There is however an invariant attached to $\widehat{S}$: the Maslov index. Denoting by $\mathrm{Inert}\, R$ the index of inertia (= the number of negative eigenvalues) of the real symmetric matrix R we have:

Proposition 14. *Let* $\widehat{S} = \widehat{S}_{\mathcal{A},m}\widehat{S}_{\mathcal{A}',m'} = \widehat{S}_{\mathcal{A}'',m''}\widehat{S}_{\mathcal{A}''',m'''}$. *We have*

$$m+m'-\mathrm{Inert}(P'+Q) \equiv m''+m'''-\mathrm{Inert}(P'''+Q'') \mod 4. \tag{202}$$

Proof. See Leray [64], de Gosson [42, 65]. □

It follows from formula (202) that the class modulo 4 of the integer $m + m' - \mathrm{Inert}(P' + Q)$ does not depend on the way we write $\widehat{S} \in \mathrm{Mp}(n)$ as a product $\widehat{S}_{\mathcal{A},m}\widehat{S}_{\mathcal{A}',m'}$ of quadratic Fourier transforms; this class is denoted by $m(\widehat{S})$ and called the Maslov index of $\widehat{S}$. The mapping

$$m : \mathrm{Mp}(n) \in \widehat{S} \longrightarrow m(\widehat{S}) \in \mathbb{Z}_4 \tag{203}$$

is called the Maslov index on $\mathrm{Mp}(n)$. We have $m(\widehat{S}_{\mathcal{A},m}) = m$, mod 4 [64, 65]. The theory of the Maslov index has been further developed by Arnol'd, Leray, and by the author.

5.1.3 Weyl symbol and Conley–Zehnder index of a metaplectic operator

We will call *symplectic isotopy* a C^1 mapping $\Sigma : t \longmapsto S_t \in \mathrm{Sp}(n)$ such that $S_0 = I_{\mathrm{d}}$. If t is restricted to a bounded interval $[0, T]$ we call S_0 the origin and S_T the endpoint of the symplectic isotopy.

We define the following subset of $\mathrm{Sp}(n)$:

$$\mathrm{Sp}_0(n) = \{S \in \mathrm{Sp}(n) : \det(S - I) \neq 0\}. \tag{204}$$

To $S \in \mathrm{Sp}_0(n)$ we associate the family of operators $\widehat{R}_\nu(S)$ defined, for $\nu \in \mathbb{R}$, by

$$\widehat{R}_\nu(S) = \left(\tfrac{1}{2\pi\hbar}\right)^n i^\nu \sqrt{|\det(S - I)|} \int \widehat{T}(Sz_0)\widehat{T}(-z_0)d^{2n}z_0. \tag{205}$$

One verifies that for all $S \in \mathrm{Sp}_0(n)$ and $\nu \in \mathbb{R}$ the operators $\widehat{R}_\nu(S)$ satisfy the intertwining formula

$$\widehat{T}(Sz_0) = \widehat{R}_\nu(S)\widehat{T}(z_0)\widehat{R}_\nu(S)^{-1}. \tag{206}$$

It follows, using the irreducibility of the Schrödinger representation of the Heisenberg group [66], that there exists a constant $c(S, \nu) \in \mathbb{C}$ such that $\widehat{R}_\nu(S) = c(S, \nu)\widehat{S}$ where $\pi^{\mathrm{Mp}}(\widehat{S}) = S$. It is moreover easy to check that the operators are $\widehat{R}_\nu(S)$ unitary, hence $|c(S, \nu)| = 1$.

To the Maslov index m is associated another integer index which plays an essential role in quantum holography:

Definition 15. *Let* $\Sigma = (S_t)_{t\in I}$ *be a symplectic isotopy in* $\mathrm{Sp}(n)$ *with endpoint* $S \notin \mathrm{Sp}_0(n)$ *such that* $S = S_{\mathcal{A}}$ *for some generating function* $\mathcal{A}$. *Let* $\widehat{S}_{\mathcal{A},m} \in \mathrm{Mp}(n)$ *cover* $S_{\mathcal{A}}$. *The integer*

$$\nu(\Sigma) = m - \mathrm{Inert}\, \mathcal{A}_{xx} \tag{207}$$

is called the Conley–Zehnder index of Σ.

Here $\mathcal{A}_{xx}$ is the Hessian matrix (= matrix of second derivatives) of the quadratic form $x \longmapsto \mathcal{A}(x,x)$; $\operatorname{Inert}\mathcal{A}_{xx}$ is the index of inertia (=number of negative eigenvalues) of $\mathcal{A}_{xx}$.

The following result connects the integer ν in (205) to the Conley–Zehnder index when $\widehat{R}_\nu(S) \in \operatorname{Mp}(n)$:

Proposition 16. *Let $\Sigma = (S_t)_{t\in I}$ be symplectic isotopy in $\operatorname{Sp}(n)$ leading from the identity to $S = S_{\mathcal{A}} \notin \operatorname{Sp}_0(n)$. Let $\widehat{\Sigma} = (\widehat{S}_t)_{t\in I}$ be the metaplectic isotopy covering Σ and $\widehat{S} = \widehat{S}_{\mathcal{A},m} \in \operatorname{Mp}(n)$ be its endpoint. We have $\widehat{S} = \widehat{R}_{\nu(\widehat{\Sigma})}(S)$ where $\nu(\widehat{\Sigma}) = \nu(\Sigma) \bmod 4$. That is:*

$$\widehat{S}_{\mathcal{A},m} = \widehat{R}_{m-\operatorname{Inert}\mathcal{A}_{xx}}(S) \tag{208}$$

Proof. See [41–43]. □

The statement above has the following consequences when the endpoint of the symplectic isotopy Σ is a free symplectic matrix $S_{\mathcal{A}}$.

This allows us to give a rigorous explicit formula for the twisted Weyl symbol of $\widehat{S}_{\mathcal{A},m}$:

Corollary 17. *The twisted Weyl symbol of $\widehat{S}_{\mathcal{A},m}$ with $S_{\mathcal{A}} \notin \operatorname{Sp}_0(n)$ is given by*

$$(s_{\mathcal{A}})_\sigma(z) = \frac{i^{m-\operatorname{Inert}\mathcal{A}_{xx}}}{\sqrt{|\det(S_{\mathcal{A}}-I)|}} \exp\left(\frac{i}{2\hbar}M_{\mathcal{A}}z\cdot z\right) \tag{209}$$

where $M_{\mathcal{A}}$ is the symplectic Cayley transform of $S_{\mathcal{A}}$.

Proof. See de Gosson [41, 42]. □

Proposition 16 and formula (209) suggest that the Conley–Zehnder index is related to a choice of argument of the square root of the determinant of $S - I$. This is indeed the case:

Proposition 18. *Let $\widehat{S}_{\mathcal{A},m} \in \operatorname{Mp}(n)$ have projection $S_{\mathcal{A}} \notin \operatorname{Sp}_0(n)$. We have*

$$\nu(\widehat{S}_{\mathcal{A},m}) = n + \frac{1}{\pi}\arg\det(S_{\mathcal{A}}-I) \mod 2. \tag{210}$$

that is

$$\nu(\widehat{S}_{\mathcal{A},m}) = \begin{cases} n \mod 2 & \text{if } S_{\mathcal{A}} \in \operatorname{Sp}_+(n) \\ n+2 \mod 2 & \text{if } S_{\mathcal{A}} \in \operatorname{Sp}_-(n) \end{cases}. \tag{211}$$

The projection $S_{\mathcal{A}} = \pi^{\operatorname{Mp}}(\widehat{S}_{\mathcal{A},m})$ is a free symplectic matrix, and a straightforward calculation yields the factorization

$$S_{\mathcal{A}} - I = \begin{pmatrix} 0 & B \\ I & D-I \end{pmatrix}\begin{pmatrix} C-(D-I)B^{-1}(A-I) & 0 \\ B^{-1}(A-I) & I \end{pmatrix}. \tag{212}$$

Since $S_{\mathcal{A}} \in \operatorname{Sp}(n)$ we have $C - DB^{-1}A = -(B^T)^{-1}$ and hence

$$C-(D-I)B^{-1}(A-I) = B^{-1}A + DB^{-1} - (B^T)^{-1} = \mathcal{A}_{xx} \tag{213}$$

so that

$$S_{\mathcal{A}} - I = \begin{pmatrix} 0 & B \\ I & D-I \end{pmatrix}\begin{pmatrix} \mathcal{A}_{xx} & 0 \\ B^{-1}(A-I) & I \end{pmatrix}. \tag{214}$$

It follows that

$$\det(S_{\mathcal{A}} - I) = (-1)^n \det B \det \mathcal{A}_{xx} \tag{215}$$

and hence

$$\arg\det(S_{\mathcal{A}}-I) = n\pi + \arg\det B + \arg\det\mathcal{A}_{xx} \mod 2\pi. \tag{216}$$

Noticing that $\arg\det\mathcal{A}_{xx} = \pi\operatorname{Inert}\mathcal{A}_{xx}$ and that this is

$$\arg\det(S_{\mathcal{A}}-I) = n\pi + \arg\det B + \pi\operatorname{Inert}\mathcal{A}_{xx} \mod 2\pi. \tag{217}$$

We have $\arg\det(B) = m\pi$ and hence

$$\arg\det(S_{\mathcal{A}}-I) = (n+m-\operatorname{Inert}\mathcal{A}_{xx})\pi \mod 2\pi \tag{218}$$

that is

$$\arg\det(S_{\mathcal{A}}-I) = (n+\nu(\widehat{S}_{\mathcal{A},m}))\pi \mod 2\pi \tag{219}$$

which yields (210).

5.2 Symplectic covariance properties

5.2.1 A conjugation property for the displacement and reflection operators

Everything here stems from the following observation: let $\widehat{S} \in \operatorname{Mp}(n)$ have projection $S \in \operatorname{Sp}(n)$ (the symplectic matrix S is thus *covered* by the two metaplectic operators $\pm\widehat{S}$). Then for every phase space point $z_0 = (x_0, p_0)$ the displacement operators $\widehat{D}(Sz_0)$ and $\widehat{D}(z_0)$ are related by the conjugation formula

$$\widehat{D}(Sz_0) = \widehat{S}\widehat{D}(z_0)\widehat{S}^\dagger \tag{220}$$

(recall that $\widehat{S}^\dagger = \widehat{S}^{-1}$ since metaplectic operators are unitary). This formula is most easily proven using the generators of $\operatorname{Sp}(n)$ and the corresponding generators of $\operatorname{Mp}(n)$; for a complete proof see for instance de Gosson [44, §8.1.3], also [47]. It easily follows from (220) that the reflection operator (95) satisfies a similar relation:

$$\widehat{\Pi}(Sz_0) = \widehat{S}\widehat{\Pi}(z_0)\widehat{S}^\dagger. \tag{221}$$

In fact, recalling (formula (151)) that $\widehat{\Pi}(z_0) = \widehat{D}(z_0)\Pi\widehat{D}(z_0)^\dagger$ we have

$$\widehat{\Pi}(Sz_0) = \widehat{D}(Sz_0)\Pi\widehat{D}(Sz_0)^\dagger = \widehat{S}\widehat{D}(z_0)(\widehat{S}^\dagger\Pi\widehat{S})\widehat{D}(z_0)^\dagger\widehat{S}^\dagger; \tag{222}$$

to get (221) we have to show that $\widehat{S}^{\dagger}\widehat{\Pi}\widehat{S} = \Pi$ since we will then have

$$\widehat{\Pi}(S\,z_0) = \widehat{S}\widehat{D}(z_0)\Pi\widehat{D}(z_0)^{\dagger}\widehat{S}^{\dagger} = \widehat{S}\widehat{\Pi}(z_0)\widehat{S}^{\dagger}. \qquad (223)$$

