Symplectic Capacities and the Geometry of Uncertainty: the Irruption of Symplectic Topology in Classical and Quantum Mechanics

Maurice de Gosson[1]
Universität Wien, NuHAG
Fakultät für Mathematik
A-1090 Wien

Franz Luef[2]
University of California
Department of Mathematics
Berkeley CA 94720-3840

August 2, 2009

[1]Financed by the Austrian Research Agency FWF (Projektnummer P20442-N13).
[2]Financed by the Marie Curie Outgoing Fellowship PIOF 220464.

Contents

Abstract

This paper aims at being an analysis of various uncertainty principles from a topological point of view where the notion of symplectic capacity plays a key role. The existence of symplectic capacities follows from a deep theorem of symplectic topology, Gromov's non-squeezing theorem, which was discovered in the mid 1980's, and who has led to numerous developments whose applications to Physics are not fully understood or exploited at the time of writing. We will show that the notion of symplectic non-squeezing contains, as a watermark, not only the Robertson–Schrödinger uncertainty relations (and a classical version thereof), but also Hardy's uncertainty principle for a function and its Fourier transform. This observation will allow us to formulate the characterization of positivity for density matrices in a topological way. We also address some open questions and conjectures, whose solution cannot be given at present time due to the lack of a sufficiently developed mathematical theory.

Keywords: symplectic non-squeezing theorem, uncertainty principles, density matrix, Wigner distribution, quantum phase space
AMS subject classification: 53D22, 53D05, 81S10, 81S30, 82C10
PACS: 03.65.Ta, 05.30.Ch, 03.65.Sq, 02.30.Ik, 02.20.Qs

1 Introduction

1.1 Declaration of intents

Our everyday world is ruled by Euclidean geometry (and by its extension to curved spaces, Riemannian geometry); we can measure distances in it, and hence velocities when time is taken into account. Far away from our daily experience, and much more subtle, is the mechanical phase space world, in which all the phenomena are expressed in terms of simultaneous measure of positions and momenta. A thorough understanding of *this* world requires the recourse to a somewhat counter-intuitive geometry, the symplectic geometry of Hamiltonian mechanics. Symplectic geometry is in fact highly counter-intuitive; the notion of length does not make sense there, while the notion of area does. This "areal" nature of symplectic geometry, which was not realized until very recently, has led to unexpected mathematical developments, starting in the mid 1980's with Gromov's discovery of a "non-squeezing" phenomenon which is reminiscent of the quantum uncertainty principle–but in a totally classical setting! Gromov's discovery has been followed by a constellation of related results, which have considerably stimulated and revigorated symplectic topology, which is the study of global topological properties invariant under the action of symplectic mappings (or "canonical transformations" as they are called in physics). Unfortunately these mathematical developments have taken place almost unnoticed from the general physics community. One noticeable exception is however Scheeres and his collaborators [80, 48, 68] who apply properties related to the non-squeezing theorem (the existence of a distinguished symplectic capacity) to the study of uncertainty analysis of Hamiltonian systems. They study the implications of the non-squeezing theorem to uncertainty distributions that are used to describe particle trajectories in space, with specific applications to spacecraft [80] and orbit debris predictions [68]. Additionally, the theorem is used to develop a constraint on covariance matrices for linear Hamiltonian systems of the form used in spacecraft navigation and orbit determination. We also notice that there have been some review papers on the topic of non-squeezing theorems and their applications: Stewart's 1987 review article [91] in *Nature* and Reich's 2009 review of de Gosson [35] in the *New Scientist*.

The aim of this paper is to introduce the physics community to some of these new ideas and concepts, and to show how they can be used with profit as well in quantum as in classical mechanics. We will try to convey to the reader our belief that these new concepts from symplectic topology might really be the proper setup for a better understanding of the twilight zone between "quantum" and "classical" properties. Indeed, the borderline between true quantum effects and classical phenomena is a murky area, which is far from being fully understood. For instance, there has been during the last years an increasing interest in precise forms of both classical and quantum uncertainty principles. This interest has been triggered not only by progresses in quantum optics, but also in quantum information science, and in time-frequency and signal analysis. We will see that symplectic topology allows us to state and prove a classical multi-dimensional uncertainty principle, formally similar to the quantum uncertainty relations. In fact, one recurring theme of the present paper is that the "uncertainty principle" in its strong form due to Robertson and Schrödinger can be expressed in terms of a class of symplectic invariants whose definition is made possible by Gromov's non-squeezing theorem. These invariants are the *symplectic capacities* of subsets of phase space. Symplectic capacities can be viewed as a measure of uncertainty, not related to volume, but rather to area, to which it reduces in some cases when the phase space is a plane; they are moreover intensive quantities, that is, they do not depend on the number of degrees of freedom. (The possibility of using symplectic capacities for multi-dimensional Hamiltonian systems was already noted by the first author in [23, 24, 35].)

Our applications of symplectic topology are not limited to the study of uncertainty principles. We will see that a careful use of the notion of symplectic capacity allows to recover the semiclassical ground energy levels of all Liouville integrable Hamiltonian systems; to do this we actually have to use the full power of symplectic topology, namely the invariance of symplectic capacities under the action of nonlinear canonical transformations (in the present case, the passage to action-angle variables). We also use define a notion of "quantum blob" which is a symplectically invariant substitute to the cubic quantum cells used in statistical quantum mechanics; the set of all such "blobs" form a what we call a "quantum phase space" whose elements are in one-to-one correspondence with the squeezed states familiar from quantum optics.

1.2 Structure of the paper

Let us outline the main parts of this work:

- We begin by collecting, in Section 2, the main results about the symplectic group and hamiltonian mechanics we will use throughout the paper, with a special emphasis on Williamson's symplectic diagonalization theorem, which plays a key role in practical calculations. We also briefly review the theory of the metaplectic representation which is so important when one wants to use a symplectically covariant operator theory (Weyl calculus);

- In Section 3 we review the concepts from symplectic topology we will use. We put an emphasis on Gromov's non-squeezing theorem (of which we give a proof in the linear case), and of the derived notion of symplectic capacity. We describe the notion of symplectic capacity by listing a series of "natural" properties they should satisfy, and construct a few explicit examples. We also give an up-to-date discussion of some symplectic embedding properties, and mention several open questions, which might be of genuine physical interest if answered positively).

- In Section 4 we apply the notion of symplectic capacity to a thorough study of two uncertainty principles: the already mentioned Robertson–Schrödinger principle, which we express in the lapidary form $c(W_\Sigma) \geq \frac{1}{2}h$ where c stands for an arbitrary symplectic capacity, and W_Σ for the Wigner ellipsoid $\frac{1}{2}\Sigma^{-1}z^2 \leq 1$ associated with a covariance matrix Σ. We thereafter study an uncertainty principle well-known in time-frequency analysis but unfortunately most of the time ignored by physicists, namely Hardy's theorem on the simultaneous sharp location of a function and of its Fourier transform. We apply our results to a study of sub-Gaussian mixed quantum states, and again express our results concisely in terms of symplectic capacities.

- In Section 5 we introduce a notion of quantum phase space, whose points are "quantum blobs"; the symplectic capacity of a quantum blob is $\frac{1}{2}h$ that is, one half of the quantum of action. Quantum blobs can be viewed as the geometric (or topological) equivalents of squeezed states, and as such saturate the Robertson–Schrödinger uncertainty inequalities.

- In Section 6 we show that surprisingly enough the notion of symplectic capacity can be used to recover the EBK ground energy level of integrable Hamiltonian systems. For the sake of completeness we begin the section by reviewing the notion of EBK quantization of Lagrangian tori. We also show in detail that EBK quantization is exact for systems with quadratic Hamiltonian.

- We end our paper by listing a few fascinating questions whose answer seems difficult to give until new mathematical developments have made them more tractable; one particularly intriguing topic seems to be Guth's catalyst map.

1.3 Notation and terminology

The letter n will stand for an (arbitrary) integer ≥ 1; the product vector space $\mathbb{R}^n \times \mathbb{R}^n$ is identified with $\mathbb{R}^{2n}$ and the generic point of $\mathbb{R}^{2n}$ is usually denoted z or (x,p) where we have set $x = (x_1, ..., x_n)$, $p = (p_1, ..., p_n)$. When matrix calculations are performed, x, p, z are viewed as column vectors. The standard Euclidean star product is denoted by a dot $\cdot$ in all dimensions.

When M is a matrix we will denote the product $z^T M z$ indifferently by $Mz \cdot z$ or Mz^2.

We will use the following unitary Fourier transforms:

- The $\hbar$-Fourier transform on $\mathbb{R}^n$ defined by

$$F^\hbar \psi(p) = \left(\tfrac{1}{2\pi\hbar}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\frac{i}{\hbar} p \cdot x} \psi(x) dx$$

  and we have $(F^\hbar)^{-1}\psi = \overline{F^\hbar \overline{\psi}}$.

- The symplectic $\hbar$-Fourier transform on $\mathbb{R}^{2n}$ defined by

$$F_\sigma^\hbar a(z) = F^\hbar a(-Jz) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int_{\mathbb{R}^{2n}} e^{-\frac{i}{\hbar}\sigma(z,z')} a(z') dz';$$

  $F_\sigma^\hbar$ is its own inverse: $F_\sigma^\hbar F_\sigma^\hbar$ is the identity.

The Weyl operator $\widehat{A}$ associated with an observable (or "symbol") a is formally defined by

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

this is equivalent to the much more tractable formula

$$\widehat{A}\psi(x) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int_{\mathbb{R}^{2n}} a_\sigma(z_0)\widehat{T}(z_0)\psi(x)dz_0 \tag{2}$$

where $a_\sigma = F_\sigma^\hbar a$ and $\widehat{T}(z_0)$ is the *Heisenberg–Weyl* operator:

$$\widehat{T}(z_0)\psi(x) = e^{\frac{i}{\hbar}(p_0\cdot x - \frac{1}{2}p_0\cdot x_0)}\psi(x - x_0). \tag{3}$$

We will use the shorthand notation $\widehat{A} \overset{\text{Weyl}}{\longleftrightarrow} a$ to signify that $\widehat{A}$ is the Weyl operator with symbol a (we refer to Littlejohn [61] for a very good review of Weyl calculus from a physicist's point of view).

We will use the notation $\widehat{X_j}$ and $\widehat{P_j}$ for the Weyl operators with symbols x_j and p_j, respectively. We have, explicitly, $\widehat{X_j}\psi = x_j\psi$ and $\widehat{P_j}\psi = -i\hbar\partial\psi/\partial x_j$.

2 Symplectic Group and Hamiltonian Dynamics

Hamiltonian systems (that is, dynamical systems whose time-evolution is governed by Hamilton's equations of motion) are particularly adequate for giving a physical illustration of the power of symplectic geometry in physical problems. Their consideration will allow us in Subsection 3.1.1 to give a dynamical description of the "principle of the symplectic camel", a central and recurring theme in this paper.

There are numerous references for the material of this preliminary subsection; classical sources are for instance Arnol'd [2] or Goldstein [22]; on a more abstract level Hofer and Zehnder [47] give a concise introduction mainly accessible to mathematicians non familiar with the topic.

2.1 The symplectic group

Here we collect a few basic results on symplectic geometry; the notion of symplectic spectrum is analyzed in some detail.

2.1.1 General definitions

A real $2n \times 2n$ matrix S is said to be "symplectic" if it satisfies the two (equivalent) conditions $S^TJS = J$ and $SJS^T = J$ where J is the block matrix

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

(J is the "standard symplectic matrix; it is sometimes called "Weyl matrix" in optics). Symplectic matrices form a group, the symplectic group $\mathrm{Sp}(2n,\mathbb{R})$. The standard symplectic form on $\mathbb{R}^{2n}$ is the non-degenerate antisymmetric bilinear form defined by

$$\sigma(z,z') = Jz\cdot z' = p\cdot x' - p'\cdot x \tag{4}$$

for $z=(x,p)$, $z'=(x',p')$; in differential notation

$$\sigma = dp_1 \wedge dx_1 + \cdots + dp_n \wedge dx_n. \tag{5}$$

A linear automorphism s of $\mathbb{R}^{2n}$ is symplectic if $\sigma(sz,sz')=\sigma(z,z')$ for all vectors z,z' in $\mathbb{R}^{2n}$. Identifying s with its matrix S in the canonical basis, s is symplectic if and only if S is a symplectic matrix. We will always implicitly use the identification $s=S$. Symplectic matrices have determinant equal to one and $\mathrm{Sp}(2n,\mathbb{R})$ is a connected subgroup of the special linear real group $SL(2n,\mathbb{R})$.

If one writes S in block matrix form

$$S = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where A,B,C,D are $n\times n$ matrices, the condition $S\in\mathrm{Sp}(2n,\mathbb{R})$ is equivalent to the conditions[1]

$$A^TC,\ B^TD \ \text{ are symmetric, and } A^TD - C^TB = I \tag{6}$$

$$AB^T,\ CD^T \ \text{ are symmetric, and } AD^T - BC^T = I. \tag{7}$$

It follows from the second of these sets of conditions that the inverse of S is given by

$$S^{-1} = \begin{bmatrix} D^T & -B^T \\ -C^T & A^T \end{bmatrix}. \tag{8}$$

The inhomogeneous (or affine) symplectic group $\mathrm{ISp}(2n,\mathbb{R})$ is the semidirect product of $\mathrm{Sp}(2n,\mathbb{R})$ and of the translation group of $\mathbb{R}^{2n}$; its elements are the product (in any order) of symplectic matrices and of phase-space translations.

[1] They are sometimes called the "Lüneburg relations" in optics

2.1.2 The subgroup $U(n)$ of $\mathrm{Sp}(2n,\mathbb{R})$

The unitary group $U(n,\mathbb{C})$ is identified with a subgroup $U(n)$ of $\mathrm{Sp}(2n,\mathbb{R})$ by the monomorphism

$$A+iB\longmapsto\begin{bmatrix}A & -B\\ B & A\end{bmatrix}$$

and we have

$$U\in U(n)\text{ if and only if }UJ=JU. \tag{9}$$

The following property of $U(n)$ is important; it says that $U(n)$ consists of *symplectic rotations*:

$$U(n)=\mathrm{Sp}(2n,\mathbb{R})\cap O(2n). \tag{10}$$

A basis $\mathcal{B}=(e_1,...,e_n;f_1,...,f_1)$ of $(\mathbb{R}^{2n},\sigma)$ is symplectic if we have

$$\sigma(e_i,e_j)=\sigma(f_i,f_j)=0\ ,\ \sigma(f_i,e_j)=-\delta_{ij} \tag{11}$$

for all $i,j=1,2,...,n$. The canonical basis $\mathcal{C}$ of $\mathbb{R}^{2n}$, defined by $e_i=((\delta_{ij})_{1\leq j\leq n};0)$, $f_i=(0;(\delta_{ij})_{1\leq j\leq n})$ is symplectic. An orthosymplectic basis of $(\mathbb{R}^{2n},\sigma)$ is a basis of $\mathbb{R}^{2n}$ which is both symplectic and orthonormal (for the standard scalar product on $\mathbb{R}^{2n}$); the canonical basis $\mathcal{C}$ is orthosymplectic. A linear mapping $\mathbb{R}^{2n}\longrightarrow\mathbb{R}^{2n}$ is symplectic if and only if it takes one (and hence all) symplectic bases $\mathcal{B}$ to a symplectic basis $\mathcal{B}'$; the elements of the subgroup $U(n)$ of $\mathrm{Sp}(2n,\mathbb{R})$ are precisely those who take one (and hence all) orthosymplectic basis to an orthosymplectic basis.

Any square matrix can be written as the product of a positive-definite matrix and a rotation: this is the polar decomposition theorem. In the symplectic case this theorem becomes (see for instance de Gosson [32], §2.2.1):

Proposition 1 *Every $S\in\mathrm{Sp}(2n,\mathbb{R})$ can be written in the form $S=RU$ where $R\in\mathrm{Sp}(2n,\mathbb{R})$ is positive-definite and $U\in U(n)$.*

Positive definite symplectic matrices themselves can be diagonalized using a symplectic rotation:

Proposition 2 *Let $S\in\mathrm{Sp}(2n,\mathbb{R})$ be symmetric and positive definite. There exists $U\in U(n)$ such that*

$$S=U^T\Delta U\quad,\quad\Delta=\begin{bmatrix}\Lambda & 0\\ 0 & \Lambda^{-1}\end{bmatrix} \tag{12}$$

where $\Lambda=\mathrm{diag}[\lambda_1,...,\lambda_n]$ and $0<\lambda_1\leq\lambda_2\leq\cdots\leq\lambda_n\leq 1$ are the n first eigenvalues of S (counted with their multiplicities).

Proof. Since S is symmetric and positive definite its eigenvalues occur in pairs $(\lambda, 1/\lambda)$ with $\lambda > 0$ so that if $\lambda_1 \leq \cdots \leq \lambda_n$ are n eigenvalues then $1/\lambda_1, ..., 1/\lambda_n$ are the other n eigenvalues. Let now U be an orthogonal matrix such that $S = U^T \Delta U$ where

$$D = \mathrm{diag}[\lambda_1, ..., \lambda_n; 1/\lambda_1, ..., 1/\lambda_n];$$

let us prove that $U \in U(n)$. Let $u_1, ..., u_n$ be orthonormal eigenvectors of U corresponding to the eigenvalues $\lambda_1, ..., \lambda_n$. Since $SJ = JS^{-1}$ (because S is both symplectic and symmetric) we have

$$SJu_k = JS^{-1}u_k = \frac{1}{\lambda_j} Ju_k$$

for $1 \leq k \leq n$, hence $\pm Ju_1, ..., \pm Ju_n$ are the orthonormal eigenvectors of U corresponding to the remaining n eigenvalues $1/\lambda_1, ..., 1/\lambda_n$. Write now the $2n \times n$ matrix $[u_1, ..., u_n]$ as

$$[u_1, ..., u_n] = \begin{bmatrix} A \\ B \end{bmatrix}$$

where A and B are of order $n \times n$; we have

$$[-Ju_1, ..., -Ju_n] = -J \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} -B \\ A \end{bmatrix}$$

hence U is of the type

$$U = [u_1, ..., u_n; -Ju_1, ..., -Ju_n] = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}.$$

Since $U^T U = I$ the blocks A and B are such that

$$AB^T = B^T A \ \ , \ \ AA^T + BB^T = I \tag{13}$$

hence U is also symplectic, that is $U \in U(n)$. ■

2.1.3 The metaplectic representation of the symplectic group

We will need some simple result about the theory of the *metaplectic group* $\mathrm{Mp}(2n, \mathbb{R})$ (see de Gosson [32], Chapter 7, for an up-to-date and complete exposition). The metaplectic group is a faithful (but not irreducible) unitary representation on $L^2(\mathbb{R}^n)$ of the double covering $\mathrm{Sp}_2(2n, \mathbb{R})$ of the symplectic group $\mathrm{Sp}(2n, \mathbb{R})$. Below we list a set of generators of $\mathrm{Mp}(2n, \mathbb{R})$ and we give the images of these generators under the covering projection $\pi^\hbar$: $\mathrm{Mp}(2n, \mathbb{R}) \longrightarrow \mathrm{Sp}(2n, \mathbb{R})$, which to every pair of operators $\pm\widehat{S} \in \mathrm{Mp}(2n, \mathbb{R})$ associates the symplectic matrix $\pi^\hbar(\pm\widehat{S}) = S$.

- *The modified Fourier transform* $\widehat{J} = i^{-n/2}F^{\hbar}$ *where* $F^{\hbar}$ *is the* $\hbar$-*Fourier transform:*

$$\widehat{J}\psi(x) = \left(\tfrac{1}{2\pi\hbar}\right)^{n/2}\int_{\mathbb{R}^n} e^{-\frac{i}{\hbar}x\cdot x'}\psi(x')dx' \quad , \quad \pi^{\hbar}(\pm\widehat{J}) = J \tag{14}$$

- *The "chirp" operators* $\widehat{V}_{-P}$ *with* $P = P^T$ *a real* $n\times n$ *matrix:*

$$\widehat{V}_{-P}\psi(x) = e^{\frac{i}{2\hbar}Px^2}\psi(x) \quad , \quad \pi^{\hbar}(\pm\widehat{V}_{-P}) = \begin{bmatrix} I & 0 \\ P & I \end{bmatrix} \tag{15}$$

- *The rescaling operators* $\widehat{M}_{L,m}$ *with* $\det L \neq 0$, $m\pi \equiv \arg\det L \bmod 2\pi$*:*

$$\widehat{M}_{L,m}\psi(x) = i^m\sqrt{|\det L|}\psi(Lx) \quad , \quad \pi^{\hbar}(\pm\widehat{M}_{L,m}) = \begin{bmatrix} L^{-1} & 0 \\ 0 & L^T \end{bmatrix}. \tag{16}$$

The existence of the metaplectic group is motivated in many texts by invoking the Heisenberg–Weyl operators and Stone–von Neumann's theorem; however such an explanation is not very conclusive, because not only is it hard to see what operators $\mathrm{Mp}(2n,\mathbb{R})$ should contain but, worse, such an argument only guarantees the existence of a unitary projective representation of the symplectic group (admittedly, the latter point does not have earthshaking consequences as long as metaplectic operators only appear in a passive way, that is, as conjugation operators, since all difficulties related to phase ambiguities then disappear).

The relevance of the metaplectic group for Weyl calculus (and in particular the theory of density matrices) comes from the following property, known as metaplectic (or, symplectic covariance):

Proposition 3 *Let* $\widehat{A} \overset{\mathit{Weyl}}{\longleftrightarrow} a$ *be a Weyl operator and* $\widehat{S} \in \mathrm{Mp}(2n,\mathbb{R})$ *such that* $\pi^{\hbar}(\widehat{S}) = S$*. Then* $\widehat{S}\widehat{A}\widehat{S}^{-1} \overset{\mathit{Weyl}}{\longleftrightarrow} a\circ S^{-1}$*.*

This result actually justifies the use of Weyl calculus in quantum mechanics, because one can prove (Shale [84]) that it is, among all pseudo-differential calculi, the only which enjoys the metaplectic covariance property. Weyl calculus thus imposes itself in any quantization procedure of observables which is meant to preserve the symplectic character of classical (Hamiltonian) mechanics.

2.2 Williamson's theorem and symplectic spectrum

Let M be a real $m \times m$ symmetric matrix. Elementary linear algebra tells us that all the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ of M are real, and that M can be diagonalized using an orthogonal transformation. Williamson's theorem provides us with the symplectic variant of this result. It says that every symmetric and positive definite matrix M can be diagonalized using symplectic matrices, and this in a very particular way.

2.2.1 Symplectic diagonalization

Let M be a positive definite real $2n \times 2n$ matrix; the eigenvalues of JM are those of the antisymmetric matrix $M^{1/2}JM^{1/2}$ and are thus of the type $\pm i\lambda_j^\sigma$ with $\lambda_j^\sigma > 0$. We will always arrange the λ_j^σ in *decreasing* order[2]:

$$\lambda_1^\sigma \geq \lambda_2^\sigma \geq \cdots \geq \lambda_n^\sigma. \tag{17}$$

Definition 4 *The positive numbers λ_j^σ are the "symplectic eigenvalues" of the positive definite matrix M. The decreasing sequence*

$$\operatorname{Spec}_\sigma(M) = (\lambda_1^\sigma, ..., \lambda_n^\sigma)$$

is called the "symplectic spectrum" of M.

We will study some important properties of the symplectic spectrum in a moment, but let us first state and prove an essential diagonalization result going back to Williamson's pioneering work [95] in 1933 on normal forms. This result has been rediscovered several times by various authors (Hofer and Zehnder [47] claim that it can already be found, albeit in disguise, in the 1858 work [92] of Weierstrass[3]).

Theorem 5 (Williamson) *There exists $S \in \operatorname{Sp}(2n, \mathbb{R})$ such that*

$$S^T M S = \begin{bmatrix} \Lambda^\sigma & 0 \\ 0 & \Lambda^\sigma \end{bmatrix} \tag{18}$$

where $\Lambda^\sigma = \operatorname{diag}[\lambda_1^\sigma, ..., \lambda_n^\sigma]$ the λ_j^σ being the symplectic eigenvalues of M.

Proof. (See [47], §1.7, Theorem 8, for an alternative proof using a variational argument; for a constructive proof see Narcowich [74], Theorem 4.2).

[2] One sometimes finds the opposite convention in the literature.

[3] The present authors haven't verified this claim.

Let us denote by $\langle \cdot, \cdot \rangle$ the scalar product associated with the symmetric matrix M^{-1}, that is $\langle z, z' \rangle = M^{-1}z \cdot z'$. Setting $K = -MJ$; we have

$$\left\langle z, Kz' \right\rangle = -z \cdot Jz' = \sigma(z, z') \tag{19}$$

for all z, z'. We now observe that $-MJ$ and $-JM$ have the same eigenvalues (both matrices are equivalent to $-M^{-1/2}JM^{-1/2}$) hence the eigenvalues of K are the numbers $\pm i\lambda_j^\sigma$. The eigenvectors of K occurring in conjugate pairs $e_j' \pm if_j'$ we thus obtain a $\langle \cdot, \cdot \rangle$-orthonormal basis $\mathcal{B}' = (e_1', ..., e_n'; f_1', ..., f_1')$ of $\mathbb{R}^{2n}$ such that $Ke_i' = \lambda_j^\sigma f_i'$ and $Kf_j' = -\lambda_j^\sigma e_j'$. Notice that it follows from these relations that

$$K^2 e_i' = -(\lambda_i^\sigma)^2 e_i' \quad , \quad K^2 f_j' = -(\lambda_j^\sigma)^2 f_j'$$

and that the vectors of the basis $\mathcal{B}'$ obey, in view of (19), the relations

$$\begin{aligned}
\sigma(e_i', e_j') &= \langle e_i', Ke_j' \rangle = \lambda_j^\sigma \langle e_i', f_j' \rangle = 0 \\
\sigma(f_i', f_j') &= \langle f_i', Kf_j' \rangle = -\lambda_j^\sigma \langle f_i', e_j' \rangle = 0 \\
\sigma(f_i', e_j') &= \langle f_i', Ke_j' \rangle = \lambda_j^\sigma \langle f_i', f_j' \rangle = -\lambda_j^\sigma \delta_{ij}.
\end{aligned}$$

Setting $e_i = (\lambda_i^\sigma)^{-1/2} e_i'$ and $f_j = (\lambda_j^\sigma)^{-1/2} f_j'$, the basis $\mathcal{B}$ =$(e_1, ..., e_n; f_1, ..., f_1)$ is thus symplectic: we have

$$\sigma(e_i, e_j) = \sigma(f_i, f_j) = 0 \ , \ \sigma(f_i, e_j) = -\delta_{ij}$$

for $1 \leq i, j \leq n$. Let now S be the element of $\mathrm{Sp}(2n, \mathbb{R})$ mapping the canonical symplectic basis $\mathcal{C}$ to $\mathcal{B}$. The $\langle \cdot, \cdot \rangle$-orthogonality of $\mathcal{B}$ implies formula (18) with $\Lambda^\sigma = \operatorname{diag}[\lambda_1^\sigma, ..., \lambda_n^\sigma]$. ■

Remark 6 *Observe that the diagonalization formula (12) for positive-definite symplectic matrices is not a Williamson diagonalization. However, writing*

$$S = (\Delta^{1/2}U)^T(\Delta^{1/2}U) \tag{20}$$

where $\Delta^{1/2}$ is the positive square root of Δ we obtain the Williamson diagonalization of S. This is because the eigenvalues of JS are those of $S^{1/2}JS^{1/2} = J$ ($S^{1/2}$ is symplectic and symmetric), and hence their moduli are all equal to $+1$.