It suffices for that purpose to show that $\widehat{S}^{\dagger}_{\mathcal{A},m}\Pi\widehat{S}_{\mathcal{A},m} = \Pi$ since the generalized Fourier transforms $\widehat{S}_{\mathcal{A},m}$ generate the metaplectic group. Now, $\Pi\widehat{S}_{\mathcal{A},m}\psi(x) = \widehat{S}_{\mathcal{A},m}\psi(-x)$ hence, by (192),

$$\Pi\widehat{S}_{\mathcal{A},m}\psi(x) = \left(\tfrac{1}{2\pi\hbar}\right)^{\frac{n}{2}} i^{m-\frac{n}{2}}\sqrt{|\det B^{-1}|} \times \int e^{\frac{i}{\hbar}\mathcal{A}(-x,x')}\psi(x')d^n x'. \qquad (224)$$

Noting that $\mathcal{A}(-x,x') = \mathcal{A}(x,-x')$ (cf. formula (188)) we get, making the change of variables $x' \longmapsto -x'$,

$$\Pi\widehat{S}_{\mathcal{A},m}\psi(x) = \left(\tfrac{1}{2\pi\hbar}\right)^{\frac{n}{2}} i^{m-\frac{n}{2}}\sqrt{|\det B^{-1}|} \times \int e^{\frac{i}{\hbar}\mathcal{A}(x,x')}\psi(-x')d^n x' \qquad (225)$$

that is $\Pi\widehat{S}_{\mathcal{A},m} = \widehat{S}_{\mathcal{A},m}\Pi$; it follows that we have

$$\widehat{S}^{\dagger}_{\mathcal{A},m}\Pi\widehat{S}_{\mathcal{A},m} = \widehat{S}^{\dagger}_{\mathcal{A},m}\widehat{S}_{\mathcal{A},m}\Pi = \Pi. \qquad (226)$$

5.2.2 Symplectic covariance

Collecting the facts above we have:

Theorem 19. *Let $z = (x,p)$ be a point in the phase space $\mathbb{R}^{2n}$ and $\widehat{S}$ a metaplectic operator with projection $\pi_{\mathrm{Mp}}(\widehat{S}) = S$ in $\mathrm{Sp}(n)$. (i) We have*

$$W\psi(S\,z) = W(\widehat{S}^{-1}\psi)(z),$$
$$\mathrm{Amb}\,\psi(S\,z) = \mathrm{Amb}(\widehat{S}^{-1}\psi)(z). \qquad (227)$$

(ii) For every symbol a we have

$$\mathrm{Op_W}(a\circ S^{-1}) = \widehat{S}\,\mathrm{Op_W}(a)\widehat{S}^{\dagger} \qquad (228)$$

where $a\circ S^{-1}(z) = a(S^{-1}z)$.

Proof. (i) To prove the first identity (227) we recall (formula (133)) that

$$W\psi(z) = \left(\tfrac{1}{\pi\hbar}\right)^n \langle\psi|\widehat{\Pi}(z)\psi\rangle \qquad (229)$$

and hence, using (221) and the unitarity of metaplectic operators,

$$W\psi(S\,z) = \left(\tfrac{1}{\pi\hbar}\right)^n \langle\psi|\widehat{S}\widehat{\Pi}(z)\widehat{S}^{\dagger}\psi\rangle = \left(\tfrac{1}{\pi\hbar}\right)^n \langle\widehat{S}^{\dagger}\psi|\widehat{\Pi}(z)\widehat{S}^{\dagger}\psi\rangle \qquad (230)$$

which is precisely (227). The proof of the second identity (227) is similar using the definition (153) of the ambiguity function together with property (220). (ii) Recall (formula (94)) that the Weyl operator $\widehat{A} = \mathrm{Op_W}(a)$ can be written

$$\widehat{A} = \left(\tfrac{1}{\pi\hbar}\right)^n \int a(z)\widehat{\Pi}(z)d^{2n}z \qquad (231)$$

and hence, using (221),

$$\widehat{S}\widehat{A}\widehat{S}^{-1} = \left(\tfrac{1}{\pi\hbar}\right)^n \int a(z)\widehat{S}\widehat{\Pi}(z)\widehat{S}^{-1}d^{2n}z = \left(\tfrac{1}{\pi\hbar}\right)^n \int a(z)\widehat{\Pi}(S\,z)d^{2n}z;$$

performing the change of variables $z' = S\,z$ we have, since $\det S = 1$,

$$\widehat{S}\widehat{A}\widehat{S}^{-1} = \left(\tfrac{1}{\pi\hbar}\right)^n \int a(S^{-1}z)\widehat{\Pi}(z)d^{2n}z = \mathrm{Op_W}(a\circ S^{-1}) \qquad (232)$$

as claimed. □

Applying the machinery above to the density matrix we get:

Corollary 20. *Let $\{(\psi_j,\alpha_j)\}$ be a mixed state with density matrix $\widehat{\rho}$ and Wigner function ρ. Let $\widehat{S} \in \mathrm{Mp}(n)$. The mixed state $\{(\widehat{S}\psi_j,\alpha_j)\}$ has density matrix $\widehat{S}\widehat{\rho}\widehat{S}^{\dagger}$ and Wigner function $\rho(S^{-1}z)$, where $S = \pi_{\mathrm{Mp}}(\widehat{S})$.*

Proof. The Wigner distribution of $\{(\widehat{S}\psi_j,\alpha_j)\}$ is

$$\sum_j \alpha_j W(\widehat{S}\psi_j)(z) = \sum_j \alpha_j W\psi_j(S^{-1}z) \qquad (233)$$

because of the first formula (227). It follows that the Weyl symbol of the density matrix corresponding to $\{(\widehat{S}\psi_j,\alpha_j)\}$ is $a(S^{-1}z)$ where $a = (2\pi\hbar)^n\rho$ is the Weyl symbol of $\widehat{\rho}$; that $\widehat{S}\widehat{\rho}\widehat{S}^{\dagger}$ is the density matrix of $\{(\widehat{S}\psi_j,\alpha_j)\}$ follows from formula (228). □

Note that the fact that $\widehat{S}\widehat{\rho}\widehat{S}^{\dagger}$ is the density matrix of $\{(\widehat{S}\psi_j,\alpha_j)\}$ can also be proven directly using the definition (14) of the density matrix in term of projectors.

6 Variable Planck Constant

We now address a controversial theme of this review, namely the mathematical consequences of possible changes in the value of Planck's constant h.

6.1 A consequence of Moyal's identity

We begin by a few straightforward observations involving the Moyal identity discussed in Section 4.3.2. Let η be a real parameter; we assume for the moment that $\eta > 0$. This parameter will play the role of a variable $\hbar = h/2\pi$. For any square integrable ψ we define the η-Wigner

transform (or distribution) of ψ by replacing $\hbar$ by η in the usual definition:

$$W_\eta\psi(x,p) = \left(\tfrac{1}{2\pi\eta}\right)^n \int e^{-\frac{i}{\eta}py}\psi(x+\tfrac{1}{2}y)\psi^*(x-\tfrac{1}{2}y)d^ny. \tag{234}$$

Of course $W_\hbar\psi = W\psi$ (the usual Wigner transform). The mathematical properties of $W_\eta\psi$ are of course the same as those of $W\psi$, replacing $\hbar$ everywhere with η. In particular, replacing the $\hbar$-Fourier transform (121) with the η-Fourier transform

$$F_\eta\psi(p) = \left(\tfrac{1}{2\pi\eta}\right)^{n/2} \int e^{-\frac{i}{\eta}px}\psi(x)d^nx \tag{235}$$

the marginal properties (119) become

$$\begin{aligned} \int W_\eta\psi(x,p)d^np &= |\psi(x)|^2, \\ \int W_\eta\psi(x,p)d^nx &= |F_\eta\psi(p)|^2 \end{aligned} \tag{236}$$

($\psi \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$).

The Moyal identity yields

$$\int W_\eta\psi(z)W_\eta\phi(z)d^{2n}z = \left(\tfrac{1}{2\pi\eta}\right)^n |\langle\psi|\phi\rangle|^2 \tag{237}$$

which is valid for all square integrable functions ψ and ϕ (see de Gosson [46]). In particular

$$\int W_\eta\psi(z)^2d^{2n}z = \left(\tfrac{1}{2\pi\eta}\right)^n \|\phi\|^4. \tag{238}$$

Let us now address the following question: for a given ψ, can we find ϕ such that $W_\eta\phi = W\psi$ for $\eta \neq \hbar$? The answer is "no"! More generally:

Theorem 21. *(i) A pure state $|\psi\rangle$ does not remain a pure state if we vary $\hbar$: let $W\psi$ be the Wigner function of $|\psi\rangle$. There does not exist any state $|\phi\rangle$ such that $W_\eta\phi = W\psi$ if $\eta \neq \hbar$. (ii) Assume that $|\psi\rangle$ becomes a mixed state when $\hbar$ is replaced with η. Then we must have $\eta \leq \hbar$.*

Proof. (i) We have $\psi, \phi \in L^2(\mathbb{R}^n)$. Assume that $W_\eta\phi = W\psi$; then

$$\int W_\eta\phi(x,p)d^np = \int W\psi(x,p)d^np \tag{239}$$

hence, using the marginal properties (119), $|\phi(x)|^2 = |\psi(x)|^2$ so that ϕ and ψ have same norm: $\|\phi\| = \|\psi\|$. On the other hand, using the Moyal identity (238), the equality $W\psi = W_\eta\phi$ implies that

$$\begin{aligned} \int W\psi(z)^2d^{2n}z &= \left(\tfrac{1}{2\pi\hbar}\right)^n \|\psi\|^4 \\ \int W_\eta\phi(z)d^{2n}z &= \left(\tfrac{1}{2\pi\eta}\right)^n \|\phi\|^4 \end{aligned}$$

hence $\eta = \hbar$. (ii) Assume that there exists a sequence (ϕ_j) of (normalized) functions in $L^2(\mathbb{R}^n)$ and a sequence of positive constants α_j summing up to one such that $W\psi = \sum_j \alpha_j W_\eta\phi_j$. Proceeding as above we get, using again the marginal properties,

$$\|\psi\|^2 = \sum_j \alpha_j\|\phi_j\|^2. \tag{240}$$

On the other hand, squaring $W\psi$ we get

$$(W\psi)^2 = \sum_{j,k} \alpha_j\alpha_k W_\eta\phi_j W_\eta\phi_k \tag{241}$$

hence, integrating and using respectively the Moyal identity for $(W\psi)^2$ and $W_\eta\phi_jW_\eta\phi_k$, and the Cauchy–Schwarz inequality we get

$$\begin{aligned} \left(\tfrac{1}{2\pi\hbar}\right)^n \|\psi\|^4 &= \left(\tfrac{1}{2\pi\eta}\right)^n \sum_{j,k} \alpha_j\alpha_k|\langle\phi_j|\phi_k\rangle|^2 \\ &\leq \left(\tfrac{1}{2\pi\eta}\right)^n \sum_{j,k} \alpha_j\alpha_k\|\phi_j\|^2\|\phi_k\|^2 \\ &= \left(\tfrac{1}{2\pi\eta}\right)^n \left(\textstyle\sum_j \alpha_j\|\phi_j\|^2\right)^2 \\ &= \left(\tfrac{1}{2\pi\eta}\right)^n \|\psi\|^4 \end{aligned}$$

which implies, using (240), that $\left(\frac{1}{2\pi\hbar}\right)^n \leq \left(\frac{1}{2\pi\eta}\right)^n$, that is $\eta \leq \hbar$ as claimed. □

A *caveat*: property (ii) in the theorem above does not say that a pure state automatically becomes a mixed state if we decrease Planck's constant. It merely says that if a pure state becomes mixed, it can only happen if Planck's constant has decreased. We will see later that this is related to the uncertainty principle.

6.2 The Quantum Bochner Theorem

6.2.1 Bochner's theorem

We begin by recalling Bochner's theorem about the Fourier transform of a probability density. That theorem says that a (complex valued) function f on $\mathbb{R}^m$, continuous at the origin and such that $f(0) = 1$ is the characteristic function of a probability density on $\mathbb{R}^m$ if and only if it is of *positive type*, that is, if for all choices of points $z_1, \ldots, z_N \in \mathbb{R}^m$ the $N \times N$ matrix

$$F_{(N)} = (f(z_j - z_k))_{1\leq j,k\leq N} \tag{242}$$

is positive semidefinite (that is, the eigenvalues of $F_{(N)}$ are all ≥ 0).

Let us introduce two modifications of the symplectic Fourier transform F_σ. First we allow the latter to depend on an arbitrary parameter $\eta \neq 0$ and set

$$F_{\sigma,\eta}a(z) = a_{\sigma,\eta}(z) = \left(\tfrac{1}{2\pi\eta}\right)^n \int e^{-\frac{i}{\eta}\sigma(z,z')}a(z')d^{2n}z'. \tag{243}$$

It coincides with F_σ when $\eta = \hbar$. We next define the reduced symplectic Fourier transform $F_\diamond$ is by

$$a_\diamond(z) = F_\diamond a(z) = \int e^{-i\sigma(z,z')} a(z') d^{2n}z'. \tag{244}$$

Obviously $F_\diamond a$ and $F_{\sigma,\eta} a = a_{\sigma,\eta}$ are related by the simple formula

$$a_\diamond(z) = (2\pi\eta)^n a_{\sigma,\eta}(\eta z). \tag{245}$$

With this notation Bochner's theorem on Fourier transforms of probability measures can be restated in the following way: a real function ρ on $\mathbb{R}^{2n}$ is a probability density if and only if its reduced symplectic Fourier transform $\rho_\diamond$ is continuous, $\rho_\diamond(0) = 1$, and for all choices of $z_1, \ldots, z_N \in \mathbb{R}^{2n}$ the $N \times N$ matrix Λ whose entries are the complex numbers $\rho_\diamond(z_j - z_k)$ is positive semidefinite:

$$\Lambda = (\rho_\diamond(z_j - z_k))_{1 \leq j,k \leq N} \geq 0 \tag{246}$$

(A matrix is said to be positive semidefinite if all its eigenvalues are ≥ 0).

When condition (246) is satisfied one says that the reduced symplectic Fourier transform $\rho_\diamond$ is of *positive type*.

6.2.2 The notion of η-positivity

The notion of η-positivity, due to Kastler [67], generalizes Bochner's notion: let $a \in \mathcal{S}'(\mathbb{R}^{2n})$ and η a real number; we say that $a_\diamond$ is of *η-positive type* if for every integer N the $N \times N$ matrix $\Lambda_{(N)}$ with entries

$$\Lambda_{jk} = e^{-\frac{i\eta}{2}\sigma(z_j,z_k)} a_\diamond(z_j - z_k) \tag{247}$$

is positive semidefinite (which we write ≥ 0 for short) for all choices of $(z_1, z_2, \ldots, z_N) \in (\mathbb{R}^{2n})^N$:

$$\Lambda_{(N)} = (\Lambda_{jk})_{1 \leq j,k \leq N} \geq 0. \tag{248}$$

The condition (248) is equivalent to the polynomial inequalities

$$\sum_{1 \leq j,k \leq N} \lambda_j \lambda_k^* e^{-\frac{i\eta}{2}\sigma(z_j,z_k)} a_\diamond(z_j - z_k) \geq 0 \tag{249}$$

for all $N \in \mathbb{N}$, $\lambda_j, \lambda_k \in \mathbb{C}$, and $z_j, z_k \in \mathbb{R}^{2n}$. If a is of η-positive type then it is also of of $(-\eta)$-positive type as is immediately seen by taking the complex conjugate of the left-hand side of (249).

When $\eta \neq 0$ we can rewrite conditions (248)–(249) using the symplectic η-Fourier transform: replacing (z_j, z_k) with $\eta^{-1}(z_k, z_j)$ and noting that $\sigma(z_k, z_j) = -\sigma(z_j, z_k)$ the conditions (248) are equivalent to

$$\Lambda'_{(N)} = (\Lambda'_{jk}(z_j, z_k))_{1 \leq j,k \leq N} \geq 0 \tag{250}$$

where

$$\Lambda'_{jk}(z_j, z_k) = e^{\frac{i}{2\eta}\sigma(z_j,z_k)} a_{\sigma,\eta}(z_j - z_k). \tag{251}$$

The polynomial conditions (249) become in this case

$$\sum_{1 \leq j,k \leq N} \lambda_j \lambda_k^* e^{\frac{i}{2\eta}\sigma(z_j,z_k)} a_{\sigma,\eta}(z_j - z_k) \geq 0 \tag{252}$$

6.2.3 KLM condition and the quantum Bochner theorem

We are now going to prove an essential result (the *quantum Bochner theorem*) originally due to Kastler [67], and Loupias and Miracle-Sole [68, 69]; also see Parthasarathy [70, 71] and Parthasarathy and Schmidt [72] for different points of view. The proof we will give is simpler than that in [67–69], which uses the theory of C^*-algebras; our proof is partially based on the discussions in [73–75]. For this we will need a technical result from linear algebra (*Schur's Lemma*), which says that the entrywise product of two positive semidefinite matrices is also positive semidefinite:

Lemma 22 (Schur). *Let $A = (A_{jk})_{1 \leq j,k \leq N}$ and $B = (B_{jk})_{1 \leq j,k \leq N}$ be two symmetric matrices with the same finite dimension N. Defining the Hadamard product of these matrices by*

$$A \circ B = (A_{jk} B_{jk})_{1 \leq j,k \leq N} \tag{253}$$

then if A and B both are positive semidefinite, then so is $A \circ B$.

For a proof of this result see for instance Bapat [76].

Theorem 23 (Quantum Bochner). *Let $\widehat{\rho}$ be a self-adjoint trace class operator on $L^2(\mathbb{R}^n)$:*

$$\widehat{\rho}\psi = \sum_j \alpha_j \langle \psi_j | \psi \rangle \psi_j. \tag{254}$$

Let $\rho = \sum_j \alpha_j W\psi_j$ be the Wigner function of $\widehat{\rho}$. We have $\widehat{\rho} \geq 0$ if and only if the two following conditions hold: (i) The reduced symplectic Fourier transform $\rho_\diamond$ is continuous and $\rho_\diamond(0) = 1$; (ii) $\rho_\diamond$ is of η-positive type.

Proof. Let us first show that the conditions (i)–(ii) are necessary. Assume that $\widehat{\rho} \geq 0$; then

$$\rho = \sum_j \alpha_j W_\eta \psi_j \tag{255}$$

for a family of normalized functions $\psi_j \in L^2(\mathbb{R}^n)$, the coefficients α_j being ≥ 0 and summing up to one. It is thus sufficient to show that the Wigner transform $W_\eta\psi$ of an arbitrary $\psi \in L^2(\mathbb{R}^n)$ is of η-positive type. This

amounts to showing that for all $(z_1,\dots,z_N)\in(\mathbb{R}^{2n})^N$ and all $(\lambda_1,\dots,\lambda_N)\in\mathbb{C}^N$ we have

$$I_N(\psi)=\sum_{1\leq j,k\leq N}\lambda_j\lambda_k^*e^{-\frac{i}{2\eta}\sigma(z_j,z_k)}F_{\sigma,\eta}W_\eta\psi(z_j-z_k)\geq 0 \tag{256}$$

for every complex vector $(\lambda_1,\dots,\lambda_N)\in\mathbb{C}^N$ and every sequence $(z_1,\dots,z_N)\in(\mathbb{R}^{2n})^N$ (see condition (252)). Since the η-Wigner function $W_\eta\psi$ and the η-ambiguity Amb_η function are obtained from each other by the symplectic η-Fourier transform

$$F_{\sigma,\eta}a(z)=\left(\tfrac{1}{2\pi\eta}\right)^n\int e^{-\frac{i}{\eta}\sigma(z,z')}a(z')d^{2n}z' \tag{257}$$

we have

$$I_N(\psi)=\sum_{1\leq j,k\leq N}\lambda_j\lambda_k^*e^{-\frac{i}{2\eta}\sigma(z_j,z_k)}\,\mathrm{Amb}_\eta\,\psi(z_j-z_k). \tag{258}$$

Let us prove that

$$I_N(\psi)=\left(\tfrac{1}{2\pi\eta}\right)^n\|\textstyle\sum_{1\leq j\leq N}\lambda_j\widehat{D}_\eta(z_j)\psi\|^2; \tag{259}$$

the inequality (256) will follow. Taking into account the fact that $\widehat{D}_\eta(-z_k)^\dagger=\widehat{D}_\eta(z_k)$ and using the relation (155) which becomes here

$$\widehat{D}_\eta(z_0)\widehat{D}_\eta(z_1)=e^{\frac{i}{2\eta}\sigma(z_0,z_1)}\widehat{D}_\eta(z_0+z_1) \tag{260}$$

we have, expanding the square in the right-hand side of (259),

$$\|\sum_{1\leq j\leq N}\lambda_j\widehat{D}_\eta(z_j)\psi\|^2=\sum_{\substack{1\leq j,\\ k\leq N}}\lambda_j\lambda_k^*\langle\widehat{D}_\eta(z_k)\psi|\widehat{D}_\eta(z_j)\psi\rangle=\sum_{\substack{1\leq j,\\ k\leq N}}\lambda_j\lambda_k^*\langle\widehat{D}_\eta(-z_j)\widehat{D}_\eta(z_k)\psi|\psi\rangle=\sum_{\substack{1\leq j,\\ k\leq N}}\lambda_j\lambda_k^*e^{-\frac{i}{2\eta}\sigma(z_j,z_k)}\langle\widehat{D}_\eta(z_k-z_j)\psi|\psi\rangle;$$

in view of formula (153), which becomes here

$$\mathrm{Amb}_\eta\,\psi(z)=\left(\tfrac{1}{2\pi\eta}\right)^n\langle\widehat{D}_\eta(z)\psi|\psi\rangle \tag{261}$$

we thus have

$$\|\sum_{1\leq j\leq N}\lambda_j\widehat{D}_\eta(z_j)\psi\|^2=(2\pi\eta)^n\sum_{\substack{1\leq j,\\ k\leq N}}\lambda_j\lambda_k^*e^{-\frac{i}{2\eta}\sigma(z_j,z_k)}\,\mathrm{Amb}_\eta\,\psi(z_j-z_k) \tag{262}$$

proving the equality (259).