Let us complement Williamson's theorem by a uniqueness result modulo symplectic rotations:

Proposition 7 *The symplectic matrix S diagonalizing M in Williamson's theorem is unique up to a symplectic rotation: if S' is another Williamson diagonalizing symplectic matrix then $S^{-1}S' \in U(n)$.*

Proof. Assume that $S' \in \mathrm{Sp}(2n, \mathbb{R})$ is such that

$$S'^T M S' = S^T M S = D = \begin{bmatrix} \Lambda^\sigma & 0 \\ 0 & \Lambda^\sigma \end{bmatrix}$$

and set $U = S^{-1}S'$; we have $U^T D U = D$. Let us show that $UJ = JU$; it will follow that $U \in U(n)$. Setting $R = D^{1/2}UD^{-1/2}$ we have

$$R^T R = D^{-1/2}(U^T D U)D^{-1/2} = D^{-1/2}DD^{-1/2} = I$$

hence $R \in O(2n)$. Since J commutes with each power of D we have, since $JU = (U^T)^{-1}J$ (because U is symplectic),

$$\begin{aligned} JR &= D^{1/2}JUD^{-1/2} = D^{1/2}(U^T)^{-1}JD^{-1/2} \\ &= D^{1/2}(U^T)^{-1}D^{-1/2}J = (R^T)^{-1}J \end{aligned}$$

hence $R \in \mathrm{Sp}(2n, \mathbb{R}) \cap O(2n)$; in view of (10) we thus have $R \in U(n)$ hence $JR = RJ$. Now $U = D^{-1/2}RD^{1/2}$ and therefore

$$\begin{aligned} JU &= JD^{-1/2}RD^{1/2} = D^{-1/2}JRD^{1/2} \\ &= D^{-1/2}RJD^{1/2} = D^{-1/2}RD^{1/2}J \\ &= UJ \end{aligned}$$

which was to be proven. ■

Remark 8 *A natural question is whether one can find other symplectic diagonalizations of the type (18) where Λ^σ is replaced by some diagonal matrix whose diagonal entries are numbers λ_j different from the λ_j^σ. The answer is negative: if $\Lambda = \mathrm{diag}\,[\lambda_1, ..., \lambda_n]$ then the sets $\{\lambda_1^\sigma, ..., \lambda_n^\sigma\}$ and $\{\lambda_1, ..., \lambda_n\}$ are the same (see de Gosson [32], §8.3.1, Theorem 8.11(ii) for a proof).*

2.2.2 Properties of the symplectic spectrum

The symplectic spectrum of a positive-definite matrix has three important properties, the third being not quite trivial (it is the symplectic version of the Rayleigh–Courant–Fisher theorem from convex optimization; see Arnol'd [2], §24). Let us list and prove these properties:

Proposition 9 *Let M and M' be a real positive-definite matrix of order $2n$. The symplectic spectrum* $\operatorname{Spec}_\sigma(M)$ *has the following properties:*

(i) $\operatorname{Spec}_\sigma(M)$ *is a symplectic invariant:* $\operatorname{Spec}_\sigma(S^TMS) = \operatorname{Spec}_\sigma(M)$ *for every $S \in \operatorname{Sp}(2n, \mathbb{R})$;*

(ii) If $\operatorname{Spec}_\sigma(M) = (\lambda_1^\sigma, ..., \lambda_n^\sigma)$*, then* $\operatorname{Spec}_\sigma(M^{-1}) = ((\lambda_n^\sigma)^{-1}, ..., (\lambda_1^\sigma)^{-1})$*;*

(iii) If $M \leq M'$, then $\operatorname{Spec}_\sigma(M) \leq \operatorname{Spec}_\sigma(M')$*, i.e. we have $\lambda_j^\sigma \leq \lambda_j'^\sigma$ for all $j = 1, ..., n$.*

Proof. (i) follows from the fact that $J(S^TMS) = S^{-1}(JM)S$ (because $JS^T = S^{-1}J$ since S is symplectic) which implies that JM and $J(S^TMS)$ have the same eigenvalues. (ii) The eigenvalues of JM are the same as those of $M^{1/2}JM^{1/2}$; the eigenvalues of JM^{-1} are those of $M^{-1/2}JM^{-1/2}$. Now

$$M^{-1/2}JM^{-1/2} = -(M^{1/2}JM^{1/2})^{-1}$$

hence the eigenvalues of JM and JM^{-1} are obtained from each other by the transformation $t \longmapsto -1/t$. The result follows since the symplectic spectra are obtained by taking the moduli of these eigenvalues. (iii) (cf. Giedke et al. [21]). Let us introduce some notation: when two matrices A and B have the same eigenvalues we will write $A \simeq B$. Notice that when A or B is invertible we have $AB \simeq BA$. We will write $M \leq M'$ when $M' - M$ is non-negative; equivalently, $M \leq M'$ is equivalent to $z^TMz \leq z^TM'z$ for every $z \in \mathbb{R}^{2n}$. We now observe that the statement (ii) is equivalent to

$$M \leq M' \Longrightarrow (JM')^2 \leq (JM)^2$$

since the eigenvalues of JM and JM' occur in pairs $\pm i\lambda$, $\pm i\lambda'$ with λ and λ' real. Replacing z by successively $(JM^{1/2})z$ and $(JM'^{1/2})z$ in $z^TMz \leq z^TM'z$ we thus have, taking into account the fact that $J^T = -J$, that is, since $J^T = -J$,

$$M^{1/2}JM'JM^{1/2} \leq M^{1/2}JMJM^{1/2}. \tag{21}$$

$$M'^{1/2}JM'JM'^{1/2} \leq M'^{1/2}JMJM'^{1/2}. \tag{22}$$

Noting that we have $M^{1/2}JM'JM^{1/2} \simeq MJM'J$ and $M'^{1/2}JMJM'^{1/2} \simeq M'JMJ \simeq MJM'J$ we can rewrite the relations (21) and (22) as

$$MJM' \leq JM^{1/2}JM'JM^{1/2} \ , \ M'^{1/2}JM'JM'^{1/2} \leq MJM'J$$

and hence, by transitivity,

$$M'^{1/2}JM'JM'^{1/2} \leq M^{1/2}JMJM^{1/2}. \tag{23}$$

Since we have $M^{1/2}JMJM^{1/2} \simeq (MJ)^2$ and $M'^{1/2}JM'JM'^{1/2} \simeq (M'J)^2$ the relation (23) is equivalent to $(M'J)^2 \leq (MJ)^2$, which was to be proven. ■

2.3 Hamiltonian mechanics

2.3.1 Hamilton's equations

Hamiltonian system describe motions involving holonomic constraints and forces arising from a potential. The phase-space evolution of a Hamiltonian system with n degrees of freedom is governed by Hamilton's equations of motion

$$\frac{dx_j}{dt} = \frac{\partial H}{\partial p_j}(x,p,t) \quad , \quad \frac{dp_j}{dt} = -\frac{\partial H}{\partial x_j}(x,p,t) \tag{24}$$

$(j = 1, ..., n)$. This system of differential equations can be written in concise form as

$$\frac{dz}{dt} = X_H(z,t) = J\partial_z H(z,t)$$

where $\partial_z = (\partial_x, \partial_p)$ is the gradient in the variables $x_1, ..., x_n, p_1, ..., p_n$. The operator $X_H = J\partial_z H$ is the *Hamilton vector field*; it is a true vector field only when H is time-independent.

When H is of the type "kinetic energy + potential energy", that is

$$H(x,p,t) = \sum_{j=1}^{n} \frac{p_j^2}{2m_j} + U(x,t)$$

Hamilton's equations become

$$\frac{dx_j}{dt} = \frac{p_j}{m_j} \quad , \quad \frac{dp_j}{dt} = -\frac{\partial U}{\partial x_j}(x,t)$$

which is just Newton's second law. Assuming that the equations (24) have a unique solution $(x,p) = (x(t), p(t))$ for every initial condition (at time t') $x(t') = x'$, $p(t') = p'$, the phase space flow $(f^H_{t,t'})$ is globally defined: $f^H_{t,t'}$ is the mapping $\mathbb{R}^{2n} \longrightarrow \mathbb{R}^{2n}$ defined by $(x,p) = f^H_{t,t'}(x',p')$. We will use the shorthand notation $f^H_t = f^H_{t,0}$; when the Hamiltonian H is time independent (f^H_t) is the usual flow.

Definition 10 *A canonical transformation (or: symplectic diffeomorphism, or symplectomorphism) is a diffeomorphism f of $\mathbb{R}^{2n}$ whose Jacobian matrix*

$$Df(x',p') = \frac{\partial(x,p)}{\partial(x',p')}$$

is symplectic at every point (x',p'). Equivalently $f^*\sigma = \sigma$ when σ is written as the differential form $dp \wedge dx$ (formula (5)).

A characteristic property of Hamiltonian systems is that their flows consist of canonical transformations: the Jacobian of each $f_{t,t'}^H$ is a symplectic matrix everywhere where it is defined, that is $(f_t^H)^*\sigma = \sigma$. We will have much more to say about this property.

2.3.2 Hamiltonian orbits on hypersurfaces

Consider a time-independent Hamiltonian function H and denote by S_E the ("energy shell") defined by the equation $H(z) = E$ (we assume that Σ_E is non-empty, and that E is a regular value of the energy, so that S_E is a smooth hypersurface in $\mathbb{R}^{2n}$). If we choose a point z_0 on S_E then, due to the theorem of conservation of energy, the points $f_t^H(z_0)$ will also be in S_E and this for all times: energy shells "trap" Hamiltonian trajectories. Let us now reverse the situation, and consider an arbitrary hypersurface S in $\mathbb{R}^{2n}$ and assume that there exist two Hamiltonian functions H and K for which S is an energy shell for some regular values (i.e. $\partial_z H \neq 0$ and $\partial_z K \neq 0$ on Σ) of the energy:

$$\Sigma = \{z : H(z) = E\} = \{z : K(z) = F\}. \tag{25}$$

We claim that in this case H and K have same trajectories on Σ. Intuitively the idea is very simple: the vector fields $\partial_z H$ and $\partial_z K$ being both normal to the constant energy shell Σ, the Hamiltonian fields $X_H = (\partial_p H, -\partial_x H)$ and $X_K = (\partial_p K, -\partial_x K)$ must be proportional and thus have the same trajectories (up to a reparametrization). Slightly more formally: since $\partial_z H(z) \neq 0$ and $\partial_z K(z) \neq 0$ are both normal to Σ at z, there exists a function $\alpha \neq 0$ such that $X_K = \alpha X_H$ on Σ. Let now (f_t^H) and (f_t^K) be the flows of H and K, respectively, and define a function $t = t(z,s)$, $s \in \mathbb{R}$, as being the solution of the ordinary differential equation

$$\frac{dt}{ds} = \alpha(f_t(z)) \quad , \quad t(z,0) = 0$$

where z is viewed as a parameter. We claim that

$$f_s^K(z) = f_t^H(z) \quad for \quad z \in S. \tag{26}$$

In fact, by the chain rule

$$\frac{d}{ds} f_t^H(z) = \frac{d}{dt} f_t^H(z) \frac{dt}{ds} = X_H(f_t^H(z))\alpha(f_t^H(z))$$

that is, since $X_K = \alpha X_H$:

$$\frac{d}{ds} f_t^H(z) = X_K(f_t^H(z))$$

which shows that the mapping $s \longmapsto f^H_{t(z,s)}(z)$ is a solution of the differential equation $\frac{d}{dt}z = X_H(z)$ passing through z at time $s = t(z,0) = 0$. By the uniqueness theorem on solutions of systems of differential equations, this mapping must be identical to the mapping $s \longmapsto f^K_s(z)$; hence the equality (26). Both Hamiltonian functions H and K thus have the same trajectories, as claimed.

2.3.3 Liouville's theorem

One of the most well-known properties of Hamiltonian flows are that they are volume preserving: this statement is the famous Liouville's theorem from classical mechanics. This property attracted a lot of attention already more than a century ago; it served as the main stimulating force for the creation of *ergodic theory*, nowadays a well-established mathematical theory which studies various recurrence properties of volume-preserving transformations (see Polterovich [76], especially Chapter 11). All known text-book proofs of Liouville's theorem are, in one way or another, variations on the following fact: a Hamiltonian flow consists of canonical transformations and the Jacobian determinant of a canonical transformation is equal to one; alternatively (and this is sufficient for deriving Liouville's theorem) Hamiltonian vector fields are divergence-free. Liouville's theorem thus actually holds for all mechanical systems whose motion is governed by a flow associated to a divergence-free vector field, and is thus not a characteristic feature of Hamiltonian system! Liouville's theorem is actually often restated in the form of the "Liouville's equation"

$$\frac{\partial \rho}{\partial t} + \{H, \rho\} = 0$$

where $\{\cdot, \cdot\}$ is the Poisson bracket, and $\rho = \rho(z,t)$ plays the role of a "phase-space fluid density at the point z at time t". Liouville's equation is equivalent to saying that for every z the quantity $\rho(f^H_t(z), t)$ is constant in time, hence it expresses the "incompressibility" of the phase-space fluid, and is thus equivalent to Liouville's theorem itself.

A better way (from our perspective) to prove Liouville's theorem is to note that it directly follows from the fact that the f^H_t are canonical transformations. This point of view namely highlights the fact that Liouville's theorem is a rather weak statement about Hamiltonian flows as soon as $n > 1$. Let Ω be a measurable subset of $\mathbb{R}^{2n}$ and set $\Omega_t = f^H_t(\Omega)$. Defining the Lebesgue measure on $\mathbb{R}^{2n}$ by

$$dz = dpdx = \prod\nolimits_{j=1}^{n} dp_j dx_j$$

we have, using the formula of change of variables for multiple integrals,

$$\begin{aligned}\mathrm{Vol}(\Omega_t) &= \int_{\Omega_t} dpdx \\ &= \int_{\Omega} \left| \det \frac{\partial(x,p)}{\partial(x',p')} \right| dp'dx' \\ &= \mathrm{Vol}(\Omega)\end{aligned}$$

where the last equality follows from the fact that the Jacobian matrix $\partial(x,p)/\partial(x',p')$ is symplectic and thus has determinant equal to one.

Liouville's theorem is a perfect transition to the "property of the symplectic camel", the main theme of this paper, and to which it opens the door.

3 The Symplectic Camel

> "*...It is easier for a camel to pass through the eye of a needle than for one who is rich to enter the kingdom of God*[4]*...*"

As we will see this Biblical quotation applies very well (in a metaphoric form!) to the symplectic non-squeezing theorem; for various interpretations of the word "camel" see the comments to the on-line version of Reich's paper [78].

3.1 The symplectic non-squeezing theorem

Let us describe in detail Gromov's symplectic non-squeezing theorem; we begin with a rather informal and "intuitive" description from a dynamical point of view.

3.1.1 An informal dynamical description

In addition of being volume-preserving, Hamiltonian flows have an unexpected additional property as soon as the number of degrees of freedom is superior to one; this property is a consequence of the symplectic non-squeezing theorem which was proved in 1985 by M. Gromov [41].

Assume that the system $\mathcal{S}$ consists of a very large number of point-like particles forming a "cloud" filling a subset Ω of phase space $\mathbb{R}^{2n}$. Suppose that this cloud is, at time $t = 0$ spherical so Ω is a phase space ball

[4] Matthew **19**(24), St Luke **18**(25), Mk **10**(25).

$B^{2n}(r) : |z - z_0| \leq r$. The orthogonal projection of that ball on any plane of coordinates will always be a circle with area πr^2. As time evolves, the cloud of points will distort and may take after a while a very different shape, while keeping constant volume. In fact, since conservation of volume has noting to do with conservation of shape, one might very well envisage that the ball $B^{2n}(r)$ can be stretched in all directions by the Hamiltonian flow (f_t^H), and eventually get very thinly spread out over huge regions of phase space, so that the projections on any plane could *a priori* become arbitrary small after some time t. This possibility is perfectly consistent with Katok's lemma [53], which can be stated as follows: consider two bounded domains Ω and Ω' in $\mathbb{R}^{2n}$ which are both diffeomorphic to the ball $B^{2n}(r)$ and have same volume. Then, for every $\varepsilon > 0$ there exists a Hamiltonian function H and a time t such that $\mathrm{Vol}(f_t^H(\Omega)\Delta\Omega')) < \varepsilon$. Here $f_t^H(\Omega)\Delta\Omega'$ denotes the symmetric difference of the two sets $f_t^H(\Omega)$ and Ω': it is the set of all points that are in $f_t^H(\Omega)$ or Ω', but not in both. Katok's lemma thus shows that up to sets of arbitrarily small measure ε any kind of phase-space spreading is a priori possible for a volume-preserving flow.

However (and this was unknown until 1985!) the projections of the set $f_t^H(B^{2n}(r))$ on any plane of *conjugate coordinates* x_j, p_j will never decrease below its original value! That projection will in fact be a surface in this plane, and that surface will have area at least πr^2. The condition that the plane is a plane of conjugate coordinates is essential: if we had chosen a plane of coordinates x_j, p_k with $j \neq k$ or any plane of coordinates x_j, x_k or p_j, p_k then the projection could become arbitrarily small. The property just described is a consequence (in fact, an equivalent statement) of "symplectic non-squeezing theorem" Gromov's.

Gromov's non-squeezing theorem, which we discuss in next subsection, is very surprising and has many indirect consequences. For instance, not so long time ago many people used to believe that whatever could be done by a volume-preserving diffeomorphism could be done (at least with an arbitrary good approximation) by symplectic diffeomorphisms (that is, canonical transformations). This belief was an extrapolation from the case $n = 1$ where both notions in fact coincide. The first step towards a better understanding of the peculiarities of symplectic diffeomorphisms was formulated in the early 1970's under the name of "Gromov's alternative": the group $\mathrm{Symp}(2n, \mathbb{R})$ of symplectic diffeomorphisms is either C^0-closed in the group of all diffeomorphisms, or its closure in the C^0-topology is the group of volume-preserving diffeomorphisms. One consequence of the symplectic non-squeezing theorem is that it is the first possibility that holds true: in general we cannot approximate a volume-preserving diffeomorphism by symplectic diffeomorphisms. (See the detailed discussion in Hofer and Zehnder [47].)

3.1.2 Two non-squeezing theorems

Let us denote by $Z_j^{2n}(R)$ the cylinder in $\mathbb{R}^{2n}$ defined by the condition: a point (x,p) is in $Z_j^{2n}(R)$ if and only if $x_j^2+p_j^2 \leq R^2$. Thus, the plane of conjugate coordinates x_j, p_j cuts $Z_j^{2n}(R)$ orthogonally, along the disk $x_j^2+p_j^2 \leq R^2$ lying in this plane. What Gromov proved, using complex analysis (the theory of pseudo-holomorphic curves) and symplectic topology, is the following:

Theorem 11 (Gromov) *If there exists a canonical transformation f in $\mathbb{R}^{2n}$ sending the ball $B^{2n}(r)$ in some cylinder $Z_j^{2n}(R)$, then we must have $r \leq R$.*

A word of explanation for the word "some" in the statement of Gromov's theorem. Let $Z_j^{2n}(R)$ and $Z_k^{2n}(R)$ be two cylinders with $j \neq k$. A phase space point $z=(x,p)$ is in $Z_j^{2n}(R)$ if and only if $x_j^2+p_j^2 \leq R^2$ and in $Z_k^{2n}(R)$ if and only if $x_k^2+p_k^2 \leq R^2$; we can switch from one condition to the other by the coordinate permutation $\tau_{j,k}$ swapping the pairs (x_j,p_j) and (x_k,p_k) and leaving all other pairs (x_ℓ,p_ℓ) invariant. Since $\tau_{j,k}$ obviously leaves $\sigma(z,z')$ unchanged it is a (linear) canonical transformation. Thus, if f sends $B^{2n}(r)$ in $Z_j^{2n}(R)$ then $\tau_{j,k}\circ f$ sends $B^{2n}(r)$ in $Z_k^{2n}(R)$ so that there exists a canonical transformation sending $B^{2n}(r)$ in $Z_j^{2n}(R)$ if and only if there exists one sending $B^{2n}(r)$ in $Z_k^{2n}(R)$.

The property of Hamiltonian flows described in the previous subsection immediately follows: assume that the orthogonal projection of $f_t^H(B^{2n}(r))$ on the x_j,p_j plane has an area $\leq \pi r^2$. Then any cylinder based on that plane and having a radius $\geq r$ would contain $f_t^H(B^{2n}(r))$.

It is essential for the non-squeezing theorem to hold that the considered cylinder is based on a x_j,p_j plane (or, more generally, on a symplectic plane). For instance, if we replace the cylinder $Z_j^{2n}(R)$ by the cylinder $Z_{12}^{2n}(R)$: $x_1^2+x_2^2 \leq R^2$ based on the x_1,x_2 plane, it is immediate to check that the linear canonical transformation f defined by $f(x,p)=(\lambda x, \lambda^{-1}p)$ sends $B^{2n}(r)$ into $Z_{12}^{2n}(R)$ as soon as $\lambda \leq r/R$. Also, one can always "squeeze" a large ball into a big cylinder using volume preserving diffeomorphisms that are not canonical. Here is an example in the case $n=2$ that is very easy to generalize to higher dimensions: define a linear mapping f by

$$f(x_1,x_2,p_1,p_2)=(\lambda x_1, \lambda^{-1}x_2, \lambda p_1, \lambda^{-1}p_2).$$

Clearly $\det f = 1$ and f is hence volume-preserving; f is however not symplectic if $\lambda \neq 1$. Choosing again $\lambda \leq r/R$, the mapping f sends $B^{2n}(R)$ into $Z_1^{2n}(r)$.

Gromov's theorem actually holds when $Z_j^{2n}(R)$ is replaced by any cylinder with radius R based on a symplectic plane, i.e. a two-dimensional subspace $\mathcal{P}$ of $\mathbb{R}^{2n}$ such that the restriction of σ to $\mathcal{P}$ is non-degenerate (equivalently, $\mathcal{P}$ has a basis $\{e, f\}$ such that $\sigma(e, f) \neq 0$). The planes $\mathcal{P}_j$ of coordinates x_j, p_j are of course symplectic, and given an arbitrary symplectic plane $\mathcal{P}$ it is easy to construct a linear canonical transformation S_j such that $S_j(\mathcal{P})= \mathcal{P}_j$. It follows that a canonical transformation f sends $B^{2n}(r)$ in the cylinder $Z_j^{2n}(R)$ if and only if $f \circ S_j$ sends $B^{2n}(r)$ in the cylinder $Z^{2n}(R)$ with same radius based on $\mathcal{P}$.

Here is another "non-squeezing" result; it is about tori and cylinder (see Sikorav in [87] and Polterovich [76], Theorem 1.1C and §4.3).

Theorem 12 *There is no canonical transformation taking the torus*

$$\mathbb{T}^n(r) = S_1^1(r) \times S_2^1(r) \times \cdots \times S_n^1(r)$$

into a cylinder $Z_j^{2n}(R)$ if $r > R$ ($S_j^1(r)$ is a circle with radius r in the x_j, p_j plane).

Notice that the result is obvious when $n = 1$ since in this case it just says that we cannot send a circle $S_1^1(r)$ inside a circle with smaller radius using area preserving transformations. However, if $n > 1$ then $\mathbb{T}^n(r)$ is a set with finite volume, and as for Gromov's theorem, one does not see why the statement should be true: it is in fact wrong (as is Gromov's theorem) if one uses arbitrary volume-preserving transformations (see Polterovich [76], §1.1).

3.1.3 Proofs of Gromov's theorem in the affine case

All known proofs (direct, or indirect) of Gromov's theorem are notoriously difficult, whatever method one uses (this might explain that it had not been discovered earlier, even in the more "physical" framework of Hamiltonian dynamics). We note that a related heuristic justification of Gromov's theorem is given by Hofer and Zehnder in [47] p.34; their "proof" however relies on an assumption which is, if true, at least as difficult to prove as Gromov's theorem itself! We are going to be much more modest, and to give a proof (actually two) of Gromov's theorem for affine canonical transformations; a canonical transformation is affine if can be factorized as the product of a symplectic transformation (i.e. an element of $\mathrm{Sp}(2n,\mathbb{R})$) and a phase space translation. We are actually going to give two proofs, both of an elementary nature (the second is shorter, but slightly more conceptual*).* To the best of our knowledge these proofs are original (for a "variational proof" see McDuff and Salamon [71] §2.4; in [32] §3.7.2 the first author has given a variant of this proof).

Proposition 13 ("Linear Gromov") *If there exists an affine canonical transformation f in $\mathbb{R}^{2n}$ sending a ball $B^{2n}(r)$ inside the cylinder $Z_j^{2n}(R)$, then we must have $r \leq R$. Equivalently, the intersection of $f(B^{2n}(r))$ by an affine plane parallel to a plane of conjugate coordinates x_j, p_j passing through the center of $f(B^{2n}(r))$ is an ellipse with area πr^2.*

Proof. *First proof.* It relies on the fact that the form $pdx = \sum_j p_j dx_j$ is a relative integral invariant of every canonical transformation, that is: if f is a canonical transformation and γ a cycle (or loop) in $\mathbb{R}^{2n}$ then

$$\oint_\gamma pdx = \oint_{f(\gamma)} pdx \tag{27}$$

(see for instance Arnol'd [2], §44, p.239). It is of course no restriction to assume that the ball $B^{2n}(r)$ is centered at the origin, and that f is a symplectic transformation S. We claim that the ellipse $\Gamma_j = S(B^{2n}(r)) \cap \mathcal{P}_j$, intersection of the ellipsoid $S(B^{2n}(r))$ with any plane $\mathcal{P}_j$ of conjugate coordinates x_j, p_j has area πr^2; the proposition immediately follows from this property. Let γ_j be the curve bounding the ellipse Γ_j and orient it positively; the area enclosed by γ is then

$$\mathrm{Area}(\Gamma_j) = \oint_{\gamma_j} p_j dx_j = \oint_{\gamma_j} pdx \tag{28}$$

hence, using property (27),

$$\mathrm{Area}(\Gamma_j) = \oint_{S^{-1}(\gamma_j)} pdx = \pi r^2 \tag{29}$$

(because $S^{-1}(\gamma)$ is a big circle of $B^{2n}(r)$); notice that the assumption that $\mathcal{P}_j$ is a plane of conjugate coordinates x_j, p_j is essential for the second equality (28) to hold, making the use of *Eqn.* (27) possible [more generally, the argument works for any symplectic plane]. *Second proof.* With the same notation as above we note that $S^{-1}(S(B^{2n}(r)) \cap \mathcal{P}_j)$ is a big circle of $B^{2n}(r)$, and hence encloses a surface with area πr^2. Now, $\mathcal{P}_j$ is a symplectic space when equipped with the skew-linear form $\sigma_j = dp_j \wedge dx_j$ and the restriction of S to $\mathcal{P}_j$ is also canonical from $(\mathcal{P}_j, \sigma_j)$ to the symplectic plane $S(\mathcal{P}_j)$ equipped with the restriction of the symplectic form σ. Canonical transformations being volume (here: area) preserving it follows that $S(B^{2n}(r)) \cap \mathcal{P}_j$ also has area πr^2. ■

It would certainly be interesting to generalize the first proof to arbitrary canonical transformations, thus yielding a new proof of Gromov's theorem in the general case, in fact, a refinement of it! The difficulty comes from the following fact: the key to the proof in the linear case is the fact that we were able to derive the equality

$$\int_{\gamma_R} p_j dx_j = \pi R^2$$

by exploiting the fact the inverse image of the x_j, p_j plane by S was a plane cutting $B^{2n}(R)$ along a big circle, which thus encloses an area equal to πR^2. When one replaces the *linear* transformation S by a non-linear one, the inverse image of x_j, p_j plane will not generally be a plane, but rather a surface. It turns out that this surface is not quite arbitrary: it is a symplectic 2-dimensional manifold. If the following property holds:

> *The section of $B^{2n}(r)$ by any symplectic surface containing the center of $B^{2n}(r)$ has an area at least πr^2*

then we would have, by the same argument,

$$\int_{\gamma_R} p_j dx_j \geq \pi R^2,$$

hence we would have proved Gromov's theorem in the general case. At the time of the writing of this paper we do not know any proof of this property; nor do we know whether it is true!