Let us now show that, conversely, the conditions (i) and (ii) are sufficient, i.e. that they imply that $(\widehat{\rho\psi}|\psi)_{L^2}\geq 0$ for all $\psi\in L^2(\mathbb{R}^n)$; equivalently (see formula (124) in Theorem 8)

$$\int\rho(z)W_\eta\psi(z)d^{2n}z\geq 0 \tag{263}$$

for $\psi\in L^2(\mathbb{R}^n)$. Let us set, as above,

$$\Lambda'_{jk}=e^{\frac{i}{2\eta}\sigma(z_j,z_k)}a_{\sigma,\eta}(z_j-z_k) \tag{264}$$

where z_j and z_k are arbitrary elements of $\mathbb{R}^{2n}$. To say that $a_{\sigma,\eta}$ is of η-positive type means that the matrix $\Lambda'=(\Lambda'_{jk})_{1\leq j,k\leq N}$ is positive semidefinite; choosing $z_k=0$ and setting $z_j=z$ this means that every matrix $(a_{\sigma,\eta}(z))_{1\leq j,k\leq N}$ is positive semidefinite. Setting

$$\begin{aligned}\Gamma_{jk}&=e^{\frac{i}{2\eta}\sigma(z_j,z_k)}F_{\sigma,\eta}W_\eta\psi(z_j-z_k)\\&=e^{\frac{i}{2\eta}\sigma(z_j,z_k)}\,\mathrm{Amb}_\eta\,\psi(z_j-z_k)\end{aligned}$$

the matrix $\Gamma_{(N)}=(\Gamma_{jk})_{1\leq j,k\leq N}$ is positive semidefinite. Let us now write

$$M_{jk}=\mathrm{Amb}_\eta\,\psi(z_j-z_k)\rho_{\sigma,\eta}(z_j-z_k); \tag{265}$$

we claim that the matrix $M_{(N)}=(M_{jk})_{1\leq j,k\leq N}$ is positive semidefinite. In fact, M is the Hadamard product of the positive semidefinite matrices $M'_{(N)}=(M'_{jk})_{1\leq j,k\leq N}$ and $M''_{(N)}=(M''_{jk})_{1\leq j,k\leq N}$ where

$$\begin{aligned}M'_{jk}&=e^{\frac{i}{2\eta}\sigma(z_j,z_k)}\,\mathrm{Amb}_\eta\,\psi(z_j-z_k)\\M''_{jk}&=e^{-\frac{i}{2\eta}\sigma(z_j,z_k)}\rho_{\sigma,\eta}(z_j-z_k)\end{aligned}$$

and Schur's Lemma 22 implies that $M_{(N)}$ is also positive semidefinite. It follows from Bochner's theorem that the function b defined by

$$b_{\sigma,\eta}(z)=\mathrm{Amb}_\eta\,\psi(z)\rho_{\sigma,\eta}(-z)=F_{\sigma,\eta}W\psi(z)\rho_{\sigma,\eta}(-z) \tag{266}$$

is a probability density; in particular we must have $b(0)\geq 0$. Integrating the equality above with respect to z we get,

using the Plancherel formula (12)

$$\int F_{\sigma,\eta}a(z)F_{\sigma,\eta}b(-z)d^{2n}z = \int a(z)b(z)d^{2n}z \qquad (267)$$

for the symplectic η-Fourier transform

$$\begin{aligned}(2\pi\eta)^n b(0) &= \int \mathrm{Amb}_\eta\,\psi(z)\rho_{\sigma,\eta}(-z)d^{2n}z\\ &= \int W_\eta\psi(z)\rho(z)d^{2n}z\end{aligned}$$

hence the inequality (263) since $b(0) \geq 0$. □

6.3 Application to quantum states

6.3.1 The covariance matrix

Let $\rho(z)$ be a real function on phase space $\mathbb{R}^{2n}$. We assume that this function is integrable and that

$$\int \rho(z)d^{2n}z = 1. \qquad (268)$$

and the marginal identities (119) hold:

$$\int \rho(x,p)d^n p = |\psi(x)|^2, \qquad (269)$$

$$\int \rho(x,p)d^n x = |F\psi(p)|^2. \qquad (270)$$

We will in addition assume that ρ decreases sufficiently fast at infinity:

$$\int (1+|z|^2)|\rho(z)|d^{2n}z < \infty. \qquad (271)$$

Setting $z_\alpha = x_\alpha$ if $1 \leq \alpha \leq n$ and $z_\alpha = p_{\alpha-n}$ if $n+1 \leq \alpha \leq 2n$, the covariances and variances of the variables $z = (x,p)$ associated with ρ are the numbers

$$\Delta(z_\alpha, z_\beta) = \int (z_\alpha - \langle z_\alpha\rangle)(z_\beta - \left\langle z_\beta\right\rangle)\rho(z)d^{2n}z \qquad (272)$$

and

$$(\Delta z_\alpha)^2 = \Delta(z_\alpha, z_\alpha) = \int (z_\alpha - \langle z_\alpha\rangle)^2\rho(z)d^{2n}z. \qquad (273)$$

In the formulas (272) and (273) above $\langle z_\alpha\rangle$ is the average with respect to ρ of z_α; more generally one defines for an integer $k \geq 0$ the moments

$$\langle z_\alpha^k\rangle = \int z_\alpha^k\rho(z)d^{2n}z. \qquad (274)$$

Notice that our condition (271) guarantees the existence of both $\langle z_\alpha\rangle$ and $\langle z_\alpha^2\rangle$ as follows from the trivial inequalities

$$\left|\int z_\alpha\rho(z)d^{2n}z\right| \leq \int (1+|z|^2)|\rho(z)|d^{2n}z < \infty \qquad (275)$$

$$\left|\int z_\alpha z_\beta\rho(z)d^{2n}z\right| \leq \int (1+|z|^2)|\rho(z)|d^{2n}z < \infty. \qquad (276)$$

It follows that the quantities (272) and (273) are well-defined in view of condition (271). Since the integral of ρ is equal to one, formulae (272) and (273) can be rewritten as

$$\Delta(z_\alpha, z_\beta) = \langle z_\alpha z_\beta\rangle - \langle z_\alpha\rangle\langle z_\beta\rangle \qquad (277)$$

$$(\Delta z_\alpha)^2 = \Delta(z_\alpha, z_\alpha) = \langle z_\alpha^2\rangle - \langle z_\alpha\rangle^2. \qquad (278)$$

We will call the symmetric $2n \times 2n$ matrix

$$\Sigma = \left(\Delta(z_\alpha, z_\beta)\right)_{1\leq\alpha,\beta\leq 2n} \qquad (279)$$

the covariance matrix associated with ρ. For instance, when $n = 1$

$$\Sigma = \begin{pmatrix}\Delta x^2 & \Delta(x,p)\\ \Delta(p,x) & \Delta p^2\end{pmatrix} \qquad (280)$$

with

$$\Delta x^2 = \langle x^2\rangle - \langle x\rangle^2\,,\ \Delta p^2 = \langle p^2\rangle - \langle p\rangle^2$$
$$\Delta(x,p) = \langle xp\rangle - \langle x\rangle\langle p\rangle.$$

Here is an example: let $\rho(z)$ be given by the formula

$$\rho(z) = (2\pi)^{-n}\sqrt{\det\Sigma^{-1}}e^{-\frac{1}{2}\Sigma^{-1}z^2} \qquad (281)$$

where Σ is a symmetric positive definite real $2n \times 2n$ matrix. A straightforward calculation involving Gaussian integrals shows that Σ is precisely the associated covariance matrix. We will see later that $\rho(z)$ is the Wigner function of a quantum state if Σ satisfies a certain condition related to the uncertainty principle. However, $\rho(z)$ can always be viewed as a classical probability distribution.

6.3.2 Two lemmas

We are going to state two preliminary results which will be used for the proof of Theorem 26 below characterizing the covariance matrix of a quantum state. We will not prove these results here, and refer to the original sources. That we need so much preparatory material is indicative of the difficulty of the topic we address.

We begin with the well-known Williamson's symplectic diagonalization theorem. Let us first recall the following terminology [42,44]: suppose that Σ is a symmetric positive definite real $2n \times 2n$ matrix. Then the matrix ΣJ has the same eigenvalues as the antisymmetric matrix $\Sigma^{1/2}J\Sigma^{1/2}$ (It is antisymmetric because its transpose is $\Sigma^{1/2}J^T\Sigma^{1/2} = -\Sigma^{1/2}J\Sigma^{1/2}$ since $J^T = -J$) and they are therefore of the type $\pm i\lambda_1, \ldots, \pm i\lambda_n$ where $\lambda_j > 0$. These numbers λ_j are called the *symplectic eigenvalues* of the matrix Σ (or sometimes also the *Williamson invariants* of Σ). Williamson's diagonalization result generalizes to the multidimensional case the elementary observation that

every 2×2 real symmetric matrix $\Sigma = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ with $a > 0$ and $ac - b^2 > 0$ can be written $\Sigma = S^T D S$ where

$$S = \begin{pmatrix} \sqrt{a/d} & b/\sqrt{ad} \\ 0 & \sqrt{d/a} \end{pmatrix} \ , \ D = \begin{pmatrix} d & 0 \\ 0 & d \end{pmatrix} \tag{282}$$

and $d = \sqrt{ac - b^2}$.

Lemma 24 (Williamson). *Let Σ be a symmetric positive definite real $2n \times 2n$ matrix. There exists $S \in \mathrm{Sp}(n)$ such that $\Sigma = S^T D S$ where D is the diagonal matrix*

$$D = \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda \end{pmatrix} \tag{283}$$

with $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, the positive numbers λ_j being the symplectic eigenvalues of M.

Proof. See for instance [42,44] and the references therein. □

For this we will need, in addition to Williamson's symplectic diagonalization result, the following technical result:

Lemma 25 (Narcowich). *Let $f(z)$ be a twice differentiable function defined on phase space. If f is of η-positive type then*

$$-f''(0) + \frac{i\eta}{2} J \geq 0 \tag{284}$$

where $f''(0)$ is the matrix of second derivatives (the Hessian matrix) of $f(z)$ at $z = 0$.

Proof. See Narcowich [77, Lemma 2.1]. □

6.3.3 A necessary (but not sufficient) condition for a state to be quantum

We are going to prove an essential result, which goes back to Narcowich [77]. It says that the covariance matrix of a quantum state must satisfy a certain condition which implies—but is stronger than—the Robertson–Schrödinger uncertainty principle.

We recall that it is assumed that the Wigner function ρ satisfies the condition

$$\int (1 + |z|^2)|\rho(z)| d^{2n}z < \infty. \tag{285}$$

This implies, among other things, that the Fourier transform (and hence also the symplectic Fourier transform) of ρ is twice continuously differentiable. In fact, writing

$$F\rho(z) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int e^{-\frac{i}{\hbar}zz'} \rho(z') d^{2n}z' \tag{286}$$

we have

$$\partial_{z_\alpha} F\rho = -\frac{i}{\hbar} F[z_\alpha \rho] \ , \ \partial_{z_\alpha}\partial_{z_\beta} F\rho = \left(-\frac{i}{\hbar}\right)^2 F[z_\alpha z_\beta \rho] \tag{287}$$

and hence

$$|\partial_{z_\alpha} F\rho(z)| \leq \frac{1}{\hbar}\left|\int z_\alpha \rho(z) d^{2n}z\right| < \infty$$

$$|\partial_{z_\alpha}\partial_{z_\beta} F\rho(z)| \leq \left(\frac{1}{\hbar}\right)^2 \left|\int z_\alpha z_\beta \rho(z) d^{2n}z\right| < \infty$$

in view of the inequalities (275), (276).

Theorem 26. *Suppose that the phase space function ρ with associated covariance matrix Σ is the η-Wigner transform of a density matrix $\widehat{\rho}$. (i) We have*

$$\Sigma + \frac{i\eta}{2} J \geq 0. \tag{288}$$

(i.e. the Hermitian matrix $\Sigma + \frac{i\eta}{2}J$ is positive semidefinite); this condition is equivalent to the inequality

$$|\eta| \leq 2\lambda_{\min} \tag{289}$$

where $\lambda_{\min}$ is the smallest symplectic eigenvalue of Σ. (ii) If (288) and (289) hold, then they hold for every $\eta' < \eta$.