We urge the reader to notice that the assumption that we are cutting $S(B^{2n}(r))$ with a plane of *conjugate coordinates* is essential, because it is this assumption that allowed us to identify the area of the section with action. Here is a counterexample which shows that the property does not hold for arbitrary sections of the ellipsoid $S(B^{2n}(r))$. Taking $n = 2$ we define a symplectic matrix

$$S = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & 1/\lambda_1 & 0 \\ 0 & 0 & 0 & 1/\lambda_2 \end{bmatrix} \quad ,$$

where $\lambda_1 > 0$, $\lambda_2 > 0$, and $\lambda_1 \neq \lambda_2$. The set $S(B^{2n}(r))$ is defined by

$$\frac{1}{\lambda_1}x_1^2 + \frac{1}{\lambda_2}x_2^2 + \lambda_1 p_1^2 + \lambda_2 p_2^2 \leq r^2$$

and its section with the x_2, p_2 plane is the ellipse

$$\frac{1}{\lambda_1}x_1^2 + \lambda_1 p_1^2 \leq r^2$$

which has area πr^2 as predicted, but its section with the x_2, p_1 plane is the ellipse

$$\frac{1}{\lambda_1}x_1^2 + \lambda_2 p_2^2 \leq R^2$$

which has area $\pi(r^2\sqrt{\lambda_1/\lambda_2})$ different from πr^2 since $\lambda_1 \neq \lambda_2$.

But why were we mentioning a "symplectic camel" in the title of this section? The reason is metaphoric: Gromov's non-squeezing theorem can be restated by saying that there is no way to deform a phase space ball using canonical transformations in such a way that we can make it pass through a circular hole in a plane of conjugate coordinates x_j, p_j if the area of that hole is smaller than that of the cross-section of that ball: the biblical camel is the ball $B^{2n}(R)$ and the hole in the plane is the eye of the needle!

3.2 The notion of symplectic capacity

Symplectic capacities were introduced by Ekeland and Hofer [14] who elaborated on Gromov's non-squeezing theorem. We refer to Hofer and Zehnder [47], Polterovich [76], or to the recent monograph by Schlenk [82] for detailed studies of the theory of symplectic capacities. McDuff and Salamon [71] also contains some interesting material on this topic.

3.2.1 Definitions and general properties

The following definition is standard, and the most commonly accepted in the literature:

Definition 14 *A "normalized symplectic capacity" on* $(\mathbb{R}^{2n}, \sigma)$ *is a function assigning to every subset* Ω *of* $\mathbb{R}^{2n}$ *a number* $c(\Omega) \geq 0$*, or* $+\infty$*, and having the properties (SC1)–(SC4) listed below:*

(**SC1**) *It must be invariant under canonical transformations (and hence, in particular, under Hamiltonian flows):*

$$c(f(\Omega)) = c(\Omega) \quad \text{if} \quad f \text{ is a canonical transformation;} \tag{30}$$

(**SC2**) *It must be monotone with respect to set inclusion:*

$$c(\Omega) \leq c(\Omega') \quad \textit{if } \Omega \subset \Omega'; \tag{31}$$

(**SC3**) *It must behave like an area under dilations:*

$$c(\lambda\Omega) = \lambda^2 c(\Omega) \quad \textit{for any scalar } \lambda \tag{32}$$

($\lambda\Omega$ *is the set of all points* λz *such that* $z \in \Omega$);
(**SC4**) *It satisfies the normalization conditions*

$$c(B^{2n}(R)) = \pi R^2 = c(Z_j^{2n}(R)) \tag{33}$$

(*recall that is the cylinder based on* x_j, p_j plane *defined by the condition* $x_j^2 + p_j^2 \leq R^2$).

Notice that in condition (33) $Z_j^{2n}(R)$ can be replaced by any cylinder $Z^{2n}(R)$ with radius R based on a symplectic plane (cf. the discussion following the statement of Gromov's theorem).

Obviously symplectic capacities are unbounded (even if the symplectic capacity of an unbounded set can be bounded, cf. property (SC4). We have for instance

$$c(\mathbb{R}^{2n}) = +\infty \tag{34}$$

as immediately follows from the formula $c(B^{2n}(R)) = \pi R^2$. However, if Ω is bounded then all its symplectic capacities are finite: there exist R (perhaps very large) such that a ball $B^{2n}(R)$ contains Ω, and one then concludes using the monotonicity property (SC2) which implies that we have:

$$c(\Omega) \leq c(B^{2n}(R)) = \pi R^2.$$

In fact, more generally, it follows from the monotonicity property and from (SC4) that If $B^{2n}(R) \subset \Omega \subset Z_j^{2n}(R)$ then $c(\Omega) = \pi R^2$; this illustrates the fact that sets very different in shape and volume can have the same symplectic capacity.

Also note that a set Ω with non-empty interior cannot have symplectic capacity equal to zero: let Ω' be the interior of Ω, it is an open set, and if it is not empty it contains a (possibly) very small) ball $B^{2n}(\varepsilon)$. Using again (SC2) we have

$$\pi\varepsilon^2 = c(B^{2n}(\varepsilon)) \leq c(\Omega).$$

We will usually drop the qualification "normalized" in the definition above and just speak about "symplectic capacities"; one exception to this rule will be the Ekeland–Hofer capacities discussed in Subsection 3.2.2 below.

The reader is urged to keep in mind that the notion of symplectic capacity is not directly related to that of volume. For instance, the function c_{Vol} defined by

$$c_{\mathrm{Vol}}(\Omega) = [\mathrm{Vol}(\Omega)]^{1/n}$$

obviously satisfies the properties (SC1)–(SC4) above *except* the identity $c_{\mathrm{Vol}}(Z_j^{2n}(R)) = \pi R^2$ (as soon as $n > 1$) which is precisely the most characteristic and interesting property of a symplectic capacity! It is conjectured (see the discussion in Artstein-Avidan et al. [3]) that for any convex body K in $\mathbb{R}^{2n}$ we have the inequality

$$\frac{c(K)}{\pi} \leq \left(\frac{\mathrm{Vol}(K)}{\pi}\right)^{1/n} \tag{35}$$

with equality only when K is the symplectic image of a ball. These authors actually prove a weaker variant of (35), namely that

$$\frac{c(K)}{\pi} = C\left(\frac{\mathrm{Vol}(K)}{\pi}\right)^{1/n} \tag{36}$$

where C is a universal constant (i.e. independent of the dimension n).

We will also often consider the weaker notion of linear symplectic capacity:

Definition 15 *A "linear symplectic capacity" assigns to every subset* Ω *of* $\mathbb{R}^{2n}$ *a number* $c^{\mathrm{lin}}(\Omega) \geq 0$ *or* $+\infty$, *and having the properties (SC2)–(SC4) above, property (SC1) being replaced by:*

(**SC1Lin**) *A linear symplectic capacity* c^{lin} *is invariant under phase-space translations and under the action of* $\mathrm{Sp}(2n, \mathbb{R})$.

This definition can be restated by saying that a linear symplectic capacity is only invariant under the action of the affine (or: inhomogeneous) symplectic group $\mathrm{ISp}(2n, \mathbb{R})$:

$$c^{\mathrm{lin}}(f(\Omega)) = c(\Omega) \text{ for all } f \in \mathrm{ISp}(2n, \mathbb{R}). \tag{37}$$

(Recall that $\mathrm{ISp}(2n, \mathbb{R})$ consists of all products $ST(z)$ where $S \in \mathrm{Sp}(2n, \mathbb{R})$ and $T(z) : z' \longmapsto z' + z$ is an arbitrary phase-space translation).

We are going to see that as a consequence of the symplectic non-squeezing theorem symplectic capacities do exist (there are actually infinitely many such capacities).

Gromov's non-squeezing theorem allows us to easily construct two symplectic capacities $c_{\min}$ and $c_{\max}$. As the notation suggests, we have

$$c_{\min}(\Omega) \leq c(\Omega) \leq c_{\max}(\Omega) \tag{38}$$

for every symplectic capacity c and every set Ω (we will prove this property below); it follows from these inequalities and the fact that $c_{\min}$ and $c_{\max}$ are not identical that there exist infinitely many symplectic capacities: for every real λ in the closed interval $[0, 1]$ the formula

$$c_\lambda = \lambda c_{\max} + (1 - \lambda) c_{\min} \tag{39}$$

obviously defines a symplectic capacity, and we have $c_\lambda \neq c_{\lambda'}$ if $\lambda \neq \lambda'$. (More generally, we can always interpolate two arbitrary symplectic capacities to obtain new capacities).

By definition, $c_{\min}(\Omega)$ is the (possibly infinite) number $c_{\min}(\Omega)$ calculated as follows: assume that there exits no canonical transformation sending any phase space ball $B^{2n}(r)$ inside Ω, no matter how small its radius r is. We will then write $c_{\min}(\Omega) = 0$. Assume next that there are canonical transformations sending $B^{2n}(r)$ in Ω for some r (and hence also for all $r' < r$). The supremum R of all such radii r is called the *symplectic radius* of Ω and we define

$$c_{\min}(\Omega) = \sup_f \{\pi r^2 : f(B^{2n}(r) \subset \Omega\} = \pi R^2 \tag{40}$$

where f ranges over all the canonical transformations of $\mathbb{R}^{2n}$. Thus $c_{\min}(\Omega) = \pi R^2$ means that one can find canonical transformations sending $B^{2n}(r)$ inside Ω for all $r < R$, but that no canonical transformation will send a ball with radius larger R inside that set. It immediately follows from this definition that $c_{\min}$ satisfies the axioms (SC1)–(SC3). It is also clear that $c_{\min}(B^{2n}(R)) = \pi R^2$: first, the identity trivially maps $B^{2n}(R)$ into itself, so that $c_{\min}(B^{2n}(R)) \leq \pi R^2$. On the other hand, we cannot have $c_{\min}(B^{2n}(R)) > \pi R^2$ because this would mean that there exists a canonical transformation mapping a ball $B^{2n}(R')$ with $R' > R$ inside $B^{2n}(R)$; but this is not possible since canonical transformations are volume preserving. There remains to check the equality $c_{\min}(Z_j^{2n}(R)) = \pi R^2$. The non-squeezing theorem says that there is no way we can squeeze a ball with radius $R' > R$ inside $Z_j^{2n}(R)$; we must thus have $c_{\min}(Z_j^{2n}(R)) \leq \pi R^2$. That we actually have equality is immediate, observing that we can translate the ball $B^{2n}(R)$ inside any cylinder with same radius, and that phase space translations are canonical transformations (because their Jacobian matrices are the identity, which is trivially symplectic). We remark that $c_{\min}(\Omega)$ is sometimes called the symplectic width of the set Ω.

The symplectic capacity $c_{\max}$ is constructed in a similar way: suppose that no matter how large we choose r there exists no canonical transformation sending Ω inside a cylinder $Z_j^{2n}(r)$. We then set $c_{\max}(\Omega) = +\infty$. Suppose that, on the contrary, there are canonical transformations sending Ω inside some cylinder $Z_j^{2n}(r)$ and let R be the infimum of all such r. Thus, by definition,

$$c_{\max}(\Omega) = \inf_f \{\pi r^2 : f(\Omega) \subset Z_j^{2n}(r)\} = \pi R^2 \tag{41}$$

where f again ranges over all the canonical transformations $\mathbb{R}^{2n} \longrightarrow \mathbb{R}^{2n}$. We leave it to the reader to verify, using again the non-squeezing theorem, that $c_{\max}$ indeed is a symplectic capacity.

There remains to prove the inequalities (38). Suppose first $c_{\min}(\Omega) > c(\Omega)$ and set $c(\Omega) = \pi R^2$; thus $c_{\min}(\Omega) > \pi R^2$. It follows, by definition of $c_{\min}$, that there exists $\varepsilon > 0$ and a canonical transformation f such that $f(B^{2n}(R+\varepsilon)) \subset \Omega$. But then, in view of the monotonicity property (SC2) of c we have $c(f(B^{2n}(R+\varepsilon))) \leq c(\Omega)$, that is, in view of the symplectic invariance property (SC1), $c(B^{2n}(R+\varepsilon)) \leq c(\Omega)$. We thus have, taking (SC4) into account, $c(B^{2n}(R+\varepsilon)) = \pi(R+\varepsilon)^2 \leq c(\Omega)$, and this contradicts $c(\Omega) = \pi R^2$. The proof of the inequality $c(\Omega) \leq c_{\max}(\Omega)$ is similar; we leave the details to the reader.

The existence of linear symplectic capacities is proven exactly in the same way as above. In fact, for $\Omega \subset \mathbb{R}^{2n}$ set

$$c_{\min}^{\mathrm{lin}}(\Omega) = \sup_{f \in \mathrm{ISp}(2n,\mathbb{R})} \{\pi R^2 : f(B^{2n}(R)) \subset \Omega\} \tag{42}$$

$$c_{\max}^{\mathrm{lin}}(\Omega) = \inf_{f \in \mathrm{ISp}(2n,\mathbb{R})} \{\pi R^2 : f(\Omega) \subset Z_j^{2n}(R)\}; \tag{43}$$

it is immediate to show that $c_{\min}^{\mathrm{lin}}$ and $c_{\max}^{\mathrm{lin}}$ are linear symplectic capacities, which can be interpreted as follows: for every $\Omega \subset \mathbb{R}^{2n}$ the number $c_{\min}^{\mathrm{lin}}(\Omega)$ (which can be infinite) is the supremum of all the πR^2 of phase space balls $B^{2n}(R)$ that can be sent in Ω using elements of $\mathrm{ISp}(2n,\mathbb{R})$; similarly $c_{\min}^{\mathrm{lin}}(\Omega)$ is the infimum of all πR^2 such that a cylinder $Z_j^{2n}(R)$ can contain the deformation of Ω by elements of the inhomogeneous symplectic group $\mathrm{ISp}(2n,\mathbb{R})$ (the group generated by phase space translations and the elements of $\mathrm{Sp}(2n,\mathbb{R})$). We have moreover

$$c_{\min}^{\mathrm{lin}}(\Omega) \leq c^{\mathrm{lin}}(\Omega) \leq c_{\max}^{\mathrm{lin}}(\Omega) \tag{44}$$

for every $\Omega \subset \mathbb{R}^{2n}$ and every linear symplectic capacity c_{lin}; the proof is similar to that of the inequalities (38).

The homogeneity property (SC2) satisfied by every symplectic capacity (linear or not) together with the fact that $c(B^{2n}(R)) = \pi R^2$ suggests that symplectic capacities have something to do with the notion of area. In fact, the following is true: the symplectic capacity $c_{\min}(\Omega)$ of a subset in the phase plane $\mathbb{R}^2$ is the area of Ω when the latter is connected (Siburg [86]; also see the proof in Hofer and Zehnder [47], §3.5, Theorem 4). Note that the result in general no longer holds when Ω is disconnected: suppose for instance that Ω is the union of two disjoint disks with radii R and R' such that $R' < R$. Then $c_{\min}(\Omega) = \pi R^2 < \text{Area}(\Omega)$. The symplectic capacity $c_{\max}(\Omega)$ is the area when Ω is simply connected. Summarizing, it follows from the inequalities (38) that:

Proposition 16 *Let c be a symplectic capacity on the phase plane $\mathbb{R}^2$. We have $c(\Omega) = \text{Area}(\Omega)$ when Ω is a connected and simply connected surface.*

The reader may easily convince himself that $c_{\min}(\Omega)$ is not the area when Ω is disconnected, and that $c_{\max}(\Omega)$ is in general not the area when Ω fails to be simply connected (a typical counterexample is the annulus $r \leq x^2 + p^2 \leq R^2$).

3.2.2 The Hofer–Zehnder capacity

In [47] (Chapter 3) Hofer–Zehnder construct a symplectic capacity c^{HZ} which measures sets in a dynamical way. It is defined as follows. Let Ω be an open set in $\mathbb{R}^{2n}$ and consider the class $\mathcal{H}(\Omega)$ of all Hamiltonians functions $H \geq 0$ having the following three properties:

- H vanishes outside Ω (and is hence H bounded);
- The critical values of H are 0 and $\max H$;
- The flow (f_t^H) has no constant periodic orbit with period $T \leq 1$.

Then, by definition,

$$c^{\text{HZ}}(\Omega) = \sup\{\max H : H \in \mathcal{H}(\Omega)\} \tag{45}$$

The Hofer–Zehnder capacity has the property that whenever Ω is a compact convex set in phase space then

$$c^{\text{HZ}}(\Omega) = \oint_{\gamma_{\min}} pdx \tag{46}$$

where $pdx = p_1 dx_1 + \cdots + p_n dx_n$ and $\gamma_{\min}$ is the shortest (positively oriented) Hamiltonian periodic orbit carried by the boundary $\partial\Omega$ of Ω. (For *Eqn.* (46) the condition that Ω be compact and convex is essential (see Hofer and Zehnder's very illustrative "Bordeaux bottle" example in [47], p. 99).

Notice that this formula generalizes the observation made earlier in this section that symplectic capacities agree with the usual notion of area in the case $n = 1$ for connected and simply connected surfaces. In fact, let Ω be such a surface in the phase plane, and assume that the boundary $\gamma = \partial\Omega$ is smooth, and given the positive orientation. We then have, by Stoke's theorem,

$$\operatorname{Area}(\Omega) = \tfrac{1}{2} \oint_\gamma (pdx - xdp) = \oint_\gamma pdx.$$

Formula (46) in particular implies the inequalities

$$c_{\min}(\Omega) \leq c^{\mathrm{HZ}}(\Omega) \leq \left| \oint_\gamma pdx \right| \tag{47}$$

for every periodic orbitγ on $\partial\Omega$.

For *Eqn.* (46) to be unambiguous, one has of course to justify two (somewhat related) points:

- The first is that there indeed exists periodic Hamiltonian orbits on $\partial\Omega$; in the present situation this is actually a very general result, going back to the early history of the theory of Hamiltonian periodic orbits: it turns out that if an energy shell bounds a convex and compact subset of phase space, then it automatically carries at least one periodic orbit (see e.g. [47]);

- The second point, which is apparently more obscure, is that we have to show that the integral in the right-hand side of (46) is independent of the choice of the Hamiltonian defining the periodic orbit! This is actually a consequence of our discussion of Hamiltonian orbits in Subsection 2.3.2, which applies to periodic orbits as well.

In [74] Narcowich notes that when studying Wigner ellipsoids associated with a covariance matrix (see Definition 27 in Section 4) area is the adequate measure of phase-space concentration only in the case of one degree of freedom; he shows that when $n > 1$ one should *not* replace area with *volume*, but rather with the *action integral* of a periodic Hamiltonian orbit carried by the boundary of the Wigner ellipsoid. Narcowich's observation indeed goes straight to the point: as we will see, all symplectic capacities agree on ellipsoids, hence the action integral is precisely the symplectic capacity of the Wigner ellipsoid in view of formula (46)!

3.3 Phase space ellipsoids and polydisks

3.3.1 Some embedding results

A basic problem in symplectic topology is the following: given to open sets[5] Ω and Ω' in $(\mathbb{R}^{2n}, \sigma)$ when does there exist a canonical transformation sending Ω in Ω'? We discuss here a few known results, and briefly mention some conjectures.

Consider two phase space ellipsoids

$$\Omega_{M,z_0} : M(z-z_0)^2 \leq 1 \quad , \quad \Omega_{M',z_0'} : M'(z-z_0')^2 \leq 1;$$

the matrices M and M' are symmetric and positive-definite. We ask the following question:

> *Under which conditions on M and M' can we find a canonical transformation sending Ω_{M,z_0}* in $\Omega_{M',z_0'}$ *?*

Clearly the answer to this question is independent of the centers z_0 and z_0': if f is a canonical transformation such that $f(\Omega_{M,z_0}) \subset \Omega_{M',z_0'}$ then $T(z_0') \circ f \circ T(z_0')$ sends $\Omega_M = \Omega_{M,0}$ in $\Omega_{M'} = \Omega_{M',0}$ hence it is sufficient to answer the question above for the ellipsoids

$$\Omega_M : Mz^2 \leq 1 \quad , \quad \Omega_{M'} : M'z^2 \leq 1.$$

We begin by proving a necessary and sufficient condition in the linear case:

Proposition 17 *There exists $S \in \mathrm{Sp}(2n, \mathbb{R})$ such that $S(\Omega_M) \subset \Omega_{M'}$ if and only if* $\mathrm{Spec}_\sigma(M) \geq \mathrm{Spec}_\sigma(M')$*.*

Proof. Assume that there exists a symplectic matrix S sending Ω_M in $\Omega_{M'}$; this is the same thing as saying that if $Mz \cdot z \leq 1$ then $M'(Sz) \cdot (Sz) \leq 1$, that is $S^T M' S z \cdot z \leq 1$. By homogeneity this is equivalent to $S^T M' S z \cdot z \leq Mz \cdot z$ for all $z \in \mathbb{R}^{2n}$, which is in turn equivalent to $S^T M' S \leq M$. In view of the properties (i) and (iii) of the symplectic spectrum listed in Proposition 9 this is equivalent to the $\mathrm{Spec}_\sigma(M) \geq \mathrm{Spec}_\sigma(M')$. ■

[5] Or more generally, symplectic manifolds.

What about the non-linear case? One could very well imagine that the sufficient and necessary condition $\mathrm{Spec}_\sigma(M) \geq \mathrm{Spec}_\sigma(M')$ for a symplectic imbedding of Ω_M in $\Omega_{M'}$ can be weakened if one uses arbitrary canonical transformations. However, this is not the case; see for instance Schlenk [82, 81] for counterexamples. In particular Schlenk shows, using higher-order symplectic capacities, that if $\mathrm{Spec}_\sigma(M) = (\lambda_1^\sigma, ..., \lambda_n^\sigma)$ is such that $\lambda_1^\sigma \leq 2\lambda_n^\sigma$ then one cannot find a canonical transformation sending a ball $B^{2n}(R)$ in Ω_M when $R < 1/\sqrt{\lambda_n^\sigma}$. In the general nonlinear case there are additional constraints which are not still completely known; see [82] for a careful study of the situation.

Let us now discuss the case of polydisks; it turns out that for these sets surprisingly little is known.

A polydisk is a product

$$\mathbb{D}^{2n}(R) = B_1^2(R_1) \times B_2^2(R_2) \times \cdots \times B_n^2(R_n) \tag{48}$$

where $B_j^2(R_j)$ is the disk $x_j^2 + p_j^2 \leq R_j^2$ in the x_j, p_j plane and R is the n-tuple $(R_1, R_2, ..., R_n)$; we will make the convention that $R_1 \leq R_2 \leq \cdots \leq R_n$. Assume that there exists a canonical transformation sending $\mathbb{D}^{2n}(R)$ into $\mathbb{D}^{2n}(R')$. Since canonical transformations are a fortiori volume preserving, we must have $\mathrm{Vol}(\mathbb{D}^{2n}(R)) \leq \mathrm{Vol}(\mathbb{D}^{2n}(R'))$ and hence $R_1 R_2 \cdots R_n \leq R_1' R_2' ... R_n'$. On the other hand, we must also have $c(\mathbb{D}^{2n}(R)) \leq c(\mathbb{D}^{2n}(R'))$ for any symplectic capacity c; since $c_{\min}(\mathbb{D}^n(R)) = \pi R_1^2$ (see *Eqn.* (63) below), and this implies that we must have $R_1 \leq R_1'$. Thus both volume and symplectic capacity can be used as obstructions for finding embeddings. Surprisingly enough, Guth has proved in [44] that these two quantities are the only obstructions. More precisely, Guth shows that:

Proposition 18 *For each integer $n \geq 1$ there exists a constant $C(n) > 0$ such that if*

$$C(n) R_1 R_2 \cdots R_n \leq R_1' R_2' \cdots R_n' \quad \textit{and} \quad C(n) R_1 \leq R_1' \tag{49}$$

then there exists a canonical transformation f such that $f(\mathbb{D}^{2n}(R)) \subset \mathbb{D}^{2n}(R')$.

In the same paper, Guth discusses a possible physical consequence of this result, the "catalyst map"; we will describe this map (which raises many interesting question) in the concluding section.

In [3] Artstein-Avidan et al. prove the following linear embedding result for convex symmetric bodies. Recall that a subset Ω of a vector space is called a convex body if it is compact and has non-empty interior; Ω is symmetric if $z \in \Omega$ if and only if $-z \in \Omega$ (in which case Ω is centered at 0). A supporting hyperplane of a convex body Ω is a hyperplane that intersects the boundary $\partial\Omega$, but not the interior of Ω.

Proposition 19 *Let Ω be a symmetric convex body in $\mathbb{R}^{2n}$ having the rotational invariance property $J\Omega = \Omega$ (that is $(x,p) \in \Omega$ if and only if $(-p,x) \in \Omega$. Then $c_{\max}^{\text{lin}}(\Omega) \geq 2\pi R^2$ where $R = \sup\{r : B^{2n}(r) \subset \Omega\}$.*

Proof. Assume that $B^{2n}(R)$ is the largest ball contained in Ω; we are going to show that Ω is then contained in a cylinder $Z^{2n}(R\sqrt{2})$ based on a symplectic plane; it will follow that

$$c_{\max}^{\text{lin}}(\Omega) \geq c_{\max}^{\text{lin}}(Z^{2n}(R\sqrt{2})) = 2\pi R^2.$$

Since Ω is symmetric there are at least two contact points z and $-z$ which belong to both the boundary $\partial\Omega$ and to the circle $S^{2n-1}(r) = \partial B^{2n}(r)$. The supporting hyperplanes to Ω at these points are the affine planes $\{\pm z\} + z^{\perp}$ where $z^{\perp}$ is the orthogonal space to the vector z. It follows that Ω lies between these two hyperplanes. Since $J\Omega = \Omega$ the points $\pm Jz$ are contact points as well, so Ω also lies between the hyperplanes $\{\pm Jz\} + (Jz)^{\perp}$. The length of the vectors z and Jz is R hence the orthogonal projection of Ω on the symplectic plane $\mathcal{P}$ spanned by z and Jz is contained in a square in a square whose edges have length $2R$, which is in turn contained in a disk of radius $R\sqrt{2}$. It follows that Ω is contained in a cylinder with radius $R\sqrt{2}$ based on a symplectic plane, therefore $c_{\max}^{\text{lin}}(\Omega) \leq \pi(R\sqrt{2})^2$ as claimed. ∎

A related topic is that of embedding of cubes; it is ironic that so little actually is known for such simple geometric loci. Assume first that $Q^{2n}(r)$ is the cube $[0,r]^n$ whose edges have length r and are all parallel to the coordinate axis. We denote this cube by $Q^{2n}_{//}(r)$. One proves, using the Ekeland–Hofer capacity c_n^{EH} that:

- There exists a canonical transformation sending $Q^{2n}_{//}(r)$ in the ball $B^{2n}(r)$ if and only if $R \geq \frac{1}{n}\pi r^2$;
- There exists a canonical transformation sending $Q^{2n}_{//}(r)$ in the cylinder $Z^{2n}_j(r)$ if and only if $r \geq \pi R^2$.

It follows that we have

$$c_{\min}(Q^{2n}_{//}(r)) = \frac{1}{n}\pi r^2 \text{ and } c_{\max}(Q^{2n}_{//}(r)) = \pi r^2. \tag{50}$$

For cubes in general position very little is known. Applying Proposition 19 above yields, for an arbitrary cube $Q^{2n}(r)$ with edge length r, the estimate $c_{\max}^{\text{lin}}(Q^{2n}(r)) \geq \frac{\pi}{2}r^2$. Together with the conjectured inequality (35) for convex bodies, this leads to the possible estimate

$$\frac{\pi}{2}r^2 \leq c_{\max}^{\text{lin}}(Q^{2n}(r)) \leq 2\pi r^2 \tag{51}$$

for general cubes. An exact formula still remains to be discovered.

3.3.2 Symplectic capacities of ellipsoids and polydisks

The following result is very important; it shows that all symplectic capacities (linear, or not) agree on phase space ellipsoids. It is actually one of the rare cases where one knows how to perform an explicit calculation of a symplectic capacity!

Proposition 20 *Let M be a $2n\times2n$ positive-definite matrix M and $\lambda_1^\sigma, ..., \lambda_n^\sigma$ its symplectic eigenvalues. Consider the ellipsoid:*

$$\Omega_{M,z_0} : M(z-z_0)^2 \leq 1. \tag{52}$$

(i) For every symplectic capacity c and every linear symplectic capacity c^{lin} on $(\mathbb{R}^{2n}, \sigma)$ we have

$$c(\Omega_{M,z_0}) = c^{lin}(\Omega_{M,z_0}) = \frac{\pi}{\lambda_{\max}^\sigma} \tag{53}$$

where $\lambda_{\max}^\sigma = \lambda_1^\sigma$ is the largest symplectic eigenvalue of M.

(ii) In particular, if A and B are real symmetric $n \times n$ matrices, then the symplectic capacity of the ellipsoid

$$\Omega_{(A,B)} : Ax^2 + Bp^2 \leq 1 \tag{54}$$

is given by

$$c(\Omega_{(A,B)}) = c^{lin}(\Omega_{(A,B)}) = \frac{\pi}{\sqrt{\lambda_{\max}}}. \tag{55}$$

where $\lambda_{\max}$ is the largest eigenvalue of AB.

Proof. Since phase space translations are canonical transformations, we may moreover assume that $z_0 = 0$ and thus take $\Omega_{M,z_0} = \Omega_M$. The statement (ii) follows from (i) in view of Lemma 34, choosing for M the block-diagonal matrix $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$. To prove (i) we first note that in view of Williamson's theorem there exists a symplectic matrix S such that

$$S^T M S = \begin{bmatrix} \Lambda^\sigma & 0 \\ 0 & \Lambda^\sigma \end{bmatrix} \text{ with } \Lambda^\sigma = \operatorname{diag}\left[\lambda_1^\sigma, ..., \lambda_n^\sigma\right]. \tag{56}$$

Since symplectic capacities are invariant by canonical transformations it follows that $c(\Omega_M) = c(S(\Omega_M))$ hence it suffices to prove formula (53) when Ω_M is replaced by $S(\Omega_M)$. We may thus assume that

$$\Omega_M \,:\, \sum_{j=1}^{n} \frac{1}{R_j^2}(x_j^2 + p_j^2) \leq 1 \tag{57}$$

where we have set $\lambda_j^\sigma = 1/R_j^2$. Suppose now that there exists a canonical transformation f sending a ball $B^{2n}(R)$ inside Ω_M. Then $f(B^{2n}(R))$ is also contained in each cylinder $Z_j^{2n}(R) : x_j^2 + p_j^2 \leq R^2$ and hence $R \leq R_{\text{min}} = \sqrt{1/\lambda_{\text{max}}^\sigma}$ in view of the non-squeezing theorem. It follows that $c_{\text{min}}(\Omega_M) \leq \pi R_{\text{min}}^2 = \pi/\lambda_{\text{max}}^\sigma$; since on the other hand $B^{2n}(R_{\text{max}})$ is anyway contained in Ω_M we must have equality: $c_{\text{min}}(\Omega_M) = \pi/\lambda_{\text{max}}^\sigma$. A similar argument shows that we also have $c_{\text{max}}(\Omega_M) = \pi/\lambda_{\text{max}}^\sigma$; formula (53) follows since c_{min} and c_{max} are the smallest and largest symplectic capacities. That we also have $c^{\text{lin}}(\Omega_{M,z_0}) = \pi/\lambda_{\text{max}}^\sigma$ follows from the proof, replacing c_{min} by $c_{\text{min}}^{\text{lin}}$ and c_{max} by $c_{\text{max}}^{\text{lin}}$. ■

Since all symplectic capacities agree on ellipsoids we have in particular

$$c^{\text{HZ}}(\Omega_M) = \frac{\pi}{\lambda_{\text{max}}^\sigma} = \oint_{\gamma_{\text{min}}} p dx \tag{58}$$

where γ_{min} is the positively oriented ellipse, intersection of Ω_M with the x_1, p_1 plane; $c^{\text{HZ}}(\Omega_{M,z_0})$ is thus the area of that ellipse and it follows that the intersection of Ω_M with any x_j, p_j plane is at least $c^{\text{HZ}}(\Omega_{M,z_0})$.

3.3.3 The dual ellipsoid

Let Q be a positive definite quadratic form on $\mathbb{R}^{2n}$ and consider the real function f_ζ defined for $\zeta \in \mathbb{R}^{2n}$ by $f_\zeta(z) = z \cdot \zeta - Q(z)$. That function has a unique critical value, which we denote by $Q^*(\zeta)$. The function Q^* thus defined is the *Legendre transform* of Q; it is also a positive definite quadratic form[6]. In fact, a straightforward calculation shows that if $Q(z) = Mz^2$

[6] Q^* is *stricto sensu* defined on the dual space of $\mathbb{R}^{2n}$ but we will gladly ignore this technicality here.

where M is a positive-definite symmetric matrix then $Q^*(\zeta) = \frac{1}{4}M^{-1}\zeta^2$. In particular if Q is the Hessian matrix of Q then the Hessian matrix of Q^* is the inverse Q^{-1}. We thus have the useful formula:

$$Q(z) = \tfrac{1}{2}Qz \cdot z \Longleftrightarrow Q^*(\zeta) = \tfrac{1}{2}Q^{-1}\zeta \cdot \zeta. \tag{59}$$

Definition 21 *Let Ω_M be the phase-space ellipsoid defined by $Q(z) \leq 1$ where Q is a positive-definite quadratic form on $\mathbb{R}^{2n}$. The dual ellipsoid Ω_M^* of Ω_M is defined by $Q^*(\zeta) \leq 1$ where Q^* is the Legendre transform of Q.*

Note that it immediately follows from (59) that we have $(\Omega_M^*)^* = \Omega_M$ (on a more conceptual level it follows from the fact that the Legendre transformation is involutive).

Remark 22 *Recently Artstein-Avidan and Milman characterized duality transforms in [4]. Among other things they showed that every involutive transform on the class of lower semi-continuous functions that is order reversing must be (up to linear terms) be the Legendre transform.*

The following result relates the symplectic capacity of an ellipsoid to that of the dual ellipsoid:

Proposition 23 *For $M > 0$ let $\Omega_M : Mz^2 \leq 1$ and $\Omega_M^* : \frac{1}{4}M^{-1}\zeta^2 \leq 1$ be dual phase-space ellipsoids. Let $\operatorname{Spec}_\sigma(M) = (\lambda_1^\sigma, ..., \lambda_n^\sigma)$ be the symplectic spectrum of M. We have*

$$c(\Omega_M^*) = \frac{\pi}{4}\lambda_{\min}^\sigma \tag{60}$$

where $\lambda_{\min}^\sigma$ is the smallest symplectic eigenvalue of M.

Proof. The dual ellipsoid Ω_M^* is determined by $2M^{-1}\zeta^2 \leq 1$. In view of Proposition 9(ii) we have

$$\operatorname{Spec}_\sigma((4M)^{-1}) = ((4\lambda_n^\sigma)^{-1}, ..., (4\lambda_1^\sigma)^{-1})$$

hence, applying formula (53) for the symplectic capacity of an ellipsoid $c(\Omega_M^*) = \pi/4(\lambda_n^\sigma)^{-1}$ which is precisely formula (60) since λ_n^σ is the smallest symplectic eigenvalue of M. ∎

Set now $R = (R_1, R_2, ..., R_n)$ where $0 < R_1 < R_2 < \cdots < R_n$ and consider the torus

$$\mathbb{T}^n(R) = S_1^1(R_1) \times S_2^1(R_2) \times \cdots \times S_n^1(R_n) \tag{61}$$

where $S^1_j(R_j)$ is the circle $x_j^2 + p_j^2 = R_j^2$ (such tori play a crucial role in the study of integrable Hamiltonian systems; see Section 6). The torus $\mathbb{T}^n(R)$ is contained in the boundary of the polydisk

$$\mathbb{D}^{2n}(R) = B_1^2(R_1) \times B_2^2(R_2) \times \cdots \times B_n^2(R_n). \tag{62}$$

We have

$$c_{\min}(\mathbb{D}^n(R)) = \pi R_{\min}^2 \tag{63}$$

where $R_{\min} = R_1$ (this follows from the non-squeezing theorem 12; see Polterovich [76] Theorem 1.1C and §4.3).

3.3.4 The Ekeland–Hofer capacities

The normalized symplectic capacities we have been using sofar do not generally allow to distinguish between ellipsoids: formula (53) implies that if $\Omega_M : Mz^2 \leq 1$ and $\Omega_{M'} : M'z^2 \leq 1$ are such that M and M' have the same smallest symplectic eigenvalue then $c(\Omega_M) = c(\Omega_{M'})$. This can however be achieved by introducing a slightly more general notion of symplectic capacity. To do this we slightly relax the normalization condition (SC4) for symplectic capacities and replace it with the simpler non-triviality requirement:

(SC4bis) $c(B^{2n}(R)) > 0$ *and* $c(Z_j(R)) < \infty$.

We will call a mapping $\Omega \longmapsto c(\Omega)$ a *generalized symplectic capacity* if it satisfies properties (SC1), (SC2), (SC3), (SC4bis).

In [14] Ekeland and Hofer construct a sequence of generalized symplectic capacities c_k^{EH} having the following properties:

(**EH1**) *The sequence* $(c_k^{\mathrm{EH}})_{k\geq 1}$ is *increasing:*

$$c_1^{\mathrm{EH}}(\Omega) \leq c_2^{\mathrm{EH}}(\Omega) \leq \cdots \leq c_k^{\mathrm{EH}}(\Omega) \leq \cdots \tag{64}$$

for all $\Omega \subset \mathbb{R}^{2n}$;

(**EH2**) *If* Ω *is convex with boundary* $\partial\Omega$, *then*

$$c_k^{\mathrm{EH}}(\Omega) = c_k^{\mathrm{EH}}(\partial\Omega) \tag{65}$$

(hence $c_k^{\mathrm{EH}}(\Omega)$ *is determined by the boundary* $\partial\Omega$*);*

(**EH3**) *If* Ω *is convex, then*

$$c_1^{\mathrm{EH}}(\Omega) = c^{\mathrm{HZ}}(\Omega) \text{ and } c_k^{\mathrm{EH}}(\Omega) = \oint_\gamma p\,dx \tag{66}$$

where γ is a periodic Hamiltonian orbit carried by $\partial\Omega$.

The values of the capacities c_k^{EH} on balls and cylinders are given by

$$c_k^{\mathrm{EH}}(B^{2n}(r)) = \left[\frac{k+n-1}{n}\right]\pi r^2 \tag{67}$$

$$c_k^{\mathrm{EH}}(Z_j^{2n}(r)) = \pi r^2; \tag{68}$$

in the first formula $[x]$ is the integer part of $x \in \mathbb{R}$.

The Ekeland–Hofer capacities c_k^{EH} allow us to classify phase-space ellipsoids. In fact, it readily follows from their properties (EH1) and (EH3) that the non-decreasing sequence of numbers $c_k^{\mathrm{EH}}(\Omega_M)$ is determined as follows: if $\mathrm{Spec}_\sigma(M) = (\lambda_1^\sigma, ..., \lambda_n^\sigma)$ write the numbers $k\pi/\lambda_1^\sigma$ in increasing order *with repetition* if a number occurs several times; we thus obtain a sequence $c_1 \leq c_2 \leq \cdots$ and we have

$$c_k^{\mathrm{EH}}(\Omega_M) = c_k. \tag{69}$$

We said above that the Ekeland–Hofer capacities c_k^{EH} allow to distinguish between ellipsoids.

It follows from the considerations above that:

Proposition 24 *An ellipsoid $\Omega_M : Mz^2 \leq 1$ is uniquely determined (up to a symplectic transformation) by the sequence of its Ekeland–Hofer capacities $c_k^{EH}(\Omega_M)$.*

Proof. (See e.g. Cieliebak et al. [10]). Suppose that Ω_M and $\Omega_{M'}$ are two ellipsoids with $\lambda_j^\sigma = \lambda_j'^\sigma$ for $1 \leq j < k$ and $\lambda_k^\sigma > \lambda_k'^\sigma$. then the multiplicity of λ_k^σ in the sequence of Ekeland–Hofer capacities is one higher for Ω_M than for $\Omega_{M'}$ hence not all of these capacities agree on Ω_M and $\Omega_{M'}$. ∎

We saw above that the Ekeland–Hofer capacities allow to distinguish between phase space ellipsoids. What do the c_k^{EH} tell us about polydisks? In [14] Ekeland and Hofer prove that

$$c_k^{\mathrm{EH}}(\mathbb{D}^{2n}(R)) = k\pi R_1^2. \tag{70}$$

Since the boundary of $\mathbb{D}^{2n}(R)$ is the torus $\mathbb{T}^{2n}(R)$ property (EH2) of c_k^{EH} implies that we also have

$$c_k^{\mathrm{EH}}(\mathbb{T}^{2n}(R)) = k\pi R_1^2. \tag{71}$$

Thus, the capacities c_k^{EH} only "see" the smallest disk $B_1^2(R_1)$ in $\mathbb{D}^{2n}(R)$.

4 Two Uncertainty Principles

Uncertainty principles are generally notoriously difficult to interpret properly. One of the most hackneyed lines is that it is impossible to obtain useful knowledge of the simultaneous values of two non-commuting observables; for instance Heisenberg's inequality $\Delta x_j \Delta p_j \gtrsim h$ is supposed to tell us that there is a lower *universal* bound for the product of conjugate position and momentum uncertainties. But how should these "uncertainties" be understood? Are they of a purely statistical nature, as the textbook derivation of the Heisenberg inequalities seems to suggest? Or does this principle say that "God throws dices" in the sense that there is no intrinsic limitation on the accuracy with which we can simultaneously measure x_j and p_j, but that these values are subject to fluctuations due to pure stochastic randomness? There are other possible "interpretations" (see Wick's lovely discussion in [93], Chapter 16). See the "Research Tutorial" by Folland and Sitaram [20] for the point of view of signal analysts; that paper contains in addition numerous historical references, but ignores properties of symplectic invariance). Such questions are also discussed in detail in the recent work [9] by Busch et al.

We will put a particular emphasis on two types of uncertainty principles:

- The Robertson–Schrödinger uncertainty principle

$$(\Delta X_j)^2(\Delta P_j)^2 \geq \Delta(X_j, P_j)^2 + \tfrac{1}{4}\hbar^2 \ , \ 1 \leq j \leq n \tag{72}$$

  where $\Delta(X_j, P_j)$ (the covariance) is a term taking into account the correlations; the inequalities above implies the textbook Heisenberg uncertainty inequalities $\Delta X_j \Delta P_j \geq \frac{1}{4}\hbar^2$. As Luo and Zhang observe in [67] the difference between the Robertson–Schrödinger uncertainty relations and Heisenberg's inequalities is fundamental: the contribution due to the correlation, expressed in terms of covariance, should be put at least on equal footing with that due to the incompatibility of the variables, expressed by the commutators $[X_j, P_j] = i\hbar$. In our case, we will see that it is precisely the use of covariances which will allow us to restate the uncertainty principle in the very lapidary form

$$c(W_\Sigma) \geq \tfrac{1}{2}h \tag{73}$$

  where W_Σ is a phase space ellipsoid associated to the covariances ("Wigner ellipsoid") and c a *symplectic capacity*. We will also see that the inequality above makes sense, under certain assumptions, in classical mechanics.

- Hardy's uncertainty principle, which is a precise statement of the folk meta-theorem "a function and its Fourier transform cannot be simultaneously sharply concentrated". We will actually prove a multi-dimensional version of Hardy's result; as a consequence of this result we will see that if a square integrable function and its $\hbar$-Fourier transform are such that $|\psi(x)| \leq Ce^{-\frac{1}{2\hbar}Ax^2}$ and $|F^{\hbar}\psi(p)| \leq Ce^{-\frac{1}{2\hbar}Bp^2}$ for some constant C and symmetric matrices A, B, then the symplectic capacity of the phase space ellipsoid $\Omega_{A,B} : Ax^2 + Bp^2 \leq \hbar$ then the following analogue of the inequality (73) must hold:

$$c(\Omega_{A,B}) \geq \tfrac{1}{2}h.$$

- The condition $c(\Omega) \geq \frac{1}{2}h$ underlying all the uncertainty principles we are considering, this suggests that the limiting case $c(\Omega) = \frac{1}{2}h$ should put a kind of lower bound on the "symplectic size" of quantum-mechanically admissible phase space subsets. This leads us to replace the usual coarse-graining of phase space by cubic cells with volume h^n by "quantum blobs" $\mathcal{B}$, which are by definition the images of balls with radius $\sqrt{\hbar}$ by canonical transformations. It follows from the properties of symplectic capacities that we always have $c(\mathcal{B}) = \frac{1}{2}h$. We show that when a quantum blob is linear (i.e. obtained by only using linear or affine symplectic transformations) one can associate to it a unique phase-space Gaussian, which is the Wigner transform of a Gaussian quantum state. We exploit this possibility by proving various positivity and averaging results, which a posteriori justify the coarse-graining by quantum blobs.

4.1 The Robertson–Schrödinger uncertainty relations

The textbook Heisenberg inequalities $\Delta P_j \Delta X_j \geq \frac{1}{2}\hbar$ are a weak form (in fact, a consequence) of the Schrödinger–Robertson uncertainty relations, which have the property of being invariant under the action of symplectic transformations (i.e. linear canonical transformations). It turns out that these relations have a precise meaning even in the classical world.

4.1.1 The covariance matrix of a quasi-distribution

In what follows ρ is a real valued-function defined on $\mathbb{R}^{2n}$ and such that

$$\int_{\mathbb{R}^{2n}} \rho(z)dz = 1 \tag{74}$$

and

$$\int_{\mathbb{R}^{2n}}(1+|z|^2)|\rho(z)|dz<\infty. \tag{75}$$

We do not assume that $\rho\geq 0$ so ρ is not in general a true probability density; typically ρ is the Wigner transform of a mixed quantum state, so we will call ρ a "quasi-distribution", even if nothing is "quantum" in the discussion that follows. Notice that condition (75) ensures us that the Fourier transform $F^{\hbar}\rho$ is twice continuously differentiable; this property will be used later on when we discuss positivity properties.

Introducing the notation $z_j=x_j$ if $1\leq j\leq n$ and $z_j=p_{j-n}$ if $n+1\leq j\leq 2n$ we define the covariances

$$\Delta(Z_j,Z_k)=\int_{\mathbb{R}^{2n}}(z_j-\langle z_j\rangle)(z_k-\langle z_k\rangle)\rho(z)dz \tag{76}$$

and

$$(\Delta Z_j)^2=\Delta(Z_j,Z_j)=\int_{\mathbb{R}^{2n}}(z_j-\langle z_j\rangle)^2\rho(z)dz \tag{77}$$

where the moments $\langle z_j^k\rangle$ are the averages given by

$$\langle z_j^k\rangle=\int_{\mathbb{R}^{2n}}z_j^k\rho(z)dz. \tag{78}$$

The quantities $Z_1,Z_2,...,Z_{2n}$ should be viewed as random variables whose values are the phase-space coordinates $z_1,z_2,...,z_n$. Since the integral of ρ is equal to one, formulae (76) and 77) can be rewritten in the familiar form

$$\Delta(Z_j,Z_k)=\langle z_jz_k\rangle-\langle z_j\rangle\langle z_k\rangle \tag{79}$$

$$(\Delta Z_j)^2=\Delta(Z_j,Z_j)=\langle z_j^2\rangle-\langle z_j\rangle^2. \tag{80}$$

The quantities (76), (77), and (78) are well-defined in view of condition (75): the integrals above are all absolutely convergent in view of the trivial estimates

$$\left|\int_{\mathbb{R}^{2n}}z_j\rho(z)dz\right|\leq\int_{\mathbb{R}^{2n}}(1+|z|^2)|\rho(z)|dz<\infty$$

$$\left|\int_{\mathbb{R}^{2n}}z_jz_k\rho(z)dz\right|\leq\int_{\mathbb{R}^{2n}}(1+|z|^2)|\rho(z)|dz<\infty.$$

Definition 25 *We will call the symmetric $2n\times 2n$ matrix*

$$\Sigma=\left[\Delta(Z_j,Z_k)\right]_{1\leq j,k\leq 2n}$$

the covariance matrix associated with ρ. When $\det\Sigma \neq 0$ the inverse Σ^{-1} is called the precision matrix[7] *(or: information matrix).*

For instance, when $n = 1$, the covariance matrix is

$$\Sigma = \begin{bmatrix} \Delta X^2 & \Delta(X,P) \\ \Delta(P,X) & \Delta P^2 \end{bmatrix}$$

where, by definition, the variances ΔX^2 and ΔP^2 and the covariances $\Delta(X,P) = \Delta(P,X)$ are the quantities

$$\begin{aligned} \Delta X^2 &= \iint_{\mathbb{R}^2} x^2\rho(x,p)dpdx - \left(\iint_{\mathbb{R}^2} x\rho(x,p)dpdx\right)^2 \\ \Delta P^2 &= \iint_{\mathbb{R}^2} p^2\rho(x,p)dpdx - \left(\iint_{\mathbb{R}^2} p\rho(x,p)dpdx\right)^2 \\ \Delta(X,P) &= \iint_{\mathbb{R}^2} xp\rho(z)dpdx - \iint_{\mathbb{R}^2} x\rho(z)dpdx \iint_{\mathbb{R}^2} p\rho(z)dpdx. \end{aligned}$$

When the averages (first moments) vanish these equations become

$$\Delta X^2 = \iint_{\mathbb{R}^2} x^2\rho(x,p)dpdx \quad , \quad \Delta P^2 = \iint_{\mathbb{R}^2} p^2\rho(x,p)dpdx \tag{81}$$

and

$$\Delta(X,P) = \iint_{\mathbb{R}^2} xp\rho(z)dpdx - \iint_{\mathbb{R}^2} x\rho(z)dpdx \iint_{\mathbb{R}^2} p\rho(z)dpdx. \tag{82}$$

The following result is essential. The first statement (i) was apparently first noted by Narcowich in [74] (Lemma 2.3), and part (ii) goes back to Narcowich [72], Narcowich and O'Connell [75], Yuen [96], Simon et al. [88, 89], and implicitly in Lindblad [60]. The third part (iii) is a way of expressing the symplectic covariance of the uncertainty principle.

Theorem 26 *Let Σ be a real symmetric $2n \times 2n$ matrix and η a real number. Then $\Sigma + i\eta J$ is a Hermitian matrix. Suppose that there exists a real number $\eta \neq 0$ such that $\Sigma + i\eta J \geq 0$. Then:*

(i) The matrix Σ must be positive definite and we have $\Sigma + i\eta' J \geq 0$ for every $\eta' \leq \eta$;

[7] This terminology is consistent with the usage in multivariate statistics.

(ii) The inequalities

$$(\Delta X_j)^2(\Delta P_j)^2 \geq \Delta(X_j, P_j)^2 + \eta^2 \tag{83}$$

hold for $1 \leq j \leq n$.

(iii) Let $S \in \operatorname{Sp}(2n, \mathbb{R})$ *and define* $(X_j^S, P_j^S) = S(X_j, P_j)$. *Then*

$$(\Delta X_j^S)^2(\Delta P_j^S)^2 \geq \Delta(X_j^S, P_j^S)^2 + \eta^2. \tag{84}$$

Proof. We first note that the fact that $\Sigma + i\eta J$ is Hermitian is clear since Σ is real symmetric and $(iJ)^* = (-i)(-J) = iJ$. *Proof of (i).* Let us begin by showing that Σ is non-negative. Suppose indeed that Σ has a negative eigenvalue λ, and let z_λ be a real eigenvector corresponding to λ (such an eigenvector exists because Σ is real and symmetric). Since $z_\lambda^T J z_\lambda = 0$ we have

$$z_\lambda^T(\Sigma + i\eta J)z_\lambda = z_\lambda^T \Sigma z_\lambda = \lambda |z_\lambda|^2 < 0$$

which contradicts the assumption $\Sigma + i\eta J \geq 0$. We next show that 0 cannot be an eigenvalue of Σ; this will prove the proposition. Suppose indeed that 0 is an eigenvalue, and let z_0 be a real eigenvector. For $\varepsilon > 0$ set $z(\varepsilon) = (I + i\varepsilon J)z_0$. Using the relations $\Sigma z_0 = 0$, $z_0^T \Sigma = 0$, and $z_0^T J z_0 = \sigma(z_0, z_0) = 0$ we get, after a few calculations,

$$z(\varepsilon)^T(\Sigma + i\eta J)z(\varepsilon) = \varepsilon\eta|z_0|^2 + \varepsilon^2 (Jz_0)^T \Sigma (Jz_0).$$

Choose now ε opposite in sign to η; then $\varepsilon\eta|z_0|^2 < 0$ and if $|\varepsilon|$ is small enough we have $z(\varepsilon)^T(\Sigma + i\eta J)z(\varepsilon) < 0$, which contradicts the fact that $\Sigma + i\eta J \geq 0$. To show that $\Sigma + i\eta' J \geq 0$ for every $\eta' \leq \eta$ it suffices to set $\eta' = r\eta$ with $0 < r \leq 1$ and to note that

$$\Sigma + i\eta' J = (1 - r)\Sigma + r(\Sigma + i\eta J) \geq 0$$

because it is the sum of two non-negative matrices. *Proof of (ii).* The non-negativity of the Hermitian matrix $\Sigma + i\eta J$ can be expressed in terms of the submatrices

$$\Sigma_{ij} = \begin{bmatrix} (\Delta X_j)^2 & \Delta(X_j, P_j) + i\eta \\ \Delta(P_j, X_j) - i\eta & (\Delta P_j)^2 \end{bmatrix}$$

which are non-negative provided that $\Sigma + i\eta J$ is. Since $\operatorname{Tr}(\Sigma_{ij}) \geq 0$ we have $\Sigma_{ij} \geq 0$ if and only if

$$\det \Sigma_{ij} = (\Delta X_j)^2(\Delta P_j)^2 - \Delta(X_j, P_j)^2 - \eta^2 \geq 0$$

which is equivalent to the inequality (83). *Proof of (iii).* Set $\Sigma^S = S^T \Sigma S$; it is the covariance matrix of the vector

$$(X_1^S, ..., X_n^S; P_1^S, ..., P_n^S) = S(X_1, ..., X_n; P_1, ..., P_n).$$

Since S is symplectic we have $S^T JS = J$ and hence

$$\Sigma^S + i\eta J = S^T(\Sigma + i\eta J)S \geq 0;$$

the inequalities (84) follow from the inequalities (83). ■

At this point it is appropriate to notice that (except in the case $n = 1$) the condition $\Sigma + i\eta J \geq 0$ is *not equivalent* to the uncertainty inequalities (83); it is in fact a *stronger* condition. Suppose first $n = 1$. In this case we have

$$\Sigma = \begin{bmatrix} \Delta X^2 & \Delta(X, P) \\ \Delta(P, X) & \Delta P^2 \end{bmatrix}$$

and since $\operatorname{Tr}(\Sigma + i\eta J) = \Delta X^2 + \Delta P^2 \geq 0$ the condition $\Sigma + i\eta J \geq 0$ is equivalent to $\det(\Sigma + i\eta J) \geq 0$ that is to

$$\Delta X^2 \Delta P^2 - (\Delta(X, P)^2 + \eta^2) \geq 0$$

which is (83) in the case $n = 1$. Here is a counterexample when $n = 2$: choose $\eta = 1$ and

$$\Sigma = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

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 (83) are trivially satisfied (they are in fact equalities). The matrix $\Sigma + iJ$ is nevertheless indefinite (its determinant is -1).

Let us introduce the following terminology, due to Littlejohn [61]:

Definition 27 *Let Σ be a covariance matrix. The phase space ellipsoid W_Σ defined by the condition $\frac{1}{2}\Sigma^{-1}z \cdot z \leq 1$ is called the "Wigner ellipsoid" associated with Σ. The dual ellipsoid $W_\Sigma^* : \frac{1}{2}\Sigma z \cdot z \leq 1$ of the Wigner ellipsoid*[8] *is called the "precision ellipsoid" (cf. Definition 25).*

The following geometric lemma is the key to our formulation of the Robertson–Schrödinger uncertainty principle in terms of symplectic capacities:

[8] W_Σ^* is actually just the dual ellipsoid of W_Σ; see Definition 21.

Lemma 28 *The three following conditions are equivalent:*

(i) The Hermitian matrix $\Sigma + i\eta J$ is non-negative (in which case the uncertainty inequalities (83) hold).