Proof. (i) That $\Sigma + \frac{i\eta}{2}J$ is Hermitian is clear: since the adjoint of J is $-J$ we have

$$(\Sigma + \tfrac{i\eta}{2}J)^\dagger = \Sigma^\dagger + (\tfrac{i\eta}{2}J)^\dagger = \Sigma + \tfrac{i\eta}{2}J. \tag{290}$$

We next remark that $\Sigma = \Sigma_0$ where Σ_0 is the covariance matrix of $\rho_0(z) = \rho(z + \langle z\rangle_\rho)$: we have $\langle z\rangle_0 = 0$ and hence

$$\Delta(x_j, x_k)_0 = \int x_j x_k \rho_0(z) d^{2n}z = \Delta(x_j, x_k)_\rho; \tag{291}$$

similarly $\Delta(x_j, p_k)_0 = \Delta(x_j, p_k)_\rho$ and $\Delta(p_j, p_k)_0 = \Delta(p_j, p_k)_\rho$. It is thus sufficient to prove the result for the density operator $\widehat{\rho}_0$. We are going to use Lemma 25 with $f(z) = F_\sigma\rho_0(z) = \rho_{0,\sigma}(z)$. Observing that the (symplectic) Fourier transform $\rho_{0,\sigma}$ is twice continuously differentiable in view of the argument preceding the statement of the theorem, we have

$$\hbar^2 \rho''_{0,\sigma}(0) = (2\pi\hbar)^{-n} \begin{pmatrix} -\Sigma_{0,pp} & \Sigma_{0,xp} \\ \Sigma_{0,px} & -\Sigma_{0,xx} \end{pmatrix} \tag{292}$$

and hence

$$\eta^2 \rho''_{0,\sigma}(0) = \left(\tfrac{1}{2\pi\eta}\right)^n J\Sigma_0 J. \tag{293}$$

Since $\widehat{\rho}$ is a density matrix we have

$$M = -2\eta^{-1} J\Sigma J + iJ \geq 0; \tag{294}$$

the condition $M \geq 0$ being equivalent to $J^T M J \geq 0$ the inequality (288) follows. Let us finally show that the

conditions (289) and (288) indeed are equivalent. Let $\Sigma = S^T DS$ be a symplectic diagonalization of Σ (Lemma 24). Since $S^T JS = J$ condition (288) is equivalent to

$$D + \frac{i\eta}{2} J \geq 0 \ , \ D = \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda \end{pmatrix}. \tag{295}$$

The characteristic polynomial of $D + \frac{i\eta}{2} J$ is

$$\begin{aligned} P(\lambda) &= \begin{vmatrix} \Lambda - \lambda I_n & \frac{i\eta}{2} I_n \\ -\frac{i\eta}{2} I_n & \Lambda - \lambda I_n \end{vmatrix} \\ &= \det\left[(\Lambda - \lambda I_n)^2 - \tfrac{1}{4}\eta^2 I_n\right]; \end{aligned}$$

the matrix Λ being diagonal, the zeroes λ of $P(\lambda)$ are the solutions of the n equations $(\lambda_j - \lambda)^2 - \frac{1}{4}\eta^2 = 0$ that is $\lambda_j - \lambda = \pm\frac{1}{2}|\eta|$. Since $\lambda \geq 0$ we must have $\lambda_j \geq \frac{1}{2}|\eta|$ for all j hence (289). Property (ii) immediately follows from (289); it can also be proved directly: setting $\eta' = r\eta$ with $0 < r \leq 1$ we have

$$\Sigma + \frac{i\eta'}{2} J = (1 - r)\Sigma + r(\Sigma + \frac{i\eta}{2} J) \geq 0. \tag{296}$$

□

The relation of the result above with the uncertainty principle comes from the following observation (see Narcowich [77], de Gosson [44, Chapter 13], de Gosson and Luef [78]): the relation

$$\Sigma + \tfrac{1}{2} i\eta J \geq 0; \tag{297}$$

(which we will call the *strong uncertainty principle*) implies the Robertson–Schrödinger inequalities

$$(\Delta x_j)^2_{\widehat{\rho}} (\Delta p_j)^2_{\widehat{\rho}} \geq (\Delta(x_j, p_j)_{\widehat{\rho}})^2 + \tfrac{1}{4}\eta^2 \tag{298}$$

$(j = 1, \ldots, n)$ and $(\Delta x_j)^2_\psi (\Delta p_k)^2_\psi \geq 0$ if $j \neq k$. It is however essential to note that (297) and (298) are *not equivalent*. Here is a counterexample in the case $n = 2$. Consider the symmetric matrix

$$\Sigma = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{299}$$

Assume that Σ is a covariance matrix; we thus have $(\Delta x_1)^2 = (\Delta x_2)^2 = 1$ and $(\Delta p_1)^2 = (\Delta p_2)^2 = 1$, and also $\Delta(x_1, p_1) = \Delta(x_2, p_2) = 0$ so that the inequalities (298) are trivially satisfied for the choice $\eta = 1$ (they are in fact equalities). The matrix $\Sigma + iJ$ is nevertheless indefinite (its determinant is -1); Σ is not even invertible.

It is also essential to realize that condition (297) is *necessary* but *not sufficient* for a phase space function to be the Wigner function of a quantum state. Let us check this on the following example due Narcowich and O'Connell [74], further discussed in de Gosson and Luef [79]. Consider the function $f(x, p)$ defined for $n = 1$ by

$$f(x, p) = (1 - \tfrac{1}{2}ax^2 - \tfrac{1}{2}bx^2)e^{-(a^2x^4 - b^2p^4)} \tag{300}$$

where a and b are positive constants such that $ab \geq \frac{1}{4}\hbar^2$. Now define

$$\rho(x, p) = \frac{1}{2\pi} \iint e^{-i(xx' + pp')} f(x', p') dp' dx'; \tag{301}$$

since $f(x, p)$ is an even function $\rho(x, p)$ is real; in view of the Fourier inversion formula we have

$$f(x, p) = \iint e^{-i(xx' + pp')} \rho(x', p') dp' dx' \tag{302}$$

and hence

$$\iint \rho(x, p) dp dx = f(0, 0) = 1 \tag{303}$$

so that $\rho(x, p)$ is a candidate for being the Wigner function of some density matrix. Calculating the covariance matrix Σ associated with ρ, one finds after some tedious calculations [74, pp. 4–5] that $\Sigma + \frac{i\hbar}{2} J \geq 0$. However, ρ cannot be the Wigner function of a density matrix $\widehat{\rho}$ for if this were the case we would have $\langle p^4 \rangle_{\widehat{\rho}} \geq 0$; but by definition of ρ

$$\iint p^4 \rho(x, p) dp dx = \frac{\partial^4 f}{\partial x^4}(0, 0) = -24a^2 < 0. \tag{304}$$

6.4 Gaussian Bosonic states

We focus now on the case where all involved states are Gaussian: we say that a pure or mixed quantum state is Gaussian if its Wigner function is of the type

$$\rho(z) = (2\pi)^{-n} \sqrt{\det \Sigma^{-1}} e^{-\frac{1}{2}\Sigma^{-1}(z - z_0)\cdot(z - z_0)} \tag{305}$$

where Σ (the *covariance matrix*) is a positive definite $2n \times 2n$ matrix satisfying a certain condition that will be stated later (intuitively speaking Σ cannot be *too small* because then $\rho(z)$ would be too sharply peaked and thus violate the uncertainty principle of quantum mechanics). Gaussian states appear naturally in every quantum system which can be described or approximated by a quadratic Bosonic Hamiltonian (Wolf *et al.* [80]); because of their peculiarities they play an exceptionally important role in quantum mechanics and optics (see Barnett and Radmore [81]).

6.4.1 Definition and examples

We will define a *generalized Gaussian* as any complex function on $\mathbb{R}^n$ of the type

$$\psi_M^\hbar(x) = \left(\tfrac{1}{\pi\hbar}\right)^{n/4} (\det X)^{1/4} e^{-\frac{1}{2\hbar}M(x-x_0)^2} \tag{306}$$

where $M = X + iY$ is a complex symmetric $2n \times 2n$ invertible matrix; X and Y are real matrices such that $X = X^T > 0$ and $Y = Y^T$. The coefficient in front of the exponential is chosen so that $\psi_M^\hbar$ is normalized to unity: $\|\psi_M^\hbar\| = 1$.

Suppose that $X = I$ and $Y = 0$; then

$$\psi_M^\hbar(x) = \phi_0^\hbar(x) = \left(\tfrac{1}{\pi\hbar}\right)^{n/4} e^{-\frac{1}{2\hbar}|x|^2} \tag{307}$$

is the standard (or fiducial) coherent state

Let $\phi_M(x) = e^{-\frac{1}{2\hbar}Mx^2}$ where $M = X+iY$ is a symmetric complex $n \times n$ matrix such that $X = \operatorname{Re} M > 0$. The Fourier transform

$$F\phi_M(p) = \left(\tfrac{1}{2\pi\hbar}\right)^{n/2} \int e^{-\frac{i}{\hbar}px} \phi_M(x) d^n x \tag{308}$$

is given by

$$F\phi_M(x) = (\det M)^{-1/2} \phi_{M^{-1}}(x) \tag{309}$$

where $(\det M)^{-1/2}$ is given by the formula

$$(\det M)^{-1/2} = \lambda_1^{-1/2} \cdots \lambda_m^{-1/2} \tag{310}$$

the numbers $\lambda_1^{-1/2}, \ldots, \lambda_n^{-1/2}$ being the square roots with positive real parts of the eigenvalues $\lambda_1^{-1}, \ldots, \lambda_m^{-1}$ of M^{-1} (see e.g. Folland [66, Appendix A]). It follows that the Fourier transform of $\psi_M^\hbar$ is given by the formula

$$F\psi_M^\hbar(p) = \left(\tfrac{1}{\pi\hbar}\right)^{n/4} (\det X)^{1/4} (\det M)^{-1/2} \phi_{M^{-1}}(x). \tag{311}$$

6.4.2 The Wigner transform of $\psi_M^\hbar$

We are following here almost verbatim our discussion in [44, §11.2.1].

Theorem 27. *The Wigner transform $W\psi_M^\hbar$ is the phase space Gaussian*

$$W\psi_M^\hbar(z) = \left(\tfrac{1}{\pi\hbar}\right)^n e^{-\frac{1}{\hbar}Gz^2} \tag{312}$$

where G is the symplectic symmetric matrix

$$G = \begin{pmatrix} X + YX^{-1}Y & YX^{-1} \\ X^{-1}Y & X^{-1} \end{pmatrix}; \tag{313}$$

in fact $G = S^T S$ where

$$S = \begin{pmatrix} X^{1/2} & 0 \\ X^{-1/2}Y & X^{-1/2} \end{pmatrix} \tag{314}$$

is a symplectic matrix.

Proof. To simplify notation we set $C(X) = (\pi\hbar)^{n/4} (\det X)^{1/4}$. By definition of the Wigner transform we have

$$W\psi_M^\hbar(z) = \left(\tfrac{1}{2\pi\hbar}\right)^n C(X)^2 \int e^{-\frac{i}{\hbar}py} e^{-\frac{1}{2\hbar}F(x,y)} d^n y \tag{315}$$

where the phase F is defined by

$$\begin{aligned} F(x,y) &= (X+iY)(x+\tfrac{1}{2}y)^2 + (X-iY)(x-\tfrac{1}{2}y)^2 \\ &= 2Xx^2 + 2iYx\cdot y + \tfrac{1}{2}Xy^2 \end{aligned}$$

so we can rewrite (315) as

$$W\psi_M^\hbar(z) = \left(\tfrac{1}{2\pi\hbar}\right)^n e^{-\frac{1}{\hbar}Xx^2} C(X)^2 \int e^{-\frac{i}{\hbar}(p+Yx)y} e^{-\frac{1}{4\hbar}Xy^2} d^n y. \tag{316}$$

Using the Fourier transformation formula (309) above with x replaced by $p + Yx$ and M by $\frac{1}{2}X$ we get

$$\int e^{-\frac{i}{\hbar}(p+Yx)y} e^{-\frac{1}{4\hbar}Xy^2} d^n y = (2\pi\hbar)^{n/2} \left[\det(\tfrac{1}{2}X)\right]^{-1/2} C(X)^2 e^{-\frac{1}{\hbar}X^{-1}(p+Yx)^2}.$$

On the other hand we have

$$(2\pi\hbar)^{n/2} \left[\det(\tfrac{1}{2}X)\right]^{-1/2} C(X)^2 = \left(\tfrac{1}{\pi\hbar}\right)^n \tag{317}$$

and hence

$$W\psi_M^\hbar(z) = \left(\tfrac{1}{\pi\hbar}\right)^n e^{-\frac{1}{\hbar}Gz^2} \tag{318}$$

where

$$Gz^2 = (X + YX^{-1})x^2 + 2X^{-1}Yx\cdot p + X^{-1}p^2 \tag{319}$$

so that G is given by (313). One immediately verifies that $G = S^T S$ where S is given by (314) and that $S^T J S = J$ hence $S \in \mathrm{Sp}(n)$ as claimed. □

In particular, when $\psi_M^\hbar$ is the standard coherent state (307) we recover the well-known formula

$$W\phi_0^\hbar(z) = \left(\tfrac{1}{\pi\hbar}\right)^n e^{-\frac{1}{\hbar}|z|^2}. \tag{320}$$

6.4.3 A necessary and sufficient condition

We are going to discuss the following question: *For which values of η can the Gaussian function ρ be the η-Wigner function of a density operator?* Narcowich [82] was the first to address this question using techniques from harmonic analysis using the approach in Kastler's paper [67]; we give here a new and simpler proof using the multidimensional generalization of Hardy's uncertainty principle.