(ii) The Wigner ellipsoid $W_\Sigma : \frac{1}{2}\Sigma^{-1}z^2 \leq 1$ satisfies

$$c(W_\Sigma) \geq 2\pi|\eta| \tag{85}$$

(iii) The precision ellipsoid $W_\Sigma^* : \frac{1}{2}\Sigma z^2 \leq 1$ satisfies

$$c(W_\Sigma^*) \leq \frac{\pi}{|\eta|}. \tag{86}$$

Proof. Setting $M = \frac{1}{2}\Sigma^{-1}$ the Wigner ellipsoid is the set of all $z \in \mathbb{R}^{2n}$ such that $Mz^2 \leq 1$ and the condition $\Sigma + i\eta J \geq 0$ is equivalent to $\frac{1}{2}M^{-1} + i\eta J \geq 0$; using a symplectic diagonalization of M this is equivalent to $\frac{1}{2}D^{-1} + i\eta J \geq 0$ where

$$D = \begin{bmatrix} \Lambda^\sigma & 0 \\ 0 & \Lambda^\sigma \end{bmatrix} \quad , \quad \Lambda^\sigma = \operatorname{diag}[\lambda_1^\sigma, ..., \lambda_n^\sigma].$$

It follows that the characteristic polynomial of $\frac{1}{2}M^{-1} + i\eta J$ is the product $P(t) = P_1(t) \cdots P_n(t)$ where

$$P_j(t) = t^2 - (\lambda_j^\sigma)^{-1}t + \tfrac{1}{4}(\lambda_j^\sigma)^{-2} - \eta^2.$$

The eigenvalues of the matrix $\frac{1}{2}M^{-1} + i\eta J$ are thus the real numbers $\frac{1}{2}[(\lambda_j^\sigma)^{-1} \pm 2\eta]$ hence that matrix is non-negative if and only if $\lambda_j^\sigma \leq \frac{1}{2}|\eta|^{-1}$ for every j, that is if and only if $\lambda_{\max}^\sigma \leq \frac{1}{2}|\eta|^{-1}$; in view of formula (53) in Proposition 20 this is equivalent to $c(W_\Sigma) = \pi/\lambda_{\max}^\sigma \geq 2\pi|\eta|$. In view of formula (60) in Proposition 23 this is also equivalent to $c(W_\Sigma^*) = 2\pi\lambda_{\min}^\sigma \leq \pi/|\eta|$. ■

4.1.2 Application: a classical uncertainty principle

That the border between classical and quantum mechanics is rather fuzzy has been known for a long time; in the context of the uncertainty principle see for instance the lucid discussion in Luo [65].

Let us now perform position and momentum measurements on K identical copies of the physical system $\mathcal{S}$ and plot the results of these measurements as a set points $\{z_1, ..., z_K\}$ in the phase space $\mathbb{R}^{2n}$. In the limit $K \to \infty$ we get a cloud of points which we identify with a region Ω of $\mathbb{R}^{2n}$. Let now $\widetilde{\Omega}$ be the convex hull of Ω and denote by $\mathcal{J}$ the John–Löwner ellipsoid of $\widetilde{\Omega}$: it is the (unique) ellipsoid having minimum volume among all other ellipsoids containing $\widetilde{\Omega}$ (its existence follows from a famous theorem in convex geometry proved by John [51] in 1948; see Ball [5] for a fresh point of view and a proof of John's theorem). Let $\langle z \rangle$ be the center of $\mathcal{J}$ and define the matrix $\Sigma > 0$ by

$$\mathcal{J} : \frac{1}{2}\Sigma^{-1}(z - \langle z \rangle)^2 \leq 1. \tag{87}$$

As the notation suggests, we view Σ as a statistical covariance matrix. Let now η be any non-negative number such that

$$c(\Omega) \geq \eta \tag{88}$$

for some symplectic capacity c. The inclusions $\Omega \subset \widetilde{\Omega} \subset \mathcal{J}$ imply, in view of the monotonicity property of symplectic capacities, that we have

$$c(\mathcal{J}) \geq c(\widetilde{\Omega}) \geq c(\Omega) \geq \eta \tag{89}$$

and hence, by the same argument as above, we will have

$$(\Delta X_j)^2(\Delta P_j)^2 \geq \Delta(X_j, P_j)^2 + \eta^2. \tag{90}$$

It turns out that these conditions are conserved in time under Hamiltonian evolution –as they would be in the quantum case. Thus, if we have

$$(\Delta X_j)^2(\Delta P_j)^2 \geq \Delta(X_j, P_j)^2 + \eta^2 \tag{91}$$

at time $t = 0$, then we will have

$$(\Delta X_{j,t})^2(\Delta P_{j,t})^2 \geq \Delta(X_{j,t}, P_{j,t})^2 + \eta^2 \tag{92}$$

for all times t, both future and past. To see why it is so, let us return to the phase space cloud Ω, assuming again that $c(\Omega) \geq \eta$. The Hamiltonian flow f_t^H will deform Ω and after time t it will have become a new cloud $\Omega_t = f_t^H(\Omega)$ with same symplectic capacity (recall that symplectic capacities are invariant under canonical transformations):

$$c(\Omega_t) \geq \eta; \tag{93}$$

it follows that $c(\widetilde{\Omega_t}) \geq \eta$ where $\widetilde{\Omega_t}$ is the convex hull of Ω_t, and hence after time t the John ellipsoid $\mathcal{J}_t$ of the convex hull will also satisfy $c(\mathcal{J}_t) \geq \eta$. This condition is equivalent to the inequalities (92) where $\Delta X_{j,t}$, etc. are defined in terms of the covariance matrix

$$\Sigma_t = \begin{bmatrix} \Sigma_{XX,t} & \Sigma_{XP,t} \\ \Sigma_{PX}, t & \Sigma_{PP,t} \end{bmatrix} \tag{94}$$

determined by $\mathcal{J}_t$ via the time t version of Eqn. (87).

A caveat: there is no particular reason to claim that the quantities $(\Delta X_j)^2$, $(\Delta P_j)^2$, etc. can be here identified with (co)variances in the usual statistical sense; nevertheless one could perhaps identify these quantities with some new kind of measurement of uncertainty, expressed in terms of the topological notion of symplectic capacity; we refer to the discussion de Gosson [35] where classical uncertainty principles are studied from the point of view of robust statistical estimators; also see Reich's New Scientist paper [78] commenting [35].

4.1.3 The quantum case

If we set $\eta = \frac{1}{2}\hbar$ in the inequalities (83) we get

$$(\Delta X_j)^2(\Delta P_j)^2 \geq \Delta(X_j, P_j)^2 + \tfrac{1}{4}\hbar^2. \tag{95}$$

These inequalities are well-known in quantum mechanics as the "Robertson–Schrödinger [79, 83] uncertainty principle[9]". They imply the familiar Heisenberg uncertainty relations[10] $\Delta X_j \Delta P_j \geq \frac{1}{2}\hbar$ if one neglects the covariances $\Delta(X_j, P_j)^2$. Let us explain the meaning of these relations in a quite general setting, that of the density matrix formalism. That is, we are working with a mixed quantum state, represented by a density matrix

$$\widehat{\rho} = \sum_{j \in \mathcal{J}} \alpha_j \left|\psi_j\right\rangle \left\langle\psi_j\right| \tag{96}$$

where the α_j are nonnegative constants summing up to one. Recall that the Wigner distribution of $\widehat{\rho}$ is the function on $\mathbb{R}^{2n}$ defined by

$$\rho = \sum_{j \in \mathcal{J}} \alpha_j W\psi_j \tag{97}$$

where $W\psi_j$ is the usual Wigner distribution of ψ_j:

$$W\psi_j(z) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int_{\mathbb{R}^n} e^{-\frac{i}{\hbar}p\cdot y}\psi_j(x + \tfrac{1}{2}y)\psi_j^*(x - \tfrac{1}{2}y)dy \tag{98}$$

(it is well-defined if $\psi_j \in L^2(\mathbb{R}^n)$; more generally $W\psi_j$ makes sense in the distributional sense if ψ_j is a tempered distribution).

[9] The inequality with the commutator term only was developed in 1930 by Robertson, and Schrödinger added the anticommutator term a little later.

[10] These relations were actually formally written down by Kennard [56].

Let $\widehat{A}$ and $\widehat{B}$ be observables; we assume that the expectation values

$$\langle\widehat{A}\rangle_{\widehat{\rho}} = \operatorname{Tr}(\widehat{\rho}\widehat{A}) \quad , \quad \langle\widehat{A}^2\rangle_{\widehat{\rho}} = \operatorname{Tr}(\widehat{\rho}\widehat{A}^2) \tag{99}$$

(and similar expressions for $\widehat{B}$) exist and are finite. Setting

$$\begin{gathered}(\Delta\widehat{A})^2_{\widehat{\rho}} = \langle\widehat{A}^2\rangle_{\widehat{\rho}} - \langle\widehat{A}\rangle^2_{\widehat{\rho}} \quad , \quad (\Delta\widehat{B})^2_{\widehat{\rho}} = \langle\widehat{B}^2\rangle_{\widehat{\rho}} - \langle\widehat{B}\rangle^2_{\widehat{\rho}} \\ \Delta(\widehat{A},\widehat{B})_{\widehat{\rho}} = \frac{1}{2}\langle\widehat{A}\widehat{B} + \widehat{B}\widehat{A}\rangle_{\widehat{\rho}} - \langle\widehat{A}\rangle_{\widehat{\rho}}\langle\widehat{B}\rangle_{\widehat{\rho}}\end{gathered}$$

we have the following result:

Proposition 29 *Let $\widehat{A} \overset{\mathrm{Weyl}}{\leftrightarrow} a$ and $\widehat{B} \overset{\mathrm{Weyl}}{\leftrightarrow} b$ be two essentially self-adjoint Weyl operators on $L^2(\mathbb{R}^n)$ for which the expectation values (99) are defined. We have*

$$|\langle\widehat{A}\widehat{B}\rangle_{\widehat{\rho}}|^2 = \Delta(\widehat{A},\widehat{B})^2_{\widehat{\rho}} - \tfrac{1}{4}\langle[\widehat{A},\widehat{B}]\rangle^2_{\widehat{\rho}} \tag{100}$$

where $[\widehat{A},\widehat{B}] = \widehat{A}\widehat{B} - \widehat{B}\widehat{A}$ and hence

$$(\Delta\widehat{A})^2_{\widehat{\rho}}(\Delta\widehat{B})^2_{\widehat{\rho}} \geq \Delta(\widehat{A},\widehat{B})^2_{\widehat{\rho}} - \tfrac{1}{4}\langle[\widehat{A},\widehat{B}]\rangle^2_{\widehat{\rho}}. \tag{101}$$

Proof. Replacing, if necessary, $\widehat{A}$ and $\widehat{B}$ by $\widehat{A} - \langle\widehat{A}\rangle_{\widehat{\rho}}$ and $\widehat{B} - \langle\widehat{B}\rangle_{\widehat{\rho}}$ we may assume that $\langle\widehat{A}\rangle_{\widehat{\rho}} = \langle\widehat{B}\rangle_{\widehat{\rho}} = 0$ so that *Eqn.* (100) and (101) reduce to, respectively,

$$|\langle\widehat{A}\widehat{B}\rangle_{\widehat{\rho}}|^2 = \tfrac{1}{2}\langle\widehat{A}\widehat{B} + \widehat{B}\widehat{A}\rangle^2_{\widehat{\rho}} - \tfrac{1}{4}\langle[\widehat{A},\widehat{B}]\rangle^2_{\widehat{\rho}} \tag{102}$$

and

$$\langle\widehat{A}^2\rangle_{\widehat{\rho}}\langle\widehat{B}^2\rangle_{\widehat{\rho}} \geq \tfrac{1}{2}\langle\widehat{A}\widehat{B} + \widehat{B}\widehat{A}\rangle^2_{\widehat{\rho}} - \tfrac{1}{4}\langle[\widehat{A},\widehat{B}]\rangle^2_{\widehat{\rho}}. \tag{103}$$

Writing $\widehat{A}\widehat{B} = \frac{1}{2}(\widehat{A}\widehat{B} + \widehat{B}\widehat{A}) + \frac{1}{2}(\widehat{A}\widehat{B} - \widehat{B}\widehat{A})$ we have,

$$\langle\widehat{A}\widehat{B}\rangle_{\widehat{\rho}} = \tfrac{1}{2}\langle\widehat{A}\widehat{B} + \widehat{B}\widehat{A}\rangle_{\widehat{\rho}} + \tfrac{1}{2}\langle\widehat{A}\widehat{B} - \widehat{B}\widehat{A}\rangle_{\widehat{\rho}}.$$

Now, $\Delta(\widehat{A},\widehat{B})_{\widehat{\rho}}$ is a real number, and $\langle[\widehat{A},\widehat{B}]\rangle_{\widehat{\rho}}$ is pure imaginary (because $[\widehat{A},\widehat{B}]^* = -[\widehat{A},\widehat{B}]$ since $\widehat{A}$ and $\widehat{B}$ are essentially self-adjoint), hence formula (102). We next observe that

$$\langle\widehat{A}\widehat{B}\rangle_{\widehat{\rho}} = \sum_{j\in\mathcal{J}}\alpha_j\langle\widehat{A}\widehat{B}\psi_j|\psi_j\rangle = \sum_{j\in\mathcal{J}}\alpha_j\langle\widehat{B}\psi_j|\widehat{A}\psi_j\rangle; \tag{104}$$

applying the Cauchy–Schwarz inequality to each term $\langle\widehat{B}\psi_j|\widehat{A}\psi_j\rangle$ we get

$$|\langle\widehat{A}\widehat{B}\rangle_{\widehat{\rho}}|^2 \leq \sum_{j\in\mathcal{J}}\alpha_j||\widehat{B}\psi_j||\,||\widehat{A}\psi_j||. \tag{105}$$

Noting that since $\langle\widehat{A}\rangle_{\widehat{\rho}}^2 = \langle\widehat{B}\rangle_{\widehat{\rho}}^2 = 0$ we have

$$||\widehat{A}\psi_j|| = \langle\widehat{A}^2\rangle_{\psi_j}^{1/2} = (\Delta\widehat{A})_{\psi_j}^2 \quad , \quad ||\widehat{B}\psi_j|| = \langle\widehat{B}^2\rangle_{\psi_j}^{1/2} = (\Delta\widehat{B})_{\psi_j}^2$$

the inequality (105) is thus equivalent to

$$|\langle\widehat{A}\widehat{B}\rangle_{\widehat{\rho}}| \leq \sum_{j\in\mathcal{J}} \alpha_j \langle\widehat{A}^2\rangle_{\psi_j}^{1/2} \langle\widehat{B}^2\rangle_{\psi_j}^{1/2}.$$

Writing $\alpha_j = (\sqrt{\alpha_j})^2$ the Cauchy–Schwarz inequality for sums yields

$$|\langle\widehat{A}\widehat{B}\rangle_{\widehat{\rho}}| \leq \left(\sum_{j\in\mathcal{J}} \alpha_j \langle\widehat{A}^2\rangle_{\psi_j}^{1/2}\right) \left(\sum_{j\in\mathcal{J}} \alpha \langle\widehat{B}^2\rangle_{\psi_j}^{1/2}\right) = \langle\widehat{A}^2\rangle_{\widehat{\rho}} \langle\widehat{B}^2\rangle_{\widehat{\rho}}$$

hence the inequality (103) using *Eqn.* (102). ∎

Choosing for $\widehat{A}$ the operator of multiplication by x_j and $\widehat{B} = -i\hbar\partial/\partial x_j$ one obtains the usual Robertson–Schrödinger inequalities

$$(\Delta X_j)_{\widehat{\rho}}^2 (\Delta P_j)_{\widehat{\rho}}^2 \geq \Delta(X_j, P_j)_{\widehat{\rho}}^2 + \tfrac{1}{4}\hbar^2 \tag{106}$$

fort $1 \leq j \leq n$.

Corollary 30 *Assume that* $[\widehat{A}, \widehat{B}] = i\hbar$*; then*

$$(\Delta\widehat{A})_{\widehat{\rho}}^2 (\Delta\widehat{B})_{\widehat{\rho}}^2 \geq \Delta(\widehat{A}, \widehat{B})_{\widehat{\rho}}^2 + \tfrac{1}{4}\hbar^2. \tag{107}$$

If $\widehat{\rho}$ *represents a pure state* $\psi \in D_{\widehat{A}\widehat{B}} \cap D_{\widehat{B}\widehat{A}}$ *then we have equality if and only if the vectors* $(\widehat{A} - \langle\widehat{A}\rangle)\psi$ *and* $(\widehat{B} - \langle\widehat{B}\rangle)\psi$ *are collinear.*

Proof. The inequality (107) immediately follows from the inequality (101). Assume that we have equality in (107). It is sufficient to consider the case $\langle\widehat{A}\rangle_\psi = \langle\widehat{B}\rangle_\psi = 0$. In view of formula $\langle\widehat{A}\widehat{B}\rangle_\psi = \langle\widehat{A}\widehat{B}\psi|\psi\rangle = \langle\widehat{B}\psi|\widehat{A}\psi\rangle$ this means that the Cauchy–Schwarz inequality reduces to an equality, which implies that $\widehat{A}\psi$ and $\widehat{B}\psi$ are colinear. ∎

Assume in particular that $\widehat{A}\psi = x_1\psi$ and $\widehat{B}\psi = -i\hbar\partial\psi/\partial x_1$. The inequality (106) with $j = 1$ becomes an equality if there exists a complex constant λ_1 such that

$$-i\hbar\frac{\partial\psi}{\partial x_1} = \lambda_1 x_1 \psi;$$

it follows that we must have

$$\psi(x) = C(x_2, ..., x_n) e^{-\frac{i}{2\hbar}\lambda_1 x_1^2}$$

for some function C of only the variables $x_2, ..., x_n$. Thus, if we require all the Robertson–Schrödinger equalities (106) to become equalities we must have

$$\psi(x) = C \exp\left[\frac{i}{2\hbar}\sum_{j=1}^{n}\lambda_j x_j^2\right]$$

where C and the λ_j are complex constants; the condition that ψ be square-integrable requires that $\operatorname{Im}\lambda_j > 0$. Choosing in particular $\lambda_1 = \cdots = \lambda_n = i$ and $C = (\pi\hbar)^{-n/4}$ one obtains the usual coherent state

$$\psi(x) = (\pi\hbar)^{-n/4} e^{-\frac{1}{2\hbar}|x|^2}. \tag{108}$$

Let us now shortly address the following important, deep, and difficult question:

> *When is a real symmetric $2n \times 2n$ matrix Σ the covariance matrix of a mixed quantum state $\widehat{\rho}$?*

In the classical case the answer is simple: Σ is the covariance matrix of some probability density if and only if Σ is positive definite. In the quantum case the situation is much more subtle and difficult than it could appear at first sight, because it is plagued by positivity questions. Let in fact $\widehat{\rho}$ be a self-adjoint operator of trace class with trace $\operatorname{Tr}(\widehat{\rho}) = 1$. The operator $\widehat{\rho}$ is thus a candidate for being a density matrix. However, to be eligible, it must in addition be non-negative, that is we must have $\langle\widehat{\rho}\psi|\psi\rangle \geq 0$ for all $\psi \in L^2(\mathbb{R}^n)$, and it is this property which is usually difficult to check. Now, a remarkable result was proven by Narcowich in [74] (also see Narcowich and O'Connell [75]). Let ρ be the Wigner distribution of the Weyl operator $\widehat{\rho}$ (it is $(2\pi\hbar)^{-n}$ times its symbol) and assume that it satisfies condition (75), in which case the covariance matrix Σ is well-defined (the normalization condition (74) is automatically satisfied since $\operatorname{Tr}(\widehat{\rho}) = 1$). Narcowich proved the following:

Theorem 31 (Narcowich) *(i) If a self-adjoint trace-class operator $\widehat{\rho}$ with trace one is a density matrix, then the covariance matrix Σ must satisfy the condition*

$$\Sigma + \tfrac{1}{2}i\hbar J \geq 0$$

(hence, in particular, the Robertson–Schrödinger inequalities must hold);

(ii) Conversely, every real symmetric $2n \times 2n$ matrix Σ satisfying the condition above is the covariance matrix of a mixed quantum state, for instance the Gaussian state with Wigner distribution

$$\rho_\Sigma(z) = \left(\frac{1}{2\pi}\right)^n (\det\Sigma)^{-1/2} \exp\left(-\frac{1}{2}\Sigma^{-1}z^2\right).$$

We urge the reader to note that part (i) of the theorem above says that, given a self-adjoint trace-class operator $\widehat{\rho}$ the condition $\Sigma + \frac{1}{2}i\hbar J \geq 0$ is *necessary* if one wants $\widehat{\rho}$ to be a density matrix, but it is *not sufficient*. In fact, Narcowich's proof uses the so-called KLM conditions (so named after Kastler [52], Loupias and Miracle-Sole [62, 63]). These conditions are a "twisted" form of the usual Bochner conditions for functions of positive type (see Katznelson [54]). The KLM conditions have been discussed by many authors, and are at the origin of the notion of "Wigner spectrum" due to Narcowich [72], and further investigated by Bröcker and Werner [6], and by Costa Dias and Prata [11].

An immediate consequence of Narcowich's theorem is:

Corollary 32 *A positive definite real symmetric $2n \times 2n$ matrix Σ is the covariance matrix of a mixed quantum state if and only if the ellipsoid $W_\Sigma : \frac{1}{2}\Sigma^{-1}z^2 \leq 1$ satisfies $c(W_\Sigma) \geq \frac{1}{2}h$ for some (and hence all) symplectic capacity c.*

Proof. This is just a restatement of Theorem 31 since the condition $\Sigma + \frac{1}{2}i\hbar J \geq 0$ is equivalent to $c(W_\Sigma) \geq \frac{1}{2}h$. ∎

We mention that in our recent work [37, 38]) we have discussed related issues concerning the relationship between positivity questions and the uncertainty principle; also see Luo's interesting discussion of hybrid mixed states in [65, 66], where a distinction between classical and truly quantum uncertainties is made.

4.2 Hardy's uncertainty principle

A "meta-theorem" in harmonic analysis is that a function ψ and its Fourier transform $F^\hbar\psi$ cannot be simultaneously sharply localized. Assume for instance that ψ is of compact support; then its the Fourier transform $F^\hbar\psi$ can be extended to an entire function on the complex plane, and is hence never of compact support. A precise way to express this kind of trade-off between a function and its Fourier transform was discovered in 1933 by Hardy [45].

Hardy showed, using the Phragmén–Lindelöf principle, that if there exist constants C_a and C_b such that $\psi \in L^2(\mathbb{R})$ and its Fourier transform satisfy

$$|\psi(x)| \leq C_a e^{-\frac{a}{2\hbar}x^2} \quad , \quad |F^{\hbar}\psi(p)| \leq C_b e^{-\frac{b}{2\hbar}p^2} \tag{109}$$

(with $a, b > 0$) then the following holds true:

- *If $ab > 1$ then $\psi = 0$;*
- *If $ab = 1$ we have $\psi(x) = Ce^{-\frac{a}{2\hbar}x^2}$ for some complex constant C;*
- *If $ab < 1$ then Eqn.(109) are satisfied by every function of the type $\psi(x) = Q(x)e^{-\frac{a}{2\hbar}x^2}$ where Q is a polynomial.*

In this section we prove a multi-dimensional version of Hardy's theorem and interpret it in terms of symplectic capacities. We will also examine the case of density matrices satisfying a sub-Gaussian estimate.

We begin by complementing Williamson's symplectic diagonalization theorem.

4.2.1 A special case of Williamson's theorem

The following simultaneous diagonalization result is a refined version of Williamson's diagonalization theorem in the block-diagonal case; it can be obtained by specializing the proof given above, but we give here an independent proof. It will be very useful in our discussion of Hardy's uncertainty principle.

Proposition 33 *Let A and B.be two positive-definite symmetric $n \times n$ matrices. There exists $L \in GL(n, \mathbb{R})$ such that*

$$L^T A L = L^{-1} B (L^T)^{-1} = \Lambda \tag{110}$$

where $\Lambda = \operatorname{diag}\left[\sqrt{\lambda_1}, ..., \sqrt{\lambda_n}\right]$ is the diagonal matrix whose eigenvalues are the square roots of the eigenvalues $\lambda_1, ..., \lambda_n$ of AB.

Proof. We make the preliminary observation that if A and B are positive definite matrices then the eigenvalues of AB are real because AB has the same eigenvalues as the symmetric matrix $A^{1/2}BA^{1/2}$. We claim that there exists $R \in GL(n, \mathbb{R})$ such that

$$R^T A R = I \text{ and } R^{-1} B (R^T)^{-1} = D \tag{111}$$

where $D = \operatorname{diag}[\lambda_1, ..., \lambda_n]$. In fact, first choose $P \in GL(n, \mathbb{R})$ such that $P^T A P = I$ and set $B_1^{-1} = P^T B^{-1} P$. Since B_1^{-1} is symmetric, there exists $H \in O(n)$ such that $B_1^{-1} = H^T D^{-1} H$ where D^{-1} is diagonal. Set now $R = PH^T$; we have $R^T A R = I$ and also

$$R^{-1} B (R^T)^{-1} = H P^{-1} B (P^T)^{-1} H^T = H B_1 H^T = D$$

hence the equalities (111). Let $\Lambda = \operatorname{diag}[\sqrt{\lambda_1}, ..., \sqrt{\lambda_n}]$. Since

$$R^T A B (R^T)^{-1} = R^T A R (R^{-1} B (R^T)^{-1}) = D$$

the diagonal elements of D are indeed the eigenvalues of AB hence $D = \Lambda^2$. Setting $L = R\Lambda^{1/2}$ we have

$$L^T A L = \Lambda^{1/2} R^T A R \Lambda^{1/2} = \Lambda$$
$$L^{-1} B (L^{-1})^T = \Lambda^{-1/2} R^{-1} B (R^T)^{-1} \Lambda^{-1/2} = \Lambda$$

hence our claim. ■

The following result allows us to relate Proposition 33 above to Williamson's theorem:

Lemma 34 *Let A and B be two real positive-definite matrices of order n. The symplectic spectrum of $M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ consists of the decreasing sequence $\sqrt{\lambda_1} \geq \cdots \geq \sqrt{\lambda_n}$ of square roots of the eigenvalues λ_j of AB.*

Proof. Let $\operatorname{Spec}_\sigma(M) = (\lambda_1^\sigma, ..., \lambda_n^\sigma)$ be the symplectic spectrum of M. The λ_j^σ are the eigenvalues of

$$JM = \begin{bmatrix} 0 & B \\ -A & 0 \end{bmatrix};$$

they are thus the moduli of the zeroes of the polynomial

$$P(t) = \det(t^2 I + AB) = \det(t^2 I + D)$$

where $D = \operatorname{diag}[\lambda_1, ..., \lambda_n]$; these zeroes are the numbers $\pm i\sqrt{\lambda_j}$, $j = 1, ..., n$; the result follows. ■

For L invertible set

$$M_L = \begin{bmatrix} L^{-1} & 0 \\ 0 & L^T \end{bmatrix}. \tag{112}$$

Obviously $M_L \in \mathrm{Sp}(2n, \mathbb{R})$ (cf. *Eqn.* (15)); Proposition 33 can thus be restated by saying that if (A, B) is a pair of symmetric positive-definite matrices then there exists L such that

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = M_{L^T} \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda \end{bmatrix} M_L. \tag{113}$$

Proposition 33 is thus a precise version of Williamson's theorem for block-diagonal positive matrices.

4.2.2 A multi-dimensional version of Hardy's theorem

To generalize Hardy's uncertainty principle to the multi-dimensional case, the following elementary result about tensor products of functions will be helpful:

Lemma 35 *Let $n > 1$. For $1 \leq j \leq n$ let f_j be a function of $(x_1, .., \widetilde{x_j}, ..., x_n) \in \mathbb{R}^{n-1}$ (the tilde$\widetilde{\ }$suppresses the term it covers), and g_j a function of $x_j \in \mathbb{R}$. If*

$$h = f_1 \otimes g_1 = \cdots = f_n \otimes g_n \tag{114}$$

then there exists a constant C such that $h = Cg_1 \otimes \cdots \otimes g_n$.

Proof. Let us first prove the result for $n = 2$. In this case (114) is

$$h(x_1, x_2) = f_1(x_2)g_1(x_1) = f_2(x_1)g_2(x_2).$$

If $g_1(x_1)g_2(x_2) \neq 0$ then

$$f_1(x_2)/g_2(x_2) = f_2(x_1)/g_1(x_1) = C$$

hence $f_1(x_2) = Cg_2(x_2)$ and $h(x_1, x_2) = Cg_1(x_1)g_2(x_2)$. If $g_1(x_1)g_2(x_2) = 0$ then $h(x_1, x_2) = 0$ hence $h(x_1, x_2) = Cg_1(x_1)g_2(x_2)$ in all cases. The general case follows by induction on the dimension n: suppose that

$$h = f_1 \otimes g_1 = \cdots = f_n \otimes g_n = f_{n+1} \otimes g_{n+1};$$

for fixed x_{n+1} the function

$$k = f_1 \otimes g_1 = \cdots = f_n \otimes g_n$$

is given by

$$k(x, x_{n+1}) = C(x_{n+1})g_1(x_1) \cdots g_n(x_n)$$

where $C(x_{n+1})$ only depends on x_{n+1}. Since we also have

$$k(x, x_{n+1}) = f_{n+1}(x_1, ..., x_n)g_{n+1}(x_{n+1})$$

it follows that $C(x_{n+1})$ is a constant C. ■

Proposition 36 *Let A and B be two real positive definite matrices and $\psi \in L^2(\mathbb{R}^n)$, $\psi \neq 0$. Assume that*

$$|\psi(x)| \leq C_A e^{-\frac{1}{2\hbar}Ax^2} \quad \textit{and} \quad |F^\hbar \psi(p)| \leq C_B e^{-\frac{1}{2\hbar}Bp^2} \tag{115}$$

for some constants $C_A, C_B > 0$. Then:

(i) The eigenvalues λ_j, $j = 1, ..., n$, of AB are ≤ 1;

(ii) If $\lambda_j = 1$ for all j, then $\psi(x) = Ce^{-\frac{1}{2\hbar}Ax^2}$ for some some complex constant C;

(iii) If $\lambda_j < 1$ for some index j then each function $\psi(x) = Q(x)e^{-\frac{1}{2\hbar}Ax^2}$ (Q a polynomial) satisfies the Hardy inequalities (115).

Proof. *Proof of (i).* It is of course no restriction to assume that $C_A = C_B = C$. Let L be as in Lemma 33 and order the eigenvalues of AB decreasingly: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. It suffices to show that $\lambda_1 \leq 1$. Setting $\psi_L(x) = \psi(Lx)$ we have

$$F^\hbar \psi_L(p) = F^\hbar \psi((L^T)^{-1}p);$$

in view of (110) in Lemma 33 condition (115) is equivalent to

$$|\psi_L(x)| \leq Ce^{-\frac{1}{2\hbar}\Lambda x^2} \quad , \quad |F^\hbar \psi_L(p)| \leq Ce^{-\frac{1}{2\hbar}\Lambda p^2} \tag{116}$$

where $\Lambda = \operatorname{diag}[\lambda_1, \lambda_2, ..., \lambda_n]$. Setting $\psi_{L,1}(x_1) = \psi_L(x_1, 0, ..., 0)$ we have

$$|\psi_{L,1}(x_1)| \leq Ce^{-\frac{1}{2\hbar}\lambda_1 x_1^2}. \tag{117}$$

On the other hand, by the Fourier inversion formula,

$$\begin{aligned}\int_{\mathbb{R}^{n-1}} F^\hbar \psi_L(p) dp_2 \cdots dp_n &= (2\pi\hbar)^{n/2} \int_{\mathbb{R}^{2n-1}} e^{-\frac{i}{\hbar}p\cdot x} \psi_L(x) dx dp_2 \cdots dp_n \\ &= (2\pi\hbar)^{(n-1)/2} F^\hbar \psi_{L,1}(p_1)\end{aligned}$$

and hence

$$|F^\hbar \psi_{L,1}(p_1)| \leq C_{L,1} e^{-\frac{1}{2\hbar}\lambda_1 p_1^2} \tag{118}$$

for some constant $C_{L,1} > 0$. Applying Hardy's theorem to the inequalities (117) and (118) we must have $\lambda_1^2 \leq 1$ hence the assertion (i). *Proof of (ii).* The condition $\lambda_j = 1$ for all j means that

$$|\psi_L(x)| \leq Ce^{-\frac{1}{2\hbar}x^2} \quad , \quad |F^\hbar \psi_L(p)| \leq Ce^{-\frac{1}{2\hbar}p^2} \tag{119}$$

for some $C > 0$. Let us keep $x' = (x_2, ..., x_n)$ constant; the partial Fourier transform of ψ_L in the x_1 variable is $F_1^\hbar \psi_L = (F'^\hbar)^{-1} F^\hbar \psi_L$ where $(F'^\hbar)^{-1}$ is the inverse $\hbar$-Fourier transform in the x' variables, hence there exists $C' > 0$ such that

$$|F_1^\hbar \psi_L(x_1, x')| \leq \left(\tfrac{1}{2\pi\hbar}\right)^{\frac{n-1}{2}} \int_{\mathbb{R}^{n-1}} |F^\hbar \psi_L(p)| dp_2 \cdots dp_n \leq C' e^{-\frac{1}{2\hbar} p_1^2}.$$

Since $|\psi_L(x)| \leq C(x') e^{-\frac{1}{2\hbar} x_1^2}$ with $C(x') \leq e^{-\frac{1}{2\hbar} x'^2}$ it follows from Hardy's theorem that we can write $\psi_L(x) = f_1(x') e^{-\frac{1}{2\hbar} x_1^2}$ for some real C^∞ function f_1 on $\mathbb{R}^{n-1}$. Applying the same argument to the remaining variables $x_2, ..., x_n$ we conclude that there exist C^∞ functions f_j for $j = 2, ..., n$, such that

$$\psi_L(x) = f_j(x_1, .., \widetilde{x_j}, ..., x_n) e^{-\frac{1}{2\hbar} x_1^2}. \tag{120}$$

In view of Lemma 35 above we have $\psi_L(x) = C_L e^{-\frac{1}{2\hbar} x^2}$ for some constant C_L; since $\Lambda = I = L^T A L$ we thus have $\psi(x) = C_L e^{-Ax^2/2\hbar}$ as claimed. *Proof of (iii). Assume that* $\lambda_1 < 1$ for $j \in \mathcal{J}$, $\mathcal{J}$ a subset of $\{1, ..., n\}$. By the same argument as in the proof of part (ii) where we established formula (120), we infer, using Hardy's theorem in the case $ab < 1$, that each function

$$\psi_L(x) = f_j(x_1, .., \widetilde{x_j}, ..., x_n) Q_j(x_j) e^{-\frac{1}{2\hbar} x_j^2}$$

(Q_j a polynomial) satisfies the estimates (115) for some constants $C_A, C_B > 0$. One concludes the proof by using once again Lemma 35. ■

Let us give a geometric interpretation of Proposition 36. As announced, it allows us to restate Hardy's uncertainty principle in a very simple geometric way in terms of the symplectic capacity of a certain phase-space ellipsoid. let us first have a look on the case $n = 1$. In this case Hardy's uncertainty principle can be restated by saying that if $\psi \neq 0$ then the conditions $|\psi(x)| \leq C_a e^{-\frac{1}{2\hbar} ax^2}$ and $|F^\hbar \psi(p)| \leq C_b e^{-\frac{1}{2\hbar} bp^2}$ imply that the ellipse $\Omega_{a,b} : ax^2 + bp^2 \leq \hbar$ has area

$$\text{Area}(\Omega_{a,b}) \geq \tfrac{1}{2} h$$

and if the area of the ellipse $\Omega_{a,b}$ is smaller than $\frac{1}{2}h$ then $\psi = 0$. In higher dimensions, once again, the notion of area is replaced by that of symplectic capacity:

Corollary 37 *Let $\psi \in L^2(\mathbb{R}^n)$, $\psi \neq 0$. Assume that*

$$|\psi(x)| \leq C_A e^{-\frac{1}{2\hbar} Ax^2} \quad \textit{and} \quad |F^\hbar \psi(p)| \leq C_B e^{-\frac{1}{2\hbar} Bp^2}. \tag{121}$$

Then the symplectic capacity of the ellipsoid

$$\Omega_{A,B} : Ax^2 + Bp^2 \leq \hbar$$

satisfies $c(\otimes_{A,B}) \geq \frac{1}{2}h$ *where* $h = 2\pi\hbar$*.*

Proof. Setting $M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ the equation of $\Omega_{A,B}$ is $Mz^2 \leq \hbar$. Let $(\lambda_1^\sigma, \lambda_2^\sigma, ..., \lambda_n^\sigma)$ be the symplectic spectrum of M; by *Eqn.* (53) together with the homogeneity property (SC3) of symplectic capacities, we have $c(\Omega_{A,B}) = \pi\hbar/\lambda_1^\sigma$. In view of Lemma 34 we have $\lambda_j^\sigma = \sqrt{\lambda_j}$ where the λ_j are the eigenvalues of AB, and by Proposition 36 we must have $\lambda_j \leq 1$, hence $c(\Omega_{A,B}) \geq \pi\hbar$. ■

4.2.3 Application to sub-Gaussian states

The following formula is well-known[11] Littlejohn [61]; a detailed proof is given in de Gosson [32], §8.5.1). Let ψ be a normalized complex Gaussian ("squeezed state"), written in the form

$$\psi(x) = \left(\frac{\det X}{(\pi\hbar)^n}\right)^{1/4} e^{-\frac{i}{2\hbar}(X+iY)x\cdot x} \tag{122}$$

where X and Y are real symmetric $n \times n$ matrices, X positive definite. Then

$$W\psi(z) = (\pi\hbar)^{-n} e^{-\frac{1}{\hbar}Gz\cdot z} \tag{123}$$

where

$$G = S^T S \quad , \quad S = \begin{bmatrix} X^{1/2} & 0 \\ X^{-1/2}Y & X^{-1/2} \end{bmatrix} \tag{124}$$

that is

$$G = \begin{bmatrix} X + YX^{-1}Y & YX^{-1} \\ X^{-1}Y & X^{-1} \end{bmatrix}. \tag{125}$$

Clearly $S \in \mathrm{Sp}(2n, \mathbb{R})$ so G is also symplectic.

Remark 38 *In a joint paper [64] with Yuri I. Manin the second author has shown that the Wigner transforms of generalized Gaussians are essential in the construction of quantum theta functions.*

[11] It is reputedly due to Bastiaans.

Setting $\Sigma = \frac{\hbar}{2}G^{-1}$ we consider the "Wigner ellipsoid" $W_\Sigma : \frac{1}{2}\Sigma^{-1}z^2 \leq 1$; it is the image by S^{-1} of the ball $B^{2n}(\sqrt{\hbar})$ and has hence symplectic capacity

$$c(W_\Sigma) = c(B^{2n}(\sqrt{\hbar}) = \tfrac{1}{2}h.$$

This suggests that symplectic capacities may be of some use for characterizing phase space functions which are Wigner distributions. We are going to see that this is indeed the case. We will need for the proof the following metaplectic covariance result for mixed states:

Lemma 39 *Let $\widehat{\rho}$ be the density matrix of the state $\psi = \sum_{j\in\mathcal{J}} \alpha_j\psi_j$ with Wigner distribution $\rho = \sum_{j\in\mathcal{J}} \alpha_j W\psi_j$. Then $\widehat{\rho \circ S} = \widehat{S}^{-1}\widehat{\rho}\widehat{S}$ is the density matrix of the state $\widehat{S}\psi = \sum_{j\in\mathcal{J}} \alpha_j \widehat{S}\psi_j$ where $\widehat{S} \in \mathrm{Mp}(2n,\mathbb{R})$ is such that $\pi^\hbar(\widehat{S}) = S$. In particular we have*

$$W(\widehat{S}\psi)(z) = W\psi(S^{-1}z) \tag{126}$$

for every $\psi \in L^2(\mathbb{R}^n)$.

For a proof of this classical property (which is a particular case of the metaplectic covariance of Weyl calculus, see Proposition 3) see for instance de Gosson [32], §7.1.3.

Proposition 40 *Let ψ be square integrable and assume that there exists a constant $C > 0$ and a positive-definite symmetric matrix M such that*

$$W\psi(z) \leq Ce^{-\frac{1}{\hbar}Mz\cdot z} \tag{127}$$

for all z. Then we must have

$$c(W_\Sigma) \geq \tfrac{1}{2}h \tag{128}$$

where W_Σ is the Wigner ellipsoid corresponding to the choice $\Sigma = \frac{\hbar}{2}M^{-1}$.

Proof. (Cf. de Gosson and Luef [37, 38]). In view of Williamson's theorem we can find $S \in \mathrm{Sp}(2n,\mathbb{R})$ such that

$$W\psi(S^{-1}z) \leq Ce^{-\frac{1}{\hbar}(\Lambda x^2 + \Lambda p^2)} \tag{129}$$

where $\Lambda^\sigma = \mathrm{diag}[\lambda_1^\sigma, ..., \lambda_n^\sigma]$ and $\mathrm{Spec}_\sigma(M) = (\lambda_1^\sigma, ..., \lambda_n^\sigma)$. In view of the symplectic covariance formula (126) we have $W\psi(S^{-1}z) = W(\widehat{S}\psi)(z)$ for any $\widehat{S} \in \mathrm{Mp}(2n,\mathbb{R})$ with projection S. Since $\widehat{S}\psi \in L^2(\mathbb{R}^n)$ and $c(\mathcal{W}_\Sigma)$ is a symplectic invariant we may thus assume that $\widehat{S}$ is the identity operator (and hence $S = I$); this choice reduces the estimate to $W\psi(z) \leq Ce^{-\frac{1}{\hbar}Mz^2}$ to the diagonal case

$$W\psi(z) \leq Ce^{-\frac{1}{\hbar}(\Lambda^\sigma x^2 + \Lambda^\sigma p^2)}. \tag{130}$$

Integrating this inequality in x and p, respectively, we get, using formulae (134),

$$|\psi(x)| \leq C_1 e^{-\frac{1}{\hbar}\Lambda^\sigma x^2} \text{ and } |F^\hbar \psi(p)| \leq C_1 e^{-\frac{1}{\hbar}\Lambda^\sigma x^2} \tag{131}$$

for some constant $C_1 > 0$. In view of the multi-dimensional Hardy uncertainty principle (Proposition 36) the eigenvalues λ_j of $(\Lambda^\sigma)^2$ must be ≤ 1; since $\lambda_j = (\lambda_j^\sigma)^2$ we must thus have $\lambda_1^\sigma \leq 1$; in view of *Eqn.* (53) this is equivalent to the inequality $c(\mathcal{B}_M) \geq \frac{1}{2}h$. ■

The following result generalizes Proposition 40 to mixed states:

Proposition 41 *Let $M > 0$ and ρ be a smooth real function on $\mathbb{R}^{2n}$ such that $\int_{\mathbb{R}^{2n}} \rho(z)dz = 1$. Assume that*

$$\rho(z) \leq Ce^{-\frac{1}{\hbar}Mz\cdot z} \tag{132}$$

for some constant $C \geq 0$ and consider the ellipsoid $\mathcal{B}_M : Mz \cdot z \leq \hbar$. If $c(\mathcal{B}_M) < \frac{1}{2}h$ then ρ cannot be the Wigner distribution of a quantum state.

Proof. As in the proof of Proposition 40 we can assume, taking into account Williamson's theorem and the invariance of symplectic capacities under canonical transformations, that

$$\rho(z) \leq Ce^{-\frac{1}{\hbar}(\Lambda^\sigma x^2 + \Lambda^\sigma p^2)}. \tag{133}$$

Let us write $\rho(z) = \sum_{j\in\mathcal{J}} \alpha_j W\psi_j(z)$; integrating the inequality (133) in x and p, respectively we thus have

$$\sum_j \alpha_j \int_{\mathbb{R}^n} W\psi_j(z)dp \leq C_1 e^{-\frac{1}{\hbar}\Lambda^\sigma x^2}$$
$$\sum_j \alpha_j \int_{\mathbb{R}^n} W\psi_j(z)dp \leq C_1 e^{-\frac{1}{\hbar}\Lambda^\sigma p^2}.$$

In view of the marginal properties

$$\int_{\mathbb{R}^n} W\psi_j(z)dp = |\psi(x)|^2 \quad , \quad \int_{\mathbb{R}^n} W\psi_j(z)dx = |F^\hbar \psi(p)|^2 \tag{134}$$

these inequalities imply in particular the existence of constants $C_j > 0$ such that $|\psi_j(x)| \leq C_j e^{-\frac{1}{2\hbar}\Lambda^\sigma x^2}$ and $|F^\hbar \psi_j(p)| \leq C_j e^{-\frac{1}{2\hbar}\Lambda^\sigma p^2}$ and one concludes as in the proof of Proposition 40. ■

As a consequence we get the following result:

Corollary 42 *The Wigner distribution ρ of a mixed quantum state cannot have compact support.*

Proof. Suppose that the support of ρ is contained in some ball $B^{2n}(R) \subset \mathbb{R}^{2n}$. Let λ be a real number such that $0 < \lambda < 1$. We can find $C > 0$ such that $\rho(z) \leq Ce^{-\frac{1}{\hbar}\lambda|z|^2}$ for all z, which contradicts the statement in Proposition 41. ■

5 A Quantum Phase Space

The possibility of using a notion of quantum cell expressed in terms of symplectic capacities was first suggested in de Gosson [27]. This notion is analyzed in depth in the present section.

5.1 Quantum blobs

In [43] Hall and Regginatto remark that

> "...uncertainty relations expressed as imprecise inequalities are not enough to pin down the essence of what is not classical in quantum mechanics..."

In this subsection we introduce a notion which *could* provide a better understanding, by providing a symplectically invariant coarse-graining of phase space. We suggest to replace the notion of cubic quantum cell familiar from statistical quantum mechanics and thermodynamics by a notion invariant under canonical transformations.

5.1.1 Definition and first properties

It is usual in statistical quantum mechanics to "coarse grain" phase space in cubic cells $Q(\sqrt{\hbar})$ with edge length equal to $\sqrt{h} = \sqrt{2\pi\hbar}$. The consideration of such cells, whose volume is h^n, is consistent with the Weyl rule for the number $N(\lambda \leq E)$ of quantum states with energy $\lambda \leq E$, which says that

$$N(\lambda \leq E) \overset{E\to\infty}{\sim} \frac{\mathrm{Vol}(\Omega)}{(\sqrt{2\pi\hbar})^n}$$

where Ω is the phase space volume within the corresponding energy shell $\partial\Omega$. The integer $N(\lambda \leq E)$ is thus roughly the number of quantum cells that can be "packed" inside Ω. One of their main drawbacks with the use of cubic cells is that they have few symmetries and behave badly under canonical transformations. We note that the symplectic capacity of $Q(\sqrt{\hbar})$ can be estimated using Proposition 19; this easily leads to the inequality $c(Q(\sqrt{\hbar})) \geq \frac{1}{2}\pi h$. This suggests to replace the usual notion of quantum cell by a more general notion, invariant under canonical transformations.

Definition 43 *The image $\mathcal{B}$ of a ball $B^{2n}(\sqrt{\hbar})$ by a canonical transformation is called a "quantum blob"; the image $\mathcal{B}^{lin}$ of $B^{2n}(\sqrt{\hbar})$ by a linear canonical transformation (i.e. an element of* $\mathrm{Sp}(2n,\mathbb{R})$ *is called a "linear quantum blob"). [One may of course replace* $\mathrm{Sp}(2n,\mathbb{R})$ *by* $\mathrm{ISp}(2n,\mathbb{R})$ *since the center of the ball is arbitrary in the definition].*

Linear quantum blobs are phase space ellipsoids. A (linear) quantum blob has (linear) symplectic capacity $\frac{1}{2}h$ for every (linear) symplectic capacity: this is an immediate consequence of the canonical invariance of the notion of symplectic capacity together with the fact that

$$c(B^{2n}(\sqrt{\hbar})) = c^{\mathrm{lin}}(B^{2n}(\sqrt{\hbar})) = \tfrac{1}{2}h.$$

The symplectic capacity of a quantum blob $\mathcal{B}$ (linear, or not) is of the same order of magnitude as that of a cubic cell; but quantum blobs do not have the shortcomings of the latter since the image of a quantum blob by a canonical transformation is again a quantum blob.

Conversely, any subset Ω of $\mathbb{R}^n$ such that $c_{\min}(\Omega) \geq \frac{1}{2}h$ for some symplectic capacity c contains a quantum blob $\mathcal{B}$: by definition of $c_{\min}$ there exists a canonical transformation f such that $\mathcal{B} = f(B^{2n}(\sqrt{\hbar})) \subset \Omega$. Similarly, suppose that $c^{\mathrm{lin}}_{\min}(\Omega) \geq \frac{1}{2}h$, then there exists $f \in \mathrm{ISp}(2n,\mathbb{R})$ such that $f(B^{2n}(\sqrt{\hbar})) \subset \Omega$, hence Ω contains a linear quantum blob. The symplectic capacity of a quantum blob $\mathcal{B}$ (linear, or not) being $\frac{1}{2}h$ it is of the same order of magnitude as that of a cubic cell; but quantum blobs do not have the shortcomings of the latter since the image of a quantum blob by a canonical transformation is again a quantum blob. It turns out that the limiting case $c^{\mathrm{lin}}_{\min}(\Omega) = \frac{1}{2}h$ is particularly interesting:

Proposition 44 *Let Ω be a phase space ellipsoid. Assume that $c(\Omega) = \frac{1}{2}h$. Then Ω contains a unique linear quantum blob $\mathcal{B}^{lin}$.*

Proof. (Cf. de Gosson [29]). Since all, symplectic capacities (linear or not) agree on ellipsoids, the condition $c(\Omega) = \frac{1}{2}h$ is equivalent to $c^{\mathrm{lin}}_{\min}(\Omega) = \frac{1}{2}h$.

Translating if necessary Ω it is sufficient to prove the claim in the case where all the considered ellipsoids are centered at zero and the elements of $\mathrm{ISp}(2n, \mathbb{R})$ are linear (i.e. in $\mathrm{Sp}(2n, \mathbb{R})$). The assumption $c_{\min}^{\mathrm{lin}}(\Omega) = \frac{1}{2}h$ thus means that $B^{2n}(\sqrt{\hbar})$ is the largest ball $B^{2n}(R)$ such that $S(B^{2n}(\sqrt{\hbar})) \subset \Omega$ for some $S \in \mathrm{Sp}(2n, \mathbb{R})$. Thus, $\mathcal{B}^{\mathrm{lin}} = S(B^{2n}(\sqrt{\hbar}))$ is a linear quantum blob contained in Ω. Assume now that there exists $S' \in \mathrm{Sp}(2n, \mathbb{R})$ such that $(\mathcal{B}^{\mathrm{lin}})' = S'(B^{2n}(\sqrt{\hbar}))$ is also contained in the ellipsoid Ω and let us show that $(\mathcal{B}^{\mathrm{lin}})' = \mathcal{B}^{\mathrm{lin}}$; the uniqueness statement will follow. Conjugating if necessary S and S' with an adequately chosen symplectic matrix we may assume, using Williamson's diagonalization theorem that

$$\Omega : \sum_{j=1}^{n} \lambda_j^{\sigma}(x_j^2 + p_j^2) \leq 1$$

where the λ_j^{σ} are the symplectic eigenvalues of the matrix of Ω. The condition $c(\Omega) = \frac{1}{2}h$ means that $\lambda_1^{\sigma} = 1$. Setting $U = S'S^{-1}$ we have $S' = US$. Let us show that $U \in U(n)$; the proposition will follows since then $U(B^{2n}(\sqrt{\hbar})) = B^{2n}(\sqrt{\hbar})$ so that $S(B^{2n}(\sqrt{\hbar})) = S'(B^{2n}(\sqrt{\hbar}))$. Clearly $U \in \mathrm{Sp}(2n, \mathbb{R})$ so it suffices, in view of the identity (10) to show that U is in addition a rotation; for this it is sufficient to check that $UJ = JU$. Set $R = D^{1/2}UD^{-1/2}$ where $D = \mathrm{diag}[\lambda_1^{\sigma}, ..., \lambda_n^{\sigma}]$; we have $U^T DU = D$ hence $R^T R = I$ so that R is orthogonal. Let us prove that R is in addition symplectic. Since J commutes with every power of the diagonal matrix D we have, taking into account the relation $JU = (U^T)^{-1}J$ (because U is symplectic):

$$\begin{aligned} JR &= D^{1/2}JUD^{-1/2} = D^{1/2}(U^T)^{-1}JD^{-1/2} \\ &= D^{1/2}(U^T)^{-1}D^{-1/2}J = (R^T)^{-1}J \end{aligned}$$

hence $R^T JR = J$ so that R is indeed symplectic. Since R is also a rotation we have $R \in U(n)$ and thus $JR = RJ$. Since $U = D^{-1/2}RD^{1/2}$ we have

$$\begin{aligned} JU &= JD^{-1/2}RD^{1/2} = D^{-1/2}JRD^{1/2} \\ &= D^{-1/2}RJD^{1/2} = D^{-1/2}RD^{1/2}J = UJ \end{aligned}$$

so that $U \in U(n)$ as claimed. ∎

5.1.2 Linear quantum phase space

Let us denote by $\mathrm{QPS}^{\mathrm{lin}}(2n)$ the set of all linear quantum blobs in $\mathbb{R}^{2n}$ (the acronym QPS stands for Quantum Phase Space). We will call $\mathrm{QPS}^{\mathrm{lin}}(2n)$

the "linear quantum phase space". It is very easy to equip $\mathrm{QPS}^{\mathrm{lin}}(2n)$ with a natural topology. Consider first the case of a linear quantum blob $\mathcal{B}^{\mathrm{lin}} = S(B^{2n}(\sqrt{\hbar}))$ centered at the origin; $\mathcal{B}^{\mathrm{lin}}$ is a centered ellipsoid uniquely represented by the inequality $(S^T S)z \cdot z \leq \hbar$. The symplectic matrix $S^T S$ is positive-definite and, conversely, every positive-definite symplectic matrix can be written in the form $S^T S$. There is thus a one-to-one correspondence between centered linear quantum blobs and positive definite symmetric matrices. These matrices form a topological manifold which is homeomorphic to the coset space $\mathrm{Sp}(2n, \mathbb{R})/U(n)$; this immediately follows from the symplectic polar decomposition theorem. As a manifold $\mathrm{Sp}(2n, \mathbb{R})/U(n)$ has dimension

$$\begin{aligned}\dim(\mathrm{Sp}(2n, \mathbb{R})/U(n)) &= \dim \mathrm{Sp}(2n, \mathbb{R}) - \dim U(n) \\ &= n(2n+1) - n^2 \\ &= n(n+1).\end{aligned}$$

We can thus identify the set of all linear quantum blobs centered at the origin with the Euclidean space $\mathbb{R}^{n(n+1)}$. Now, an arbitrary linear quantum blob is obtained from a centered quantum blob by a phase space translation; such translations form a group isomorphic to $\mathbb{R}^{2n}$. It follows, by dimension count, that we have the identification

$$\mathrm{QPS}^{\mathrm{lin}}(2n) \equiv \mathbb{R}^{n(n+1)} \times \mathbb{R}^{2n} \equiv \mathbb{R}^{n(n+3)}.$$

In the case $n = 1$, corresponding to the phase plane, this identifies the linear quantum phase space with $\mathbb{R}^4$.

Notice that since canonical transformations are volume preserving the volume of a quantum blob is

$$\mathrm{Vol}(\mathcal{B}) = \mathrm{Vol}(B^{2n}(\sqrt{\hbar})) = \frac{1}{n!}\left(\frac{h}{2}\right)^n;$$

their volume is thus much smaller than that of a traditional quantum cell (which is h^n), and decreases very quickly with n. Thus, for a fixed of h, the phase space "graining" rapidly becomes less and less "coarse" when the number of degrees of freedom increases.