Let us begin with what is usually called in the literature *Hardy's uncertainty principle*. In what follows we denote by $F_\eta\psi$ the η-Fourier transform given, for $\psi \in L^2(\mathbb{R}^n)$, by

$$F_\eta\psi(p) = \left(\tfrac{1}{2\pi\eta}\right)^n \int e^{-\frac{i}{\eta}px}\psi(x)d^nx. \tag{321}$$

An old result (1933) due to Hardy [83] quantifies the *folk theorem* following which a function and its Fourier transform cannot be simultaneously arbitrarily sharply peaked. In fact Hardy proved that if $\psi \in L^2(\mathbb{R})$ satisfies

$$|\psi(x)| \leq Ce^{-\frac{1}{2\eta}ax^2} \text{ and } |F_\eta\psi(p)| \leq Ce^{-\frac{1}{2\eta}bx^2} \tag{322}$$

then we must have $ab \leq 1$. In particular, if $ab = 1$ then $\psi(x) = Ne^{-ax^2/2\eta}$ for some constant N (we are thus here in the presence of a particular quantum tomography result, which says that it suffices, in the Gaussian case, to know the position and momentum probabilities to determines the state). We will use the following generalization to the multidimensional case of Hardy's result:

Lemma 28 (Hardy). *Let A and B be two real positive definite matrices and $\psi \in L^2(\mathbb{R}^n)$, $\psi \neq 0$. Assume that*

$$|\psi(x)| \leq Ce^{-\frac{1}{2}Ax^2} \text{ and } |F_\eta\psi(p)| \leq Ce^{-\frac{1}{2}Bp^2} \tag{323}$$

for a constant $C > 0$. Then: (i) The eigenvalues λ_j, $j = 1, \ldots, n$, of the matrix AB are all $\leq 1/\eta^2$; (ii) If $\lambda_j = 1/\eta^2$ for all j, then $\psi(x) = ke^{-\frac{1}{2}Ax^2}$ for some complex constant k.

Proof. See de Gosson and Luef [78, 84], de Gosson [44]. □

We will also need the following positivity result:

Lemma 29. *If R is a symmetric positive semidefinite $2n \times 2n$ matrix, then*

$$P_{(N)} = \left(Rz_j \cdot z_k\right)_{1\leq j,k\leq N} \tag{324}$$

is a symmetric positive semidefinite $N \times N$ matrix for all $z_1, \ldots, z_N \in \mathbb{R}^{2n}$.

Proof. There exists a matrix L such that $R = L^*L$ (Cholesky decomposition). Denoting by $\langle z|z'\rangle = z \cdot \overline{z'}$ the inner product on $\mathbb{C}^{2n}$ we have, since the z_j are real vectors,

$$L^*z_j \cdot z_k = \langle L^*z_j|z_k\rangle = \langle z_j|Lz_k\rangle = z_j \cdot (Lz_k)^* \tag{325}$$

hence $Rz_j \cdot z_k = Lz_j \cdot (Lz_k)^*$. It follows that

$$\sum_{1\leq j,k\leq N} \lambda_j\lambda_k^* Rz_j \cdot z_k = \sum_{1\leq j\leq N} \lambda_j Lz_j \left(\sum_{1\leq j\leq N} \lambda_j Lz_j\right)^* \geq 0 \tag{326}$$

hence our claim. □

We now have the tools needed to give a complete characterization of Gaussian η-Wigner functions. Recall from Theorem 26 that a necessary condition for a matrix Σ to be the covariance matrix of a quantum state is that it satisfies the condition $\Sigma + \frac{i\eta}{2}J \geq 0$. It turns out that in the Gaussian case this condition is also *sufficient*:

Theorem 30. *The Gaussian function*

$$\rho(z) = (2\pi)^{-n}\sqrt{\det\Sigma^{-1}}e^{-\frac{1}{2}\Sigma^{-1}z^2} \tag{327}$$

is the η-Wigner transform of a positive trace class operator if and only if it satisfies

$$|\eta| \leq 2\lambda_{\min} \tag{328}$$

where $\lambda_{\min}$ is the smallest symplectic eigenvalue of Σ; equivalently

$$\Sigma + \frac{i\eta}{2}J \geq 0. \tag{329}$$

Proof. Let us give a direct proof of the necessity of condition (328) for the Gaussian (327) to be the η-Wigner transform of a positive trace class operator. Let $\widehat{\rho} = (2\pi\eta)^n \operatorname{Op}_\eta^{\mathrm{W}}(\rho)$ and set $a(z) = (2\pi\eta)^n\rho(z)$. Let $\widehat{S} \in \mathrm{Mp}(n)$; the operator $\widehat{\rho}$ is of trace class if and only if $\widehat{S}\widehat{\rho}\widehat{S}^{-1}$ is, in which case $\mathrm{Tr}(\widehat{\rho}) = \mathrm{Tr}(\widehat{S}\widehat{\rho}\widehat{S}^{-1})$. Choose $\widehat{S}$ with projection $S \in \mathrm{Sp}(n)$ such that $\Sigma = S^TDS$ is a symplectic diagonalization of Σ. This choice reduces the proof to the case $\Sigma = D$, that is to

$$\rho(z) = (2\pi)^{-n}(\det\Lambda^{-1})e^{-\frac{1}{2}(\Lambda^{-1}x^2+\Lambda^{-1}p^2)}. \tag{330}$$

Suppose now that $\widehat{\rho}$ is of trace class; then there exist functions $\psi_j \in L^2(\mathbb{R}^n)$ $(1 \leq j \leq n)$ such that

$$\rho(z) = \sum_j \alpha_j W_\eta\psi_j(z) \tag{331}$$

where the $\alpha_j > 0$ sum up to one. Integrating with respect to the p and x variables, respectively, the marginal conditions satisfied by the η-Wigner transform and formula (330) imply that we have

$$\sum_j \alpha_j|\psi_j(x)|^2 = (2\pi)^{-n/2}(\det\Lambda)^{1/2}e^{-\frac{1}{2}\Lambda^{-1}x^2}$$
$$\sum_j \alpha_j|F_\eta\psi_j(p)|^2 = (2\pi)^{-n/2}(\det\Lambda)^{1/2}e^{-\frac{1}{2}\Lambda^{-1}p^2}.$$

In particular, since $\alpha_j \geq 0$ for every $j = 1, 2, \ldots, n$,

$$|\psi_j(x)| \leq C_je^{-\frac{1}{4}\Lambda^{-1}x^2}, \ |F_\eta\psi_j(p)| \leq C_je^{-\frac{1}{4}\Lambda^{-1}p^2} \tag{332}$$

here $C_j = (2\pi)^{-n/4}(\det\Lambda)^{1/4}/\alpha_j^{1/2}$. Applying Hardy's Lemma 28 with $A = B = \frac{1}{2}\eta\Lambda^{-1}$ we must have $|\eta| \leq 2\lambda_j$ for all $j = 1, \ldots, n$ which is condition (328); this

establishes the sufficiency statement. Let us finally show that, conversely, the condition (329) is sufficient. It is again no restriction to assume that Σ is the diagonal matrix $D = \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda \end{pmatrix}$; the symplectic Fourier transform of ρ is easily calculated and one finds that $\rho_\diamond(z) = e^{-\frac{1}{4}Dz^2}$. Let $\Lambda_{(N)}$ be the $N \times N$ matrix with entries

$$\Lambda_{jk} = e^{-\frac{i\eta}{2}\sigma(z_j,z_k)}\rho_\diamond(z_j - z_k); \tag{333}$$

a simple algebraic calculation shows that we have

$$\Lambda_{jk} = e^{-\frac{1}{4}Dz_j^2}e^{\frac{1}{2}(D+i\eta J)z_j\cdot z_k}e^{-\frac{1}{4}Dz_k^2} \tag{334}$$

and hence

$$\Lambda_{(N)} = \Delta_{(N)}\Gamma_{(N)}\Delta_{(N)}^* \tag{335}$$

where $\Delta_{(N)} = \mathrm{diag}(e^{-\frac{1}{4}Dz_1^2},\dots,e^{-\frac{1}{4}Dz_N^2})$ and $\Gamma_{(N)} = (\Gamma_{jk})_{1\leq j,k\leq N}$ with $\Gamma_{jk} = e^{\frac{1}{2}(D+i\eta J)z_j\cdot z_k}$. The matrix $\Lambda_{(N)}$ is thus positive semidefinite if and only if $\Gamma_{(N)}$ is, but this is the case in view of Lemma 29. □

Setting $2\lambda_{\min} = \hbar$ and writing Σ in the block-matrix form $\begin{pmatrix} \Sigma_{xx} & \Sigma_{xp} \\ \Sigma_{px} & \Sigma_{pp} \end{pmatrix}$ where $\Sigma_{xx} = (\Delta(x_j,x_k))_{1\leq j,k\leq n}$, $\Sigma_{xp} = (\Delta(x_j,p_k))_{1\leq j,k\leq n}$ and so on, one shows [78] that (329) implies the generalized uncertainty relations (the *Robertson–Schrödinger inequalities*; see de Gosson and Luef [78] for a detailed discussion of these inequalities)

$$\Delta x_j^2\Delta p_j^2 \geq \Delta(x_j,p_k)^2 + \tfrac{1}{4}\hbar^2 \tag{336}$$

where, for $\leq j \leq n$, the $\Delta x_j^2 = \Delta(x_j,x_j)$, $\Delta p_j^2 = \Delta(p_j,p_j)$ are viewed as variances and the $\Delta(x_j,p_k)$ as covariances. We have given a detailed discussion of the Robertson–Schrödinger inequalities in de Gosson and Luef [78] from the symplectic point of view.

7 Some Speculations

7.1 The fine structure constant

Dirac [85] already speculated in 1937 that physical constants such as the gravitational constant or the fine structure constant might be subject to change over time. This question has since been a very active area of research (see the recent reviews [86, 87]). Some scientists have suggested that the fine structure constant, $\alpha \approx 1/137$, might not be constant, but could vary over time and space. This dimensionless constant, introduced by Sommerfeld in 1916, measures the strength of interactions between light and matter, or equivalently, how strong electrical and magnetic forces are. It can be expressed as a combination of three constants: the electron charge, the speed of light, and Planck's constant h:

$$\alpha = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{\hbar c}. \tag{337}$$

The quest for testing the non-constancy of α is ongoing. The Oklo natural nuclear reactor is known to give limits on the variation of the fine structure constant over the period since the reactor was running ($\approx$ 1.8 billion years). In 1999, a team of astronomers using a telescope in Hawaii reported that measurements of light absorbed by very distant galaxy-like objects in space called quasars—which are so far away that we see them today as they looked billions of years ago—suggest that the value of the fine structure constant was once slightly different from what it is today. Experiments can in principle only put an upper bound on the relative change per year. For the fine structure constant, this upper bound is comparatively low, at roughly 10^{-17} per year. That claim was controversial, and still unproven. But if true, it must mean that at least one of the three fundamental constants that constitute α must vary. The possibility that some constants of Nature could vary in space-time has remained a subject of fascination which has motivated numerous theoretical and experimental researches [88, 89].