The definition of the quantum phase space $\mathrm{QPS}^{\mathrm{lin}}(2n)$ we have given is of a purely geometric nature; it can actually be realized in a more conventional "functional". we are going to see in next subsection that one can associate to every linear quantum blob $\mathcal{B} = \mathcal{B}^{\mathrm{lin}}$ a Gaussian state

$$\psi_{\mathcal{B}}(x) \sim \widehat{T}(z_0)e^{-\frac{1}{2\hbar}(X+iY)x \cdot x}$$

where X and Y are symmetric $n \times n$ matrices, $X > 0$ and $\widehat{T}(z_0)$ the Heisenberg–Weyl operator (3) (the $\psi_{\mathcal{B}}$ are just the squeezed states familiar from quantum optics; see Littlejohn [61] for a thorough study). Now X and Y each depend on $n(n+1)/2$ independent parameters and $\widehat{T}(z_0)$ on $2n$ parameters. it follows that the states $\psi_{\mathcal{B}}$ depend on

$$\frac{1}{2}n(n+1) + \frac{1}{2}n(n+1) + 2n = n(n+3)$$

independent parameters; this is just the dimension of $\mathrm{QPS}^{\mathrm{lin}}(2n)$. We may thus identify the linear quantum phase space $\mathrm{QPS}^{\mathrm{lin}}(2n)$ with the set of all squeezed states.

We take the opportunity to remark that one already finds in Littlejohn [61] (formula (8.12) p.265) an identification between the set of centered Wigner distributions of squeezed states and the coset space $\mathrm{Sp}(2n,\mathbb{R})/U(n)$. As Littlejohn notices, that $\mathrm{Sp}(2n,\mathbb{R})/U(n)$ is not a group, but just a symmetric space, is related to the fact that there is no privileged role for the standard coherent state

$$\psi(x) = (\pi\hbar)^{-n/4} e^{-\frac{1}{2\hbar}|x|^2}$$

if one allows canonical transformation. Geometrically speaking, this means that there is no reason for which we should privilege spherical quantum blobs $B^{2n}(\sqrt{\hbar})$ among all linear quantum blobs $\mathcal{B}^{\mathrm{lin}} = S(B^{2n}(\sqrt{\hbar}))$.

From the identification of $\mathrm{QPS}^{\mathrm{lin}}(2n)$ with the squeezed also follows that the linear quantum blobs are precisely those which correspond to the saturation of the Robertson–Schrödinger uncertainty principle. That is, if the Wigner ellipsoid $W_\Sigma : \frac{1}{2}\Sigma^{-1}z^2 \leq 1$ is a quantum blob $\mathcal{B}^{\mathrm{lin}}$ then we have

$$(\Delta X_j)^2(\Delta P_j)^2 = \Delta(X_j, P_j)^2 + \tfrac{1}{4}\hbar^2$$

for $= 1, 2, ..., n$. This fact can be checked by calculation, but it is also a straightforward consequence of the well-known fact that squeezed states are states of minimum uncertainty.

5.2 The Gaussian state associated with a quantum blob

We are going to see that we can associate a unique canonical quantum state to every quantum blob. more generally, this will be possible for every phase space ellipsoid Ω such that $c(\Omega) = \frac{1}{2}h$.

5.2.1 Quantum blobs and Gaussian states

We begin by stating a technical result which gives a particular factorization result for symplectic matrices:

Lemma 45 *Let* $S = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ *be a symplectic matrix. We have*

$$S = \begin{bmatrix} A_0 & 0 \\ C_0 & A_0^{-1} \end{bmatrix} \begin{bmatrix} X_0 & -Y_0 \\ Y_0 & X_0 \end{bmatrix}$$

where

$$\begin{aligned} A_0 &= (AA^T + BB^T)^{1/2} = A_0^T \\ C_0 &= (CA^T + DB^T)(AA^T + BB^T)^{-1/2} \\ X_0 + iY_0 &= (AA^T + BB^T)^{-1/2}(A - iB). \end{aligned}$$

Proof. See de Gosson [32], Proposition 2.29, p.42 and Corollary 2.30, p.43. ■

It immediately follows that:

Proposition 46 *Every linear quantum blob* $\mathcal{B}^{lin} = S(B^{2n}(\sqrt{\hbar}))$ *centered at the origin can be written* $\mathcal{B} = S_0(B^{2n}(\sqrt{\hbar}))$ *where the symplectic matrix* S_0 *is given by*

$$S_0 = \begin{bmatrix} I & 0 \\ P & I \end{bmatrix} \begin{bmatrix} L & 0 \\ 0 & L^{-1} \end{bmatrix} \tag{135}$$

where L *and* P *are symmetric (and* L *invertible). In fact, if* $S = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ *then*

$$L = (AA^T + BB^T)^{1/2} \ , \ P = (CA^T + DB^T)(AA^T + BB^T)^{-1}. \tag{136}$$

Proof. It is a straightforward consequence of Lemma 45 since the symplectic matrix

$$U = \begin{bmatrix} X_0 & -Y_0 \\ Y_0 & X_0 \end{bmatrix}$$

is a rotation (hence $U \in U(n)$) and thus leaves $B^{2n}(\sqrt{\hbar})$ invariant since it is centered at the origin. ■

Remark 47 *The matrices (135) form a subgroup of the symplectic group* $\mathrm{Sp}(2n, \mathbb{R})$*. That subgroup is the isotropy subgroup (or "stabilizer")* $\mathrm{St}(0 \times \mathbb{R}^n)$ *of the plane of momentum coordinates. To each element* $S_0 \in \mathrm{St}(0 \times \mathbb{R}^n)$ *correspond metaplectic operators* $\pm\widehat{S}_0$ *which are products of unitary rescalings (16) and chirps (15), and are thus local operators. In that sense, the passage from a quantum blob to another is, on the quantum level, a local operation.*

The results above enable us to prove the main result of this section:

Proposition 48 *Let* $\mathcal{B}^{lin}$ *be a linear quantum blob centered at* z_0*:* $\mathcal{B}^{lin} = S_0(B^{2n}(z_0, \sqrt{\hbar}))$ *with* S_0 *given by (135) with* P *and* L *symmetric.*

(i) The function

$$W_{\mathcal{B}}(z) = (\pi\hbar)^{-n} e^{-\frac{1}{\hbar}G(z-z_0)^2} \quad , \quad G = (S_0^T)^{-1} S_0^{-1}.$$

is independent of the choice of S_0 *and is thus uniquely determined by the quantum blob* $\mathcal{B}$*.*

(ii) The function $W_{\mathcal{B}}$ *is the Wigner distribution of the Gaussian state:* $W_{\mathcal{B}} = W(\widehat{T}(z_0)\psi_{\mathcal{B}})$ *with*

$$\psi_{\mathcal{B}}(x) = \left(\frac{\det X}{(\pi\hbar)^n}\right)^{1/4} \exp\left[-\frac{1}{2\hbar}(X + iY)x \cdot x\right] \tag{137}$$

where the symmetric matrices X *and* Y *are given by*

$$X + iY = L^{-2} - iP. \tag{138}$$

Proof. Let us prove (i). Since

$$W(\widehat{T}(z_0)\psi_{\mathcal{B}})(z) = W(\widehat{T}(z_0)\psi_{\mathcal{B}})(z - z_0)$$

it is no restriction to assume that $z_0 = 0$. Let S and S' be such that

$$\mathcal{B} = S(B^{2n}(\sqrt{\hbar}) = S'(B^{2n}(\sqrt{\hbar})).$$

By homogeneity we will then have $S^{-1}S'(B^{2n}(r)) = B^{2n}(r)$ for all radii r and hence $S^{-1}S'$ preserves every ball centered at the origin; it must thus be a rotation and since it is also symplectic we therefore have $S^{-1}S' = U \in U(n)$. We have $S' = SU$ hence $(S'^T)^{-1}(S')^{-1} = (S^T)^{-1}S^{-1}$ and we are done. Proof of (ii). In view of (135) we have

$$(S_0^T)^{-1}S_0^{-1} = \begin{bmatrix} L^{-2} + PL^2P & -PL^2 \\ -L^2P & L^2 \end{bmatrix}.$$

Writing as in (125)

$$G = \begin{bmatrix} X + YX^{-1}Y & YX^{-1} \\ X^{-1}Y & X^{-1} \end{bmatrix}$$

we get formula (138). ■

An important consequence of this result is:

Corollary 49 *Let Ω be a phase space ellipsoid such that $c(\Omega) = \frac{1}{2}h$. There exists a unique Gaussian quantum state associated with Ω.*

Proof. This immediately follows from Proposition 48 using the uniqueness statement in Proposition 44. ■

Remark 50 *When Ω is a Wigner ellipsoid W_Σ the equality $c(W_\Sigma) = \frac{1}{2}h$ means that the Robertson–Schrödinger uncertainty inequalities are partially saturated (i.e., at least one of the n inequalities is an equality).*

Let us generalize this to the case of mixed states. Assume that $\widehat{\rho}$ is the density matrix of a sub-Gaussian mixed state, of the type considered in Subsection 4.2.3. that is, we assume that there exists a constant $C > 0$ such that

$$\rho(z) \leq Ce^{-\frac{1}{\hbar}Mz\cdot z} \tag{139}$$

and consider the Wigner ellipsoid $\mathcal{B}_M : Mz \cdot z \leq \hbar$. We have seen that if $c(\mathcal{B}_M) < \frac{1}{2}h$ then ρ cannot be the Wigner distribution of a quantum state. Let us consider the case where $\mathcal{B}_M$ is a quantum blob; in this case we have $M = (S^T)^{-1}S^{-1}$ and we must have equality in (139) for an appropriate choice of the constant C and we are led back to the case investigated in Proposition 48: ρ is the Wigner transform of a pure Gaussian state, and there is nothing more to say. Suppose now that we have $c(\mathcal{B}_M) = \frac{1}{2}h$. Then $\mathcal{B}_M$ contains a unique quantum blob $\mathcal{B}$ and Proposition 48 allows us again to associate to $\widehat{\rho}$ a unique pure Gaussian state (137). Notice that if $c(\mathcal{B}_M) > \frac{1}{2}h$, that is, if $\widehat{\rho}$ is the density matrix of a general (sub-)Gaussian state there is no way we can associate a unique pure Gaussian to $\widehat{\rho}$, because there is enough room in $\mathcal{B}_M$ to send a ball $B^{2n}(\sqrt{\hbar})$ inside it using many (in fact infinitely many) canonical transformations (see de Gosson [28] for a detailed discussion).

5.2.2 Quantum blobs and averaging

Let us begin by proving a positivity result for the averaging of Weyl operators over quantum blobs.

Let $\widehat{A} \overset{\text{Weyl}}{\longleftrightarrow} a$ be a Weyl operator; one of the features of Weyl calculus is the fact that the condition $a \geq 0$ is neither necessary nor sufficient to ensure non-negativity of the operator $\widehat{A}$. However, if we average a non-negative symbol over a quantum blob, then we will get a non-negative operator.

Let the function

$$\rho_\Sigma(z) = \left(\frac{1}{2\pi}\right)^n (\det \Sigma)^{-1/2} \exp\left(-\frac{1}{2}\Sigma^{-1}z^2\right) \tag{140}$$

be a normalized phase space Gaussian. We recall that for such Gaussians we have

$$\rho_\Sigma * \rho_{\Sigma'} = \rho_{\Sigma+\Sigma'} \tag{141}$$

(cf. any book on elementary probability theory).

Proposition 51 *Assume that $a \geq 0$ and let ρ_Σ be given by Eqn. (140). Let $W_\Sigma : \frac{1}{2}\Sigma^{-1}z^2 \leq 1$ be the corresponding Wigner ellipsoid. If $c(W_\Sigma) \geq \frac{1}{2}h$ then the Weyl operator $\widehat{A}_\Sigma \overset{Weyl}{\longleftrightarrow} a * \rho_\Sigma$ is non-negative: $\widehat{A}_\Sigma \geq 0$.*

Proof. We must show that $\langle \widehat{A}_\Sigma \psi | \psi \rangle \geq 0$ for every $\psi \in L^2(\mathbb{R}^n)$. In view of the density of the Schwartz space of rapidly decreasing functions in $L^2(\mathbb{R}^n)$ is is sufficient to assume that $\psi \in \mathcal{S}(\mathbb{R}^n)$. We have

$$\langle \widehat{A}_\Sigma \psi | \psi \rangle = \int_{\mathbb{R}^{2n}} (a * \rho_\Sigma)(z) W\psi(z) dz;$$

taking into account the fact that ρ_Σ is an even function we get

$$\begin{aligned} \langle \widehat{A}_\Sigma \psi | \psi \rangle &= \iint_{\mathbb{R}^{4n}} a(u)\rho_\Sigma(z-u)W\psi(z)dudz \\ &= \int_{\mathbb{R}^{2n}} a(u) \left[\int_{\mathbb{R}^{2n}} \rho_\Sigma(z-u)W\psi(z)dz\right] du \\ &= \int_{\mathbb{R}^{2n}} a(u)(\rho_\Sigma * W\psi)(u)du. \end{aligned}$$

In view of Proposition 52(i) we have $\rho_\Sigma * W\psi \geq 0$ hence $\rho_\Sigma * W\psi \geq 0$. ■

Here is an example that certainly deserves to be generalized to more general situations. Let

$$H = \frac{1}{2m}(p^2 + m^2\omega^2 x^2)$$

be the one-dimensional harmonic oscillator Hamiltonian and choose

$$\rho_\Sigma(z) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2\sigma^2}(x^2+p^2)\right]$$

where $\sigma > 0$. Set $H_\Sigma = H * \rho_\Sigma(z)$; a straightforward calculation shows that

$$H_\Sigma(z) = H(z) + \frac{\sigma^2}{2m}(1+m\omega^2) \geq H(z) + \sigma^2\omega.$$

Assume now that the ellipse $x^2+p^2 \leq 2\sigma^2$ contains a quantum blob (in this case the image of the disk $x^2+p^2 \leq \hbar$ by an area-preserving mapping of the plane). This is equivalent to the condition $\sigma^2 \geq \frac{1}{2}\hbar$ and hence $H_\Sigma(z) \geq \frac{1}{2}\hbar\omega$. We have thus recovered, as lower bound, the ground energy level of the harmonic oscillator.

It has been known since de Bruijn [7, 8] that the "average" $W\psi * \Phi_R$ of a Wigner distribution $W\psi$ over an elementary Gaussian $\Phi_R(x) = e^{-|z|^2/R}$ is such that:

$$W\psi * \Phi_R \geq 0 \;\text{ if }\; R^2 = \hbar \tag{142}$$

$$W\psi * \Phi_R > 0 \;\text{ if }\; R^2 > \hbar. \tag{143}$$

In our terminology this amounts to say that if we average a Wigner distribution over a ball containing the trivial quantum blob $B^{2n}(\sqrt{\hbar})$ we will obtain a positive function (this is essentially Husimi's [49] procedure for constructing a non-negative distribution function starting from a quasi-distribution). The general notion of quantum blob allows us to improve considerably this result.

Proposition 52 *Let $\Sigma > 0$ be a covariance matrix and $W_\Sigma : \frac{1}{2}\Sigma^{-1}z^2 \leq 1$ the associated covariance matrix. Let ρ_Σ be a Gaussian (140).*

(i) If W_Σ is a linear quantum blob (in which case $c(W_\Sigma) = \frac{1}{2}h$ for every symplectic capacity c) then $W\psi * \rho_\Sigma \geq 0$;

(ii) if $c(W_\Sigma) > \frac{1}{2}h$ then $W\psi * \rho_\Sigma > 0$.

Proof. (i) If W_Σ is a quantum blob $\mathcal{B}$ then $\Sigma = \hbar^{-1}S^TS$ for some $S \in \mathrm{Sp}(2n,\mathbb{R})$ hence there exists a Gaussian state $\psi_\mathcal{B}$ such that $\rho_\Sigma = W\psi_\mathcal{B}$. It follows from the positivity condition (144) that we have

$$W\psi * \rho_\Sigma = W\psi * W\psi_\mathcal{B} \geq 0.$$

(ii) If $c(W_\Sigma) > \frac{1}{2}h$ then there exists $S \in \mathrm{Sp}(2n,\mathbb{R})$ such that W_Σ contains $S(B^{2n}(\sqrt{\hbar}))$ as a proper subset. Let $\Sigma_0 = \hbar^{-1}S^TS$; we have $\Sigma - \Sigma_0 > 0$ and hence $\rho_\Sigma = \rho_{\Sigma-\Sigma_0} * \rho_{\Sigma_0}$ in view of formula (141). It follows, using formula (141), that

$$W\psi * \rho_\Sigma = (W\psi * \rho_{\Sigma-\Sigma_0}) * \rho_{\Sigma_0} > 0$$

as claimed. ■

A classical result is that the convolution of two Wigner distributions $W\psi$ and $W\phi$ yields a positive function; in fact

$$W\psi * W\phi = (2\pi\hbar)^n |F_\sigma^\hbar W(\psi,\phi^v)|^2 \tag{144}$$

where $\phi^v(x) = \phi(-x)$ and $W(\cdot,\cdot)$ is the cross-Wigner distribution (see for instance Folland [19], Proposition 1.99 or de Gosson [32], §8.5, for proofs). This does not mean that the convolution of two Wigner distributions is itself a Wigner distributions: see the very pertinent discussion of this fact in Narcowich [72, 73]. However:

Theorem 53 *Let $\mathcal{B} = S(B^{2n}(\sqrt{\hbar}))$ be a (linear) quantum blob and $\psi_\mathcal{B}$ the associated Gaussian state (formula (137)). The convolution product $W = \rho_\Sigma * W\psi_\mathcal{B}$ is the Wigner distribution of a Gaussian mixed state with Wigner distribution*

$$\rho_{\Sigma,\mathcal{B}}(z) = (\pi\hbar)^{-n}\sqrt{\det M'}e^{-\frac{1}{\hbar}M'z'^2} \tag{145}$$

where the symplectic spectrum of M' is the image of that of $M = \frac{1}{2}\hbar\Sigma^{-1}$ by the mapping $\lambda \longmapsto \lambda/(1+\lambda)$.

Proof. (Part of this was proven in de Gosson [32], Theorem 9.35). Let us write $M = \frac{1}{2}\hbar\Sigma^{-1}$ so that

$$\rho_\Sigma(z) = (\pi\hbar)^{-n}(\det M)^{1/2}e^{-\frac{1}{\hbar}Mz^2}.$$

Since $W\psi_\mathcal{B}(z) = (\pi\hbar)^{-n}e^{-\frac{1}{\hbar}Gz^2}$ we have, by definition of the convolution product,

$$W(z) = (\pi\hbar)^{-2n}(\det M)^{1/2}\int_{\mathbb{R}^{2n}} e^{-\frac{1}{\hbar}M(z-z')^2}e^{-\frac{1}{\hbar}Gz'^2}dz'.$$

Writing $G = (S^T)^{-1}S^{-1}$ and setting $z'' = S^{-1}z'$ we get

$$W(Sz) = (\pi\hbar)^{-2n}(\det M)^{1/2}\int_{\mathbb{R}^{2n}} e^{-\frac{1}{\hbar}S^TMS(z-z'')^2}e^{-\frac{1}{\hbar}|z''|^2}dz''.$$

Replacing if necessary S by another symplectic matrix we may assume in view of Williamson's theorem that

$$S^T MS = D = \begin{bmatrix} \Lambda^\sigma & 0 \\ 0 & \Lambda^\sigma \end{bmatrix}$$

with $\Lambda^\sigma = \operatorname{diag}[\lambda_1^\sigma, ..., \lambda_n^\sigma]$ and hence

$$\begin{aligned} W(Sz) &= (\pi\hbar)^{-2n}(\det M)^{1/2} \int_{\mathbb{R}^{2n}} e^{-\frac{1}{\hbar}D(z-z'')^2} e^{-\frac{1}{\hbar}|z''|^2} dz'' \\ &= \exp\left[-\frac{1}{\hbar}D(I+D)^{-1}z \cdot z\right] \end{aligned}$$

where the last equality is obtained by a straightforward calculation of one-dimensional Gaussian integrals (D is diagonal). It follows that

$$W(z) = e^{-\frac{1}{\hbar}M'z^2} \quad , \quad M' = (S^{-1})^T D(I+D)^{-1}S^{-1}.$$

Let us calculate the symplectic capacity of the ellipsoid $\Omega' : M'z^2 \leq \hbar$. We have $J(S^{-1})^T = SJ$ hence

$$JM' = S\left[JD(I+D)^{-1}\right]S^{-1}$$

hence the symplectic spectrum of M' is that of $D(I+D)^{-1}$; since $D(I+D)^{-1}$ already is in Williamson diagonal form the symplectic eigenvalues of M' are the numbers $\lambda_j'^\sigma = \lambda_j^\sigma(1+\lambda_j^\sigma)^{-1}$. We thus have $\lambda_j'^\sigma \leq 1$ for $j = 1, ..., n$ and hence $c(\Omega') \geq \frac{1}{2}h$. It follows, using Corollary 32, that $\widehat{\rho}_{\Sigma,\mathcal{B}}$ is a density operator. ■

6 EBK Quantization and Symplectic Capacities

Until now the notion of symplectic capacity we have been using has essentially been limited to linear one. In this section we will need the notion in its full generality since we are going to use highly nonlinear canonical transformations (the passage from Cartesian coordinates to action-angle coordinates).

6.1 A heuristic example

The fact that the ground energy level of a one-dimensional harmonic oscillator with Hamiltonian

$$H = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2 x^2$$

is different from zero can heuristically be justified using the Heisenberg uncertainty relation $\Delta p_x \Delta x \geq \frac{1}{2}\hbar$. (See Sudarshan et al. [90] for variations on this theme.) In fact, the inequality $\Delta p_x \Delta x \geq \frac{1}{2}\hbar$ prevents us from assigning simultaneously a precise value to both position and momentum, the oscillator cannot be at rest. To show that the lowest energy has the value $\frac{1}{2}\hbar\omega$ predicted by quantum mechanics one then argues that since we cannot distinguish the origin from a phase plane trajectory whose all points lie inside the double hyperbola $p_x x < \frac{1}{2}\hbar$, we must require that at least one point (x, p_x) of that trajectory is such that $|p_x x| \geq \frac{1}{2}\hbar$. Multiplying both sides of the trivial inequality

$$\frac{p_x^2}{m\omega} + m\omega x^2 \geq 2|px| \geq \hbar$$

by $\omega/2$ we then get

$$E = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2 x^2 \geq \frac{1}{2}\hbar\omega$$

which is the correct lower bound for the quantum energy. The argument above can also be reversed: since the lowest energy of an oscillator with frequency ω and mass m is $\frac{1}{2}\hbar\omega$, the minimal phase space trajectory will be the ellipse

$$\frac{p_x^2}{m\hbar\omega} + \frac{x^2}{(\hbar/m\omega)} = 1;$$

that ellipse encloses a surface with area $\frac{1}{2}h$.

We are going to extend the discussion above (in a rigorous way) to a class of Hamiltonian systems which can be explicitly integrated using "action-angle" variables. Let us first review some basic notions from the theory of integrable Hamiltonian systems and their semiclassical (EBK) quantization.

6.2 Integrable Hamiltonian systems and EBK quantization

Although integrable systems are rather the exception among all Hamiltonian systems, many important physical problems are integrable: the multidimensional harmonic oscillator, the motion in a central force field, two center Newtonian gravitational motion, Lagrange, Euler and Kovalevsky tops, to name a few (see for instance Arnol'd [2] for detailed examples). The interest of such systems in classical mechanics comes from the fact that they are (in principle) exactly solvable using the so-called "action-angle variables". These systems are also very attractive in quantum mechanics (at least in its semiclassical version) because they can be "quantized" semi-classically, that is by imposing constraints on the classical motions (the EBK conditions).

6.2.1 Action-angle variables

Let us review some basic notions about integrable Hamiltonian systems. For a complete and rigorous treatment of integrability in the sense below see for instance Arnol'd [2] or the Appendix A.2 in Hofer and Zehnder's book [47]; Goldstein [22] or Landau and Lifschitz [57] give a very readable account of the topic, perhaps more accessible to physicists less familiar with techniques from differential geometry.

Definition 54 *A time-independent Hamiltonian H on $\mathbb{R}^{2n}$ is "completely integrable" (for short: "integrable") if there exists a symplectomorphism $f : (x,p) \longmapsto (\phi, \mathrm{I})$ (in general not globally defined) such that the composed function $K = H \circ f$ only depends explicitly on the action variables $\mathrm{I} = (\mathrm{I}_1, ..., \mathrm{I}_n)$:*

$$H(z) = H(f(\phi, \mathrm{I})) = K(\mathrm{I}). \tag{146}$$

The numbers $\omega_j(\mathrm{I}) = \omega_j(\mathrm{I}_1, ..., \mathrm{I}_n)$ defined by

$$\omega_j(\mathrm{I}) = \frac{\partial K}{\partial \mathrm{I}_j}(\mathrm{I}) \tag{147}$$

are called the "frequencies of the motion". We will write $\omega(\mathrm{I}) = (\omega_1(\mathrm{I}), ..., \omega_n(\mathrm{I}))$ and call $\omega = \omega(\mathrm{I})$ the "frequency vector".

The Hamilton equations for K are

$$\frac{d\phi_j}{dt} = \frac{\partial K}{\partial \mathrm{I}_j}(\mathrm{I}) \quad , \quad \frac{d\mathrm{I}_j}{dt} = -\frac{\partial K}{\partial \phi_j}(\mathrm{I}); \tag{148}$$

since $\partial K / \partial \phi_j = 0$ for all $j = 1, 2, ..., n$ these equations can be explicitly solved by quadratures, yielding the solutions

$$\phi_j(t) = \omega_j(\mathrm{I}(0))t + \phi(0) \quad , \quad \mathrm{I}_j(t) = \mathrm{I}_j(0). \tag{149}$$

The flow (f_t^K) determined by K is thus given by

$$f_t^K(\phi, \mathrm{I}) = (\omega(\mathrm{I})t + \phi, \mathrm{I}). \tag{150}$$

A typical (but rather trivial) example is provided by the n-dimensional harmonic oscillator: consider the n-dimensional harmonic oscillator Hamiltonian in normal form

$$H = \sum_{j=1}^{n} \frac{\omega_j}{2}(p_j^2 + x_j^2)$$

an define new variables (ϕ, I) by $x_j = \sqrt{2\mathrm{I}_j}\cos\phi_j$ and $p_j = \sqrt{2\mathrm{I}_j}\sin\phi_j$ for $1 \leq j \leq n$; we assume $\mathrm{I}_j \geq 0$ and the angles ϕ_j are chosen such that $0 \leq \phi_j < 2\pi$. In these new variables the Hamiltonian H takes the simple form

$$K(\mathrm{I}) = \sum_{j=1}^{n} \omega_j \mathrm{I}_j \tag{151}$$

(observe that K does not contain ϕ); the transformation $f : (x, p) \longmapsto (\phi, \mathrm{I})$ being canonical outside the origin of $\mathbb{R}^{2n}$ the procedure described above applies.

More generally, one can always define action-angle variables when the manifold $\mathbb{V}^n$ is compact:

Theorem 55 *Assume that $F_1 = H, F_2, ..., F_n$ are n independent constants of the motion in involution on an open dense subset of $\mathbb{R}^{2n}$. Set $F = (F_1, F_2, ..., F_n)$ and assume that $F^{-1}(z_0)$ is a compact and connected n-dimensional manifold. Then:*

(i) there exists a neighborhood $\mathcal{V}$ of $F^{-1}(0)$ in $\mathbb{R}^n$, an open subset $\mathcal{U}$ of $\mathbb{R}^n$ and a canonical transformation

$$f : (\mathbb{R}/2\pi\mathbb{Z})^n \times \mathcal{V} \longrightarrow \mathcal{U}.$$

(ii) Setting $(x, p) = f(\phi, \mathrm{I})$ the Hamiltonian becomes $K(\mathrm{I}) = H(x, p)$ in the (ϕ, I) variables, and the motion thus takes place on a n-dimensional submanifold $\mathbb{V}^n$ of $\mathbb{R}^{2n}$ such that

$$f(\mathbb{V}^n) = \{\mathrm{I}_0\} \times (\mathbb{R}/2\pi\mathbb{Z})^n.$$

Notice that the last statement in *(ii)* follows from Eqn. (150) due to the identification of I-coordinates modulo 2π. The action variables $\mathrm{I} = (\mathrm{I}_1, ..., \mathrm{I}_n)$ are constructed as follows: choose a basis (γ_j) for the 1-cycles on $\mathbb{T}^n$ and set

$$\mathrm{I}_j = \frac{1}{2\pi} \oint_{\gamma_j} p dx. \tag{152}$$

for $1 \leq j \leq n$.