7.2 Planck's constant

Kentosh and Mohageg focused on h, and specifically on whether h depends on where (not when) you measure it. If h changes from place to place, so do the frequencies, and thus the *ticking rate*, of atomic clocks. And any dependence of h on location would translate as a tiny timing discrepancy between different GPS clocks. The physicist Freeman Dyson has suggested (private communication) that the increasing precision of measurements of time could lead to non-ambiguous results. Mohageg and his student Kentosh [90,91] have tested the constancy of h using the freely available data from GPS. Kentosh and Mohageg were actually motivated by the fact that h also appears in the fine structure constant, whose possible variation is a very active area of research in experimental physics. After careful analysis of the data from seven highly stable GPS satellites, Kentosh and Mohageg concluded that h is identical at different locations to an accuracy of seven parts in a thousand. In other words, if h were a one-metre measuring stick, then two sticks in different places anywhere in the world would not differ by more than seven millimeters.

At least as interesting is the possible time-variation of Planck's constant (see Mangano *et al.* [92]). This deserves to be explored because if true it could shed some light on the Early Universe, just after the Big Bang. In fact, if the

fine structure constant has been increasing since the Big Bang, this could perhaps be due to a decrease of Planck's constant. If such a variation could be experimentally detected, then it would mean, following our discussion of the quantum Bochner theorem, that the early Universe was much more *quantum* than it is now; this would of course have major implications in terms of entanglement.

7.3 Units

Testing the constancy of a physical parameter means going to extraordinary lengths in terms of precision measurements, and is intimately related to choices of *unit systems*. The physicist Michael Duff [93] remarked in 2002 (also see Duff [86, 94]) that all the fundamental physical dimensions could be expressed using only one: mass. Duff first noticed the obvious, namely that lengths can be expressed as times using c, the velocity of light, as a conversion factor. One can therefore take $c = 1$, and measure lengths in seconds. The second step was to use the relation $E = h\nu$ which relates energy to a frequency, that is to the inverse of a time. We can thus measure a time using the inverse of energy. But energy is equivalent to mass as shown by Einstein, so that time can be measured by the inverse of mass. Thus, setting $c = h = 1$ we have reduced all the fundamental dimensions to one: mass. A further step consists in choosing a reference mass such that the gravitational constant (first measured by Cavendish in 1798) is equal to one: $G = 1$. Summarizing, we have obtained a theoretical system of units in which $c = h = G = 1$. Now, a very important physical parameter is, without doubt, the fine structure constant $\alpha = e^2/2\varepsilon_0 hc$ (ε_0 the dielectric constant); it is a dimensionless number whose approximate value is $1/137$. There are other ways to define irreducible unit systems. Already Stoney, noting that electric charge is quantized, derived units of length, time, and mass in 1881 by normalizing G, c, and e to unity; Planck suggested in 1898–1899 that it would suffice to use G, c, and h to define length, mass, and time units. His proposal led to what are called today Planck's length $\ell_P = \sqrt{Gh/c^3}$ and Planck mass and time $M_P = \sqrt{hc/G}$ and $T_P = \sqrt{Gh/c^5}$.

Acknowledgements

This work has been financed by the grant P27773 N23 of the Austrian Research Foundation FWF (Fonds zur Förderung der wissenschaftlichen Forschung). It is my pleasure to thank Glen Dennis for a careful reading of the manuscript.

References

[1] Vogel K, Risken H. Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Physical Review A* 1989; **40**(5): 2847–2849. doi:10.1103/PhysRevA.40.2847

[2] de Gosson MA. *Born–Jordan Quantization: Theory and Applications*. Fundamental Theories of Physics, vol. 182, Cham, Switzerland: Springer, 2016. doi:10.1007/978-3-319-27902-2

[3] de Gosson MA, Luef F. Preferred quantization rules: Born–Jordan versus Weyl. The pseudo-differential point of view. *Journal of Pseudo-Differential Operators and Applications* 2011; **2**(1): 115–139. arXiv:1102.5132, doi:10.1007/s11868-011-0025-6

[4] Peres A. *Quantum Theory: Concepts and Methods*. Fundamental Theories of Physics, vol. 72, New York: Kluwer, 2002. doi:10.1007/0-306-47120-5

[5] von Neumann J. Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1927: 245–272. http://eudml.org/doc/59230

[6] Fano U. Description of states in quantum mechanics by density matrix and operator techniques. *Reviews of Modern Physics* 1957; **29**(1): 74–93. doi:10.1103/RevModPhys.29.74

[7] Peres A. Separability criterion for density matrices. *Physical Review Letters* 1996; **77**(8): 1413–1415. arXiv:quant-ph/9604005, doi:10.1103/PhysRevLett.77.1413

[8] Horodecki M, Horodecki P, Horodecki R. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A* 1996; **223**(1): 1–8. arXiv:quant-ph/9605038, doi:10.1016/S0375-9601(96)00706-2

[9] Horodecki M, Horodecki P, Horodecki R. Separability of n-particle mixed states: necessary and sufficient conditions in terms of linear maps. *Physics Letters A* 2001; **283**(1–2): 1–7. arXiv:quant-ph/0006071, doi:10.1016/S0375-9601(01)00142-6

[10] Leinaas JM, Myrheim J, Ovrum E. Geometrical aspects of entanglement. *Physical Review A* 2006; **74**(1): 012313. arXiv:quant-ph/0605079, doi:10.1103/PhysRevA.74.012313

[11] Rieffel MA. Quantization and C^*-algebras. *Contemporary Mathematics* 1994; **167**: 67–97. doi:10.1090/conm/167

[12] Dixmier J. *Les C-algèbres et leurs représentations*. Paris: Gauthier-Villars, 1969.

[13] Gelfand I, Neumark M. On the imbedding of normed rings into the ring of operators in Hilbert space. *Matematicheskii Sbornik* 1943; **12**(2): 197–217. http://mi.mathnet.ru/eng/msb6155

[14] Takhtajan LA. *Quantum Mechanics for Mathematicians*. Graduate Studies in Mathematics, vol. 95, Providence, Rhode Island: American Mathematical Society, 2008. doi:10.1090/gsm/095

[15] Daubechies I. Continuity statements and counterintuitive examples in connection with Weyl quantization. *Journal of Mathematical Physics* 1983; **24**(6): 1453–1461. doi:10.1063/1.525882

[16] Albeverio SA, Høegh-Krohn RJ, Mazzucchi S. *Mathematical Theory of Feynman Path Integrals: An Introduction*. Lecture Notes in Mathematics, vol. 523, Berlin: Springer, 2008. doi:10.1007/978-3-540-76956-9

[17] Nicola F. Convergence in L^p for Feynman path integrals. *Advances in Mathematics* 2016; **294**: 384–409. arXiv:1503.05863, doi:10.1016/j.aim.2016.03.003

[18] Brislawn C. Kernels of trace class operators. *Proceedings of the American Mathematical Society* 1988; **104**(4): 1181–1190.

[19] Simon B. *Trace Ideals and Their Applications*. Mathematical Surveys and Monographs, vol. 120, Providence, Rhode Island: American Mathematical Society, 2005. doi:10.1090/surv/120

[20] Shubin MA. *Pseudodifferential Operators and Spectral Theory*. Andersson SI (translator), Berlin: Springer, 2001. doi:10.1007/978-3-642-56579-3

[21] Wang Y, Xu C. Density matrix estimation in quantum homodyne tomography. *Statistica Sinica* 2015; **25**(3): 953–973. doi:10.5705/ss.2013.300

[22] D'Ariano GM. Universal quantum observables. *Physics Letters A* 2002; **300**(1): 1–6. doi:10.1016/S0375-9601(02)00752-1

[23] D'Ariano GM, Macchiavello C, Paris MGA. Detection of the density matrix through optical homodyne tomography without filtered back projection. *Physical Review A* 1994; **50**(5): 4298–4302. doi:10.1103/PhysRevA.50.4298

[24] Leonhardt U, Paul H. Realistic optical homodyne measurements and quasiprobability distributions. *Physical Review A* 1993; **48**(6): 4598–4604. doi:10.1103/PhysRevA.48.4598

[25] Bužek V, Adam G, Drobný G. Reconstruction of Wigner functions on different observation levels. *Annals of Physics* 1996; **245**(1): 37–97. doi:10.1006/aphy.1996.0003

[26] Thekkadath GS, Giner L, Chalich Y, Horton MJ, Banker J, Lundeen JS. Direct measurement of the density matrix of a quantum system. *Physical Review Letters* 2016; **117**(12): 120401. arXiv:1604.07917, doi:10.1103/PhysRevLett.117.120401

[27] Lvovsky AI, Raymer MG. Continuous-variable optical quantum-state tomography. *Reviews of Modern Physics* 2009; **81**(1): 299–332. arXiv:quant-ph/0511044, doi:10.1103/RevModPhys.81.299

[28] Paris M, Řeháček J. *Quantum State Estimation*. Lecture Notes in Physics, vol. 649, Berlin: Springer, 2004. doi:10.1007/b98673

[29] Mancini S, Man'ko VI, Tombesi P. Symplectic tomography as classical approach to quantum systems. *Physics Letters A* 1996; **213**(1): 1–6. arXiv:quant-ph/9603002, doi:10.1016/0375-9601(96)00107-7

[30] Man'ko O, Man'ko VI. Quantum states in probability representation and tomography. *Journal of Russian Laser Research* 1997; **18**(5): 407–444. doi:10.1007/bf02559668

[31] Ibort A, Man'ko VI, Marmo G, Simoni A, Ventriglia F. An introduction to the tomographic picture of quantum mechanics. *Physica Scripta* 2009; **79**(6): 065013. arXiv:0904.4439, doi:10.1088/0031-8949/79/06/065013

[32] Pauli W. *General Principles of Quantum Mechanics*. Achuthan P, Venkatesan K (translators), Berlin: Springer, 1980. doi:10.1007/978-3-642-61840-6

[33] Esposito G, Marmo G, Miele G, Sudarshan G. *Advanced Concepts in Quantum Mechanics*. Cambridge: Cambridge University Press, 2014. doi:10.1017/cbo9781139875950

[34] Berry MV. Quantal phase factors accompanying adiabatic changes. *Proceedings of the Royal Society of London, Series A. Mathematical and Physical Sciences* 1984; **392**(1802): 45–57. doi:10.1098/rspa.1984.0023

[35] Aharonov Y, Bohm D. Significance of electromagnetic potentials in the quantum theory. *Physical Review* 1959; **115**(3): 485–491. doi:10.1103/PhysRev.115.485

[36] Gabor D. Holography, 1948–1971. *Nobel Lecture*, December 11, 1971, https://www.nobelprize.org/prizes/physics/1971/gabor/lecture/

[37] Pancharatnam S. Generalized theory of interference, and its applications. *Proceedings of the Indian Academy of Sciences, Section A* 1956; **44**(5): 247–262. doi:10.1007/bf03046050

[38] Sjöqvist E, Pati AK, Ekert A, Anandan JS, Ericsson M, Oi DKL, Vedral V. Geometric phases for mixed states in interferometry. *Physical Review Letters* 2000; **85**(14): 2845–2849. arXiv:quant-ph/0005072, doi:10.1103/PhysRevLett.85.2845

[39] Nicacio F, Valdés-Hernández A, Majtey AP, Toscano F. Unified framework to determine Gaussian states in continuous-variable systems. *Physical Review A* 2017; **96**(4): 042341. doi:10.1103/PhysRevA.96.042341

[40] de Gosson MA, Nicacio F. Relative phase shifts for metaplectic isotopies acting on mixed Gaussian states. *Journal of Mathematical Physics* 2018; **59**(5): 052106. arXiv:1802.07499, doi:10.1063/1.5026586

[41] de Gosson MA. On the Weyl representation of metaplectic operators. *Letters in Mathematical Physics* 2005; **72**(2): 129–142. arXiv:math/0503708, doi:10.1007/s11005-005-4391-y

[42] de Gosson MA. *Symplectic Geometry and Quantum Mechanics*. Operator Theory: Advances and Applications, vol. 166, Basel: Birkhäuser, 2006. doi:10.1007/3-7643-7575-2

[43] de Gosson MA. Metaplectic representation, Conley–Zehnder index, and Weyl calculus on phase space. *Reviews in Mathematical Physics* 2007; **19**(10): 1149–1188. doi:10.1142/s0129055x07003152

[44] de Gosson MA. *Symplectic Methods in Harmonic Analysis and in Mathematical Physics*. Pseudo-Differential Operators, vol. 7, Basel: Birkhäuser, 2011. doi:10.1007/978-3-7643-9992-4

[45] Blanchard P, Brüning E. *Mathematical Methods in Physics: Distributions, Hilbert Space Operators, Variational Methods, and Applications in Quantum Physics*. Progress in Mathematical Physics, vol. 69, Basel: Birkhäuser, 2015. doi:10.1007/978-3-319-14045-2

[46] de Gosson MA. *The Wigner Transform*. Advanced Textbooks in Mathematics, Singapore: World Scientific, 2017. doi:10.1142/q0089

[47] Littlejohn RG. The semiclassical evolution of wave packets. *Physics Reports* 1986; **138**(4–5): 193–291. doi:10.1016/0370-1573(86)90103-1

[48] Wigner E. On the quantum correction for thermodynamic equilibrium. *Physical Review* 1932; **40**(5): 749–759. doi:10.1103/PhysRev.40.749

[49] Hillery M, O'Connell RF, Scully MO, Wigner EP. Distribution functions in physics: fundamentals. *Physics Reports* 1984; **106**(3): 121–167. doi:10.1016/0370-1573(84)90160-1

[50] Hudson RL. When is the Wigner quasi-probability density non-negative? *Reports on Mathematical Physics* 1974; **6**(2): 249–252. doi:10.1016/0034-4877(74)90007-x

[51] Janssen AJEM. A note on Hudson's theorem about functions with nonnegative Wigner distributions. *SIAM Journal on Mathematical Analysis* 1984; **15**(1): 170–176. doi:10.1137/0515014

[52] de Gosson C, de Gosson MA. The phase space formulation of time-symmetric quantum mechanics. *Quanta* 2015; **4**: 27–34. doi:10.12743/quanta.v4i1.46

[53] de Gosson C, de Gosson MA. Weak values, the reconstruction problem, and the uncertainty principle. *Journal of Physics: Conference Series* 2016; **701**(1): 012011. doi:10.1088/1742-6596/701/1/012011

[54] Grossmann A. Parity operator and quantization of δ-functions. *Communications in Mathematical Physics* 1976; **48**(3): 191–194. Euclid:1103899886, doi:10.1007/bf01617867

[55] Royer A. Wigner function as the expectation value of a parity operator. *Physical Review A* 1977; **15**(2): 449–450. doi:10.1103/PhysRevA.15.449

[56] Potoček V, Barnett SM. On the exponential form of the displacement operator for different systems. *Physica Scripta* 2015; **90**(6): 065208. doi:10.1088/0031-8949/90/6/065208

[57] Weyl H. Quantenmechanik und Gruppentheorie. *Zeitschrift für Physik* 1927; **46**(1): 1–46. doi:10.1007/bf02055756

[58] Du J, Wong MW. A trace formula for Weyl transforms. *Approximation Theory and its Applications* 2000; **16**(1): 41–45. doi:10.1007/bf02845227

[59] Gracia-Bondía JM, Várilly JC. Algebras of distributions suitable for phase-space quantum mechanics. I. *Journal of Mathematical Physics* 1988; **29**(4): 869–879. doi:10.1063/1.528200

[60] Gracia-Bondía JM, C. Várilly J. Nonnegative mixed states in Weyl–Wigner–Moyal theory. *Physics Letters A* 1988; **128**(1): 20–24. doi:10.1016/0375-9601(88)91035-3

[61] de Gosson MA. *The Principles of Newtonian and Quantum Mechanics: The Need for Planck's Constant, h*, 2nd edition. Singapore: World Scientific, 2017. doi:10.1142/10307

[62] Arvind, Dutta B, Mukunda N, Simon R. The real symplectic groups in quantum mechanics and optics. *Pramana* 1995; **45**(6): 471–497. arXiv:quant-ph/9509002, doi:10.1007/bf02848172

[63] García-Bullé M, Lassner W, Wolf KB. The metaplectic group within the Heisenberg–Weyl ring. *Journal of Mathematical Physics* 1986; **27**(1): 29–36. doi:10.1063/1.527333

[64] Leray J. *Lagrangian Analysis and Quantum Mechanics: A Mathematical Structure Related to Asymptotic Expansions and the Maslov Index*. Schroeder C (translator), Cambridge, Massachusetts: MIT Press, 1982.

[65] de Gosson MA. Maslov indices on the metaplectic group $Mp(n)$. *Annales de l'institut Fourier* 1990; **40**(3): 537–555. doi:10.5802/aif.1223

[66] Folland GB. *Harmonic Analysis in Phase Space*. The Annals of Mathematics Studies, vol. 122, Princeton: Princeton University Press, 1989.

[67] Kastler D. The C^*-algebras of a free boson field. *Communications in Mathematical Physics* 1965; **1**(1): 14–48. Euclid:1103757613, doi:10.1007/bf01649588

[68] Loupias G, Miracle-Sole S. C^*-algèbres des systèmes canoniques. I. *Communications in Mathematical Physics* 1966; **2**(1): 31–48. doi:10.1007/bf01773339

[69] Loupias G, Miracle-Sole S. C^*-algèbres des systèmes canoniques. II. *Annales de l'Institut Henri Poincaré, Physique théorique* 1967; **6**(1): 39–58.

[70] Parthasarathy KR. *An Introduction to Quantum Stochastic Calculus*. Basel: Birkhäuser, 1992. doi:10.1007/978-3-0348-0566-7

[71] Parthasarathy KR. What is a Gaussian state? *Communications on Stochastic Analysis* 2010; **4**(2): 143–160. doi:10.31390/cosa.4.2.02

[72] Parthasarathy KR, Schmidt K. *Positive Definite Kernels, Continuous Tensor Products, and Central Limit Theorems of Probability Theory*. Lecture Notes in Mathematics, vol. 272, Berlin: Springer, 1972. doi:10.1007/bfb0058340

[73] Narcowich FJ. Distributions of $\hbar$-positive type and applications. *Journal of Mathematical Physics* 1989; **30**(11): 2565–2573. doi:10.1063/1.528537

[74] Narcowich FJ, O'Connell RF. Necessary and sufficient conditions for a phase-space function to be a Wigner distribution. *Physical Review A* 1986; **34**(1): 1–6. doi:10.1103/PhysRevA.34.1

[75] Werner R. Quantum harmonic analysis on phase space. *Journal of Mathematical Physics* 1984; **25**(5): 1404–1411. doi:10.1063/1.526310

[76] Bapat RB, Raghavan TES. *Nonnegative Matrices and Applications*. Encyclopedia of Mathematics and its Applications, vol. 64, Cambridge: Cambridge University Press, 1997. doi:10.1017/cbo9780511529979

[77] Narcowich FJ. Geometry and uncertainty. *Journal of Mathematical Physics* 1990; **31**(2): 354–364. doi:10.1063/1.528922

[78] de Gosson MA, Luef F. Symplectic capacities and the geometry of uncertainty: the irruption of symplectic topology in classical and quantum mechanics. *Physics Reports* 2009; **484**(5): 131–179. doi:10.1016/j.physrep.2009.08.001

[79] de Gosson MA, Luef F. Remarks on the fact that the uncertainty principle does not determine the quantum state. *Physics Letters A* 2007; **364**(6): 453–457. arXiv:quant-ph/0703061, doi:10.1016/j.physleta.2006.12.024

[80] Wolf MM, Giedke G, Cirac JI. Extremality of Gaussian quantum states. *Physical Review Letters* 2006; **96**(8): 080502. arXiv:quant-ph/0509154, doi:10.1103/PhysRevLett.96.080502

[81] Barnett S, Radmore P. *Methods in Theoretical Quantum Optics*. Oxford Series in Optical and Imaging Sciences, vol. 15, Oxford: Oxford University Press, 1997.

[82] Narcowich FJ. Conditions for the convolution of two Wigner distributions to be itself a Wigner distribution. *Journal of Mathematical Physics* 1988; **29**(9): 2036–2041. doi:10.1063/1.527861

[83] Hardy GH. A theorem concerning Fourier transforms. *Journal of the London Mathematical Society* 1933; **8**(3): 227–231. doi:10.1112/jlms/s1-8.3.227

[84] de Gosson MA, Luef F. Quantum states and Hardy's formulation of the uncertainty principle: a symplectic approach. *Letters in Mathematical Physics* 2007; **80**(1): 69–82. arXiv:quant-ph/0703063, doi:10.1007/s11005-007-0150-6

[85] Dirac PAM. A new basis for cosmology. *Proceedings of the Royal Society of London, Series A. Mathematical and Physical Sciences* 1938; **165**(921): 199–208. doi:10.1098/rspa.1938.0053

[86] Duff MJ. How fundamental are fundamental constants? *Contemporary Physics* 2015; **56**(1): 35–47. arXiv:1412.2040

[87] Uzan J-P. Varying constants, gravitation and cosmology. *Living Reviews in Relativity* 2011; **14**(1): 2. doi:10.12942/lrr-2011-2

[88] Damour T, Dyson F. The Oklo bound on the time variation of the fine-structure constant revisited. *Nuclear Physics B* 1996; **480**(1): 37–54. doi:10.1016/S0550-3213(96)00467-1

[89] Dyson FJ. The fundamental constants and their time variation. In: *Aspects of Quantum Theory*. Salam A, Wigner EP (editors), Cambridge: Cambridge University Press, 1972, pp. 213–236. https://archive.org/details/AspectsOfQuantumTheory

[90] Kentosh J, Mohageg M. Global positioning system test of the local position invariance of Planck's constant. *Physical Review Letters* 2012; **108**(11): 110801. arXiv:1203.0102, doi:10.1103/PhysRevLett.108.110801

[91] Kentosh J, Mohageg M. Testing the local position invariance of Planck's constant in general relativity. *Physics Essays* 2015; **28**(2): 286–289. doi:10.4006/0836-1398-28.2.286

[92] Mangano G, Lizzi F, Porzio A. Inconstant Planck's constant. *International Journal of Modern Physics A* 2015; **30**(34): 1550209. arXiv:1509.02107, doi:10.1142/S0217751X15502097

[93] Duff MJ, Okun LB, Veneziano G. Trialogue on the number of fundamental constants. *Journal of High Energy Physics* 2002; **2002**(3): 23. arXiv:physics/0110060, doi:10.1088/1126-6708/2002/03/023

[94] Duff MJ. Comment on time-variation of fundamental constants. 2002: arXiv:hep-th/0208093