The manifold $\mathbb{V}^n$ is *Lagrangian*, that is, it has dimension n as a manifold, and the tangent spaces $T_z\mathbb{V}^n$ are isotropic: $\sigma(X(z), Y(z)) = 0$ for every pair of tangent vectors to $\mathbb{V}$ at z. This property follows from the fact that the constants of motion $F_1, F_2, ..., F_n$ are in involution.

6.2.2 EBK quantization of Lagrangian manifolds

Let $\mathbb{V}^n$ be a Lagrangian submanifold of $(\mathbb{R}^{2n}, \sigma)$ (typically, $\mathbb{V}^n$ will be a torus $\mathbb{T}^n$ associated to an integrable Hamiltonian system, as described above).

Definition 56 *A Lagrangian manifold $\mathbb{V}^n$ is said to be semiclassically quantized if it satisfies the "EBK" (or "Keller–Maslov") quantization conditions*

$$\frac{1}{2\pi\hbar}\oint_{\gamma} pdx - \frac{1}{4}m[\gamma] \text{ is an integer } \geq 0 \tag{153}$$

for every 1-cycle γ in $\mathbb{V}^n$. The integer $m[\gamma]$ is the Maslov index of the (homotopy class) of γ.

The condition (153) has its origin in earlier work of Einstein, Brillouin, and Keller (hence the acronym *EBK*). In its modern, and mathematically rigorous form it goes back to the work of Maslov [69] (who elaborated on the earlier work of Keller [55]), Maslov and Fedoriuk [70], Arnol'd[12] [1, 2], and Leray [58, 59]. In de Gosson [32] one of us has given a detailed study of the Maslov index and of related mathematical objects. See Littlejohn [61] for a very clear description of the Maslov index at an elementary (but rigorous) level. Roughly speaking, the Maslov index is used to characterize the "caustics" of a Lagrangian manifold through a counting procedure; for instance the Maslov index of a circle in the phase plane and described counterclockwise is $+2$.

A Lagrangian torus $\mathbb{T}^n$ being a Lagrangian manifold we can apply the EBK conditions to $\mathbb{T}^n$; in view of the definition (152) of the action variables and using the fact that the Maslov index of a positively oriented circle is equal to 2 (see [24, 32, 70]) these conditions are equivalent to

$$\mathrm{I}_j = (N_j + \tfrac{1}{2})\hbar \tag{154}$$

where N_j is any integer ≥ 0. The corresponding semi-classical energy levels are then calculated by inserting these values in $K(I)$. More precisely: let $\mathrm{I} = (\mathrm{I}_1, ..., \mathrm{I}_n)$ be the action variables corresponding to the basic one-cycles $\gamma^1, ..., \gamma^n$ on $\mathbb{V}^n$. These are defined as follows: let $\bar{\gamma}^1, ..., \bar{\gamma}^n$ be the loops in

[12] According to Arnol'd himself, the first citatiion was originally intended to be a referee report on Maslov's work.

$\mathbb{T}^n(R_1,...,R_n)$ defined, for $0 \leq t \leq 2\pi$, by

$$\begin{aligned}
\bar{\gamma}^1(t) &= R_1(\cos t, 0, ..., 0; \sin t, 0, ..., 0) \\
\bar{\gamma}^2(t) &= R_2(0, \cos t, ..., 0; 0, \sin t, ..., 0) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots.. \\
\bar{\gamma}^n(t) &= R_n(0, ..., 0, \cos t; 0, ..., 0, \sin t).
\end{aligned}$$

The basic one-cycles $\gamma^1, ..., \gamma^1$ of $\mathbb{V}^n$ are then just

$$\gamma^1 = f^{-1}(\bar{\gamma}^1), ..., \gamma^n = f^{-1}(\bar{\gamma}^n).$$

The action variables being given by

$$\mathrm{I}_j = \frac{1}{2\pi}\oint_{\bar{\gamma}^j} \mathrm{I}d\phi = \frac{1}{2\pi}\oint_{\gamma^j} pdx \quad , \quad 1 \leq j \leq n \tag{155}$$

the EBK quantization conditions (153) imply that we must have

$$\mathrm{I}_j = (N_j + \tfrac{1}{4}m(\gamma^j))\hbar \quad \text{for} \quad 1 \leq j \leq n \tag{156}$$

each N_j being an integer ≥ 0. Writing $H(x,p) = K(\mathrm{I})$ the semiclassical energy levels are then given by the formula

$$E_{N_1,...,N_n} = K((N_1 + \tfrac{1}{4}m(\gamma^1))\hbar, ..., (N_n + \tfrac{1}{4}m(\gamma^1))\hbar) \tag{157}$$

where $N_1, ..., N_n$ range over all *non-negative* integers; they correspond to the physical "quantum states" labeled by the sequence of integers $(N_1, ..., N_n)$.

We do not discuss here the ambiguity that might arise in the calculation of the energy because of the non-uniqueness of the angle action coordinates; that ambiguity actually disappears if one requires that the system under consideration is non-degenerate, that is $\partial^2 K(\mathrm{I}) \neq 0$.

Let us discuss EBK quantization when the Hamiltonian H is a positive definite quadratic form:

$$H_M(z) = \frac{1}{2}Mz \cdot z \quad , \; M > 0. \tag{158}$$

We denote by $\widehat{H}_M$ Hamiltonian operator obtained from H_M by the usual Weyl quantization procedure. We are going to prove the following claim:

Proposition 57 *The exact energy levels (= eigenvalues) of $\widehat{H}_M$ are obtained by the EBK quantization of the Lagrangian manifolds associated to H_M.*

Proof. In view of Williamson's theorem there exists $S \in \mathrm{Sp}(2n, \mathbb{R})$ such that

$$S^T M S = D = \begin{bmatrix} \Lambda^\sigma & 0 \\ 0 & \Lambda^\sigma \end{bmatrix},$$

with $\Lambda^\sigma = \mathrm{diag}[\lambda_1^\sigma, ..., \lambda_n^\sigma]$ where the λ_j^σ are the symplectic eigenvalues of M. We will use in what follows the notation $\lambda_j^\sigma = \omega_j$ which is more in conformity with physical tradition. Thus

$$H_M(z) = H_D(Sz) = \sum_{j=1}^{n} \frac{1}{2}\omega_j(x_j^2 + p_j^2)$$

and, by the symplectic covariance property of Weyl calculus $\widehat{H}_M = \widehat{S}\widehat{H}_D\widehat{S}^{-1}$ where $\widehat{S} \in \mathrm{Mp}(2n, \mathbb{R})$ is any of the two metaplectic operators associated with the symplectic matrix S. The spectra of $\widehat{H}_M$ and $\widehat{H}_D$ are the same: if $\widehat{H}_M\psi = E\psi$ for some $\psi \in L^2(\mathbb{R}^n)$ then $\widehat{H}_D(\widehat{S}^{-1}\psi) = E(\widehat{S}^{-1}\psi)$ and we have $\widehat{S}^{-1}\psi \in L^2(\mathbb{R}^n)$ because metaplectic operators are unitary. The eigenvalues of the self-adjoint operator

$$\widehat{H}_D = \sum_{j=1}^{n} \frac{1}{2}\omega_j \left(x_j^2 - \hbar^2 \frac{\partial^2}{\partial x_j^2} \right)$$

are the numbers

$$E_{N_1,...,N_n} = \sum_{j=1}^{n} (N_j + \frac{1}{2})\hbar\omega_j \qquad (159)$$

where the N_j are integers ≥ 0. Let us check that these numbers are precisely those obtained by the EBK quantization procedure; since the latter is preserved by symplectic transformations, our claim will follow. This is actually folklore, because H_D is just a sum of (rescaled) harmonic oscillator Hamiltonians. The n equations

$$x_j^2 + p_j^2 = r_j^2$$

define a Lagrangian torus $\mathbb{T}^n(R) = S_1^1(R_1) \times \cdots \times S_n^1(R_n)$; the fundamental group $\pi_1(\mathbb{T}^n) \equiv \mathbb{Z}^n$ is generated by the loops $\gamma_j^{(n)} = (0, ..., 0, \gamma_j, 0, ..., 0)$ where

$$\gamma_j(t) = R_j(\cos t, \sin t) \quad , \quad 0 \leq t \leq 2\pi;$$

since $m(\gamma_j) = 2$ the EBK conditions (153) become

$$-\frac{1}{2\hbar}R_j^2 - \frac{1}{2} = -(N_j + 1)$$

for some non-negative integer N_j. The value of H_D on the quantized torus $\mathbb{T}^n(R)$ is precisely given by *Eqn.* (159). ■

What we did above was just to use action-angle variables quantization in disguise: setting

$$x_j = \sqrt{2\mathrm{I}_j}\cos\phi_j \quad , \quad p_j = \sqrt{2\mathrm{I}_j}\sin\phi_j;$$

one easily verifies that the change of variables $(x,p) \longmapsto (\phi, I)$ is canonical and in these coordinates we have

$$H_D(z) = K(\mathrm{I}) = \sum_{j=1}^{n} \omega_j \mathrm{I}_j.$$

The Hamilton equations for K are just $\frac{d}{dt}\phi_j = \omega_j$ and $\frac{d}{dt}\mathrm{I}_j = 0$ hence the I_j are constants of the motion; clearly they are independent and in involution: $\{\mathrm{I}_j, \mathrm{I}_k\} = 0$ so the relations

$$\mathrm{I}_j = \mathrm{I}_j(0) \ , \ 1 \leq j \leq n$$

define Lagrangian tori $\mathbb{V}_{I(0)}$. The EBK quantization conditions are then given by formula (156) and one recovers the energy values (159) using formula (157).

6.3 EBK quantization and symplectic capacities

The main original results of this subsection are Theorem 58 and its consequence, Corollary 61. We begin our discussion with an elementary and informal example illustrating the relationship between the uncertainty principle and the ground energy of a harmonic oscillator.

6.3.1 Symplectic quantization of quadratic Hamiltonians

In the introduction to this section we briefly discussed a relationship between Heisenberg's uncertainty principle and the ground level of the one-dimensional harmonic oscillator. Recall that we used a heuristic argument to justify the fact that the Hamiltonian trajectory in the plane was an ellipse enclosing a surface Ω with area at least $\frac{1}{2}h$. In the case $n = 1$ this area coincides with the symplectic capacity $c(\Omega)$ in view of Proposition 16 so that we could as well write this condition $c(\Omega) \geq \frac{1}{2}h$. Let us discuss what this condition leads to in higher dimensions. For this consider the n-dimensional anisotropic harmonic oscillator described by the Hamiltonian function

$$H = \sum_{j=1}^{n} \frac{1}{2m_j}(p_j^2 + m_j^2\omega_j^2 x_j^2). \tag{160}$$

The solutions of the corresponding Hamilton equations are given by the formulae

$$x_j(t) = x_j(0)\cos\omega_j t + \frac{1}{m_j\omega_j}p_j(0)\sin\omega_j t \tag{161}$$

$$p_j(t) = -m_j\omega_j x_j(0)\sin\omega_j t + p_j(0)\cos\omega_j t \tag{162}$$

so the classical motion takes place on the subspace Ω of $\mathbb{R}^{2n}$ product of the n ellipses

$$\Omega_j : \frac{1}{2m_j}(p_j + m_j^2\omega_j^2 x_j^2) = \epsilon_j$$

where ϵ_j is the energy of the j-th oscillator:

$$\epsilon_j = \frac{1}{2m_j}(p_j(0)^2 + m_j^2\omega_j^2 x_j(0)^2).$$

Let f be the linear transformation of $\mathbb{R}^{2n}$ taking each pair of conjugate coordinates (x_j, p_j) to the re-scaled pair

$$(x_j', p_j') = ((m_j\omega_j)^{-1/2}x_j, (m_j\omega_j)^{1/2}p_j).$$

It is immediate to verify that f is symplectic, and that $f(\Omega)$ is a polydisk:

$$f(\Omega) = B_1^2(\sqrt{2\epsilon_1/\omega_1}) \times \cdots \times B_n^2(\sqrt{2\epsilon_n/\omega_n}).$$

It follows from the symplectic invariance of symplectic capacities and from formula (63) that

$$c_{\min}(\Omega) = c_{\min}(f(\Omega)) = 2\pi \min_j(\epsilon_j/\omega_j).$$

Let us now require that $c_{\min}(\Omega) \geq \frac{1}{2}h$. This condition is equivalent to $\epsilon_j \geq \frac{1}{2}\hbar\omega_j$ for $j = 1, 2, ..., n$. Since the total energy is $E = \sum_{j=1}^n \epsilon_j$ this leads to the lower bound

$$E \geq \sum_{j=1}^n \frac{1}{2}\hbar\omega_j = E_{\text{ground}} \tag{163}$$

where E_{ground} is the correct ground-energy level of the anisotropic harmonic oscillator.

Assume again that the Hamiltonian is of the type

$$H_M(z) = \frac{1}{2}Mz \cdot z$$

as in Subsection 6.2.2 (*Eqn.* (158)). Using Williamson's theorem we can reduce the study of H_M to the normal case

$$H_D(Sz) = \sum_{j=1}^{n} \frac{1}{2}\omega_j(x_j^2 + p_j^2)$$

where the frequencies ω_j are the symplectic eigenvalues of M. The same argument as for the anisotropic harmonic oscillator shows that we again have the lower bound (163) for the energy of the system with Hamiltonian H_M.

We are going to see that the procedure above can be extended to all integrable Hamiltonian systems.

6.3.2 Generalization to integrable Hamiltonians

The ground level corresponds to the case where the formulae (156) reduce to $\mathrm{I}_j = \frac{1}{4}m(\gamma^j)\hbar$ for $1 \leq j \leq n$; since $m[\gamma^j] = 2$, this means that $\mathrm{I}_j = \frac{1}{2}\hbar$. Introducing the canonical variables $X_j = \sqrt{2\mathrm{I}_j}\cos\phi_j$ and $X_j = \sqrt{2\mathrm{I}_j}\sin\phi_j$ the motion is given by $X_j = \sqrt{\hbar}\cos\phi_j$, $P_j = \sqrt{\hbar}\sin\phi_j$ and thus takes place on the torus $\mathbb{T}^n_{\text{ground}} = \mathbb{T}^n(\sqrt{\hbar}, ..., \sqrt{\hbar})$, product of n circles with radius $\sqrt{\hbar}$. Since

$$c_{\min}(\mathbb{T}^n_{\text{ground}}) = \pi\hbar = \tfrac{1}{2}h$$

we conclude, using the invariance of symplectic capacities under canonical transformations, that in the original variables x, p the motion takes place on a Lagrangian manifold $\mathbb{V}^n_{\text{ground}}$ such that $c(\mathbb{V}^n_{\text{ground}}) = \frac{1}{2}h$. The following result generalizes this observation:

Theorem 58 *Assume that $\mathbb{V}^n$ is compact and connected and that there exists a canonical transformation $f : \mathbb{V}^n \longrightarrow \mathbb{T}^n = \mathbb{T}^n(R_1, ..., R_n)$. Consider the set $\bar{\mathbb{V}}^n = f^{-1}(\mathbb{D}^n)$ where $\mathbb{D}^n = \mathbb{D}^n(R_1, ..., R_n)$ is the polydisk whose boundary is the torus $\mathbb{T}^n$.*

(i) If $\mathbb{V}^n$ satisfies the EBK condition (153), then for every symplectic capacity c we have

$$c(\bar{\mathbb{V}}^n) \geq \tfrac{1}{2}h \tag{164}$$

where the set $\bar{\mathbb{V}}^n$ is defined by $f(\bar{\mathbb{V}}^n) = \mathbb{D}^n$ (f the mapping $(x, p) \longmapsto (\phi, \mathrm{I})$). In particular, $\bar{\mathbb{V}}^n$ contains a quantum blob $\mathcal{B}$.

(ii) If conversely $\bar{\mathbb{V}}^n$ satisfies (164), and the frequencies ω_j are everywhere > 0 then the energy E of the motion carried by $\mathbb{V}^n$ is such that

$$E \geq E_0 = K(\tfrac{1}{2}\hbar, ..., \tfrac{1}{2}\hbar); \tag{165}$$

In view of Eqn. (168) we have $m(\gamma^j) \geq 2$ for every basic one-cycle γ^j on $\mathbb{V}^n$ and hence

$$E_{N_1,...,N_n} \geq E_0 = K(\tfrac{1}{2}\hbar, ..., \tfrac{1}{2}\hbar); \tag{166}$$

the number E_0 is a lower bound for the quantized energy levels $E_{N_1,...,N_n}$ given by Eqn. (157).

Proof. *(i)* Since capacities are symplectic invariants, we may assume without restricting the generality of the argument that $\mathbb{V}^n$ is the torus $\mathbb{T}^n$ itself. Since

$$\mathrm{I}_j = \frac{1}{2\pi}\oint_{\gamma^j} pdx = \frac{1}{2\pi}\oint_{\bar{\gamma}^j} \mathrm{I}_j d\phi_j = \frac{1}{2}R_j^2$$

the quantization conditions (156) are equivalent to the conditions

$$R_j^2 = (2N_j + \tfrac{1}{2}m(\gamma^j))\hbar.$$

As a manifold $\mathbb{V}^n$ (and hence $\mathbb{T}^n$) has dimension n; we must thus have $R_j > 0$ for every j, and this implies that $m(\gamma^j) > 0$ for every basic one-cycle γ^j. It follows that

$$\inf_{1\leq j\leq n} R_j^2 \geq \tfrac{1}{2}\hbar \inf_{1\leq j\leq n} m(\gamma^j) > 0. \tag{167}$$

We next observe that the torus $\mathbb{T}^n = \mathbb{T}^n(R_1, ..., R_n)$ is an oriented manifold (because it is a product of circles, which are oriented manifolds). It follows that $\mathbb{V}^n = f^{-1}(\mathbb{T}^n)$ is also oriented (symplectomorphisms are orientation preserving). The Maslov index $m(\gamma^j)$ of every basic one-cycle on $\mathbb{V}^n$ is even, and hence

$$\inf_{1\leq j\leq n} m(\gamma^j) \geq 2. \tag{168}$$

It follows from the inequalities (167) and (168) that we have

$$c(\mathbb{D}^n(R_1, ..., R_n)) = \pi \inf_{1\leq j\leq n} R_j^2 \geq \tfrac{1}{2}h$$

as was to be proven. *(ii)* Assume that conversely

$$c(\bar{\mathbb{V}}^n) = c(\mathbb{D}^n) \geq \tfrac{1}{2}h.$$

The motion takes place on a torus $\mathbb{T}^n = \mathbb{T}^n(R_1, ..., R_n)$ such that

$$\pi \inf_{1\leq j\leq n} R_j^2 \geq \tfrac{1}{2}h$$

and we thus have

$$\mathrm{I}_j = \frac{1}{2\pi}\oint_{\bar{\gamma}^j} \mathrm{I}_j d\phi_j = \frac{1}{2}R_j^2 \geq \frac{1}{2}\hbar.$$

The assumption

$$\omega_j(\mathrm{I}) = \partial_{I_j} K(\mathrm{I}) > 0$$

implies that K is an increasing function of the variables $\mathrm{I} = (\mathrm{I}_1, ..., \mathrm{I}_n)$ and we thus have

$$E = K(\mathrm{I}_1, ..., \mathrm{I}_n) \geq K(\tfrac{1}{2}\hbar, ..., \tfrac{1}{2}\hbar).$$

In view of *Eqn.* (168) we have $m(\gamma^j) \geq 2$ for every basic one-cycle γ^j on $\mathbb{V}^n$ and hence

$$E_{N_1,...,N_n} \geq K(\tfrac{1}{2}\hbar, ..., \tfrac{1}{2}\hbar)$$

which ends the proof. ■

Remark 59 *The observant reader will notice from the proof above that it is the first instance we use the invariance of symplectic capacities under the action of arbitrary (non-linear) canonical transformations.*

At this point it is perhaps useful to note that if we had imposed the EBK quantization conditions (153) on the tori $\mathbb{T}^n$ themselves (and not on the Lagrangian manifolds $\mathbb{V}^n$) then the number E_0 in (165) effectively coincides with the ground energy level; but in the general case we have $E_0 < E_{N_1,...,N_n}$. This is due to the fact that $\mathbb{V}^n$ can have more caustic points than $\mathbb{T}^n$ has: the notion of caustic is not intrinsically attached to the Lagrangian manifold itself, but is instead related to the representation of that manifolds that is used (see the lucid discussion in Littlejohn [61]).

Another obvious consequence of the considerations above is:

Proposition 60 *The EBK quantization rules (153) are equivalent to the following conditions*

Proof. This immediately follows from the identity $c_{\min}(\mathbb{D}^n(R)) = \pi R_{\min}^2$. ■

From Theorem 58 we easily deduce the following form of the uncertainty principle for integrable systems. Let us call a planar curve λ "simple" if it can be smoothly deformed into a circle. It follows from a deep result of Dacorogna and Moser [12] that if $\lambda = \partial\Lambda$ then there exists an area-preserving diffeomorphism f such that $f(\Lambda)$ is a disk with same area.

Corollary 61 *Let $\mathbb{V}^n$ be a compact and connected Lagrangian manifold associated to a Liouville-integrable Hamiltonian system. If $\mathbb{V}^n$ satisfies the EBK quantization conditions (153) then the following property holds: Let*

Λ_j be a connected subset in the x_j, p_j plane bounded by a simple curve λ_j. If Λ_j contains the projection of $\mathbb{V}^n$ then we have

$$\operatorname{Area}(\Lambda_j) \geq \tfrac{1}{2}h \tag{169}$$

for every $j = 1, 2, ...n$.

Proof. Recall from Theorem 131 that if $\mathbb{V}^n$ is quantized then $c(\bar{\mathbb{V}}^n) \geq \frac{1}{2}h$. Let us first assume that λ_j is a circle $S_j^1(R)$. Then $\bar{\mathbb{V}}^n \subset Z_j(R)$ and hence

$$\tfrac{1}{2}h \leq c(\bar{\mathbb{V}}^n) \leq c(Z_j(R)) = \pi R^2$$

proving the claim in that case. If λ_j is not a circle, choose an area-preserving diffeomorphism f of the x_j, p_j plane taking Λ_j into a circle $S_j^1(R)$. The phase space transformation F taking (x_j, p_j) into $f(x_j, p_j)$ and leaving all other coordinates unchanged is symplectic, and the projection of $F(\bar{\mathbb{V}}^n)$ lies inside $S_j^1(R)$. We have

$$c(\bar{\mathbb{V}}^n) = c(F(\bar{\mathbb{V}}^n)) \leq c(Z_j(R)) = \operatorname{Area}(\Lambda_j)$$

hence again $\operatorname{Area}(\Lambda_j) \geq \frac{1}{2}h$. ∎

We note that the result above can also be directly deduced from Gromov's theorem: if $c(\bar{\mathbb{V}}^n) \geq \frac{1}{2}h$ for every symplectic capacity c then, in particular, $c_{\min}(\bar{\mathbb{V}}^n) \geq \frac{1}{2}h$. The largest ball $B^{2n}(R)$ that can be squeezed inside $\bar{\mathbb{V}}^n$ by a symplectomorphism f has therefore radius $\sqrt{\hbar}$, and the area of the orthogonal projection of $f(B^{2n}(\sqrt{\hbar}))$ on any x_j, p_j plane is then at least $\frac{1}{2}h$.

7 Open Questions and Perspectives

Here is a –certainly non-exhaustive!– list of open topics and problems related to the main themes of this paper, and which could be (hopefully) attacked using methods from symplectic topology.

7.1 Other uncertainty principles

The usual uncertainty principle of quantum mechanics is expressed in terms of variances and covariances of the involved variables; these are calculated using a wavefunction or its Wigner transform). Less known by physicists is the uncertainty principle of Donoho and Stark [13], widely used in signal theory and time-frequency analysis. Let us say that a function $\psi \in L^2(\mathbb{R}^n)$ is ε-concentrated on a (measurable) subset X of $\mathbb{R}^n$ if

$$\int_{X^c} |\psi(x)|^2 dx \leq \varepsilon^2 ||\psi||^2.$$

The Donoho–Stark uncertainty principle says that if $\psi \neq 0$ is ε_X-concentrated on X and its Fourier transform $F\psi$ is ε_P-concentrated on $P \subset \mathbb{R}^n$ then we must have

$$\operatorname{Vol}(X) \operatorname{Vol}(P) \geq (1 - \varepsilon_X - \varepsilon_P)^2.$$

It would be interesting to reformulate this property using the notion of symplectic capacity; a first step might be to evaluate $c_{\min}(X \times P)$; a more global approach could consist in translating the Donoho–Stark uncertainty principles into a property of the Wigner distribution of ψ and to estimate the "concentration support" of $W\psi$ using some symplectic capacity c (or all).

It would also be very interesting to see whether there is a symplectic interpretation of Luo and Zhang's [67] information-theoretical approach to the Robertson–Schrödinger uncertainty principle; Luo and Zhang's approach seems to be very promising since it allows to distinguish between classical and quantum uncertainty in mixed sates. Perhaps their ideas could be recast in terms of the quantum phase space?

7.2 EBK quantization

We have seen in Subsection 6 that the minimal Lagrangian torus $\mathbb{V}^n$ carrying a quantized motion has symplectic capacity $\frac{1}{2}h$. Conversely, the condition $c(\mathbb{V}^n) = \frac{1}{2}h$ characterizes the ground energy level, and allows us to calculate it explicitly. A natural question that arises is the following: can we characterize the excited levels using symplectic capacities? Experimenting with a simple system (the one-dimensional harmonic oscillator) does apparently not give any valuable clue. It seems that the difficulty comes from the fact that we do not know of any generalized symplectic capacities which allow to distinguish between polydisks; the situation would certain clarify if we were able to find a sequence $(c_k)_{k\geq 1}$ of capacities (normalized or not) such that $c_k(\mathbb{D}^{2n}(R)) = \pi R_k^2$ if

$$\mathbb{D}^{2n}(R) = B_1^2(R_1) \times B_2^2(R_2) \times \cdots \times B_n^2(R_n)$$

(in which case $\mathbb{V}^n = \partial\mathbb{D}^{2n}(R)$). Perhaps such capacities would allow a reformulation of the Keller–Maslov conditions in terms of symplectic objects.

One of us has shown in [25] that if one assumes that the semi-classical wavefunctions are of a certain type familiar from geometric quantization (see for instance de Gosson [24], Chapter 5), then the knowledge of the ground level automatically leads to all the excited energy levels. This approach, based on method from algebraic topology (the theory of covering spaces), is however indirect, and does not solve the problem posed above.

7.3 Adiabatic invariance

The notion of adiabatic invariance in its modern form has its roots in the 1916 work of Ehrenfest. Roughly stated, "adiabatic invariance" refers to the approximate conservation of certain quantities (for instance the action of periodic orbits, or the ratio energy/frequency) under slow modifications of certain parameters of the system under consideration. The archetypical example is "Einstein's pendulum", discussed by Einstein and Lorentz during the Solvay Conference of 1911. That example goes as follows: consider a bob swinging on a string which is slowly being drawn up or down through a hole. The initial long pendulum has a Hamiltonian H_{in} whose phase space trajectories are concentric ellipses; the bob is running along one of them. As the pendulum's length is varied, the phase space trajectories become squashed horizontally so that after some time T the pendulum has a Hamiltonian H_{fin} whose phase space trajectories are again concentric ellipses, but with a different shape. The bob will again be running around one of them. The question is: *which one*? The answer is: along the one with the same *area* as the initial ellipse. In this case the area is related to the *action integral* and the Hofer–Zehnder capacity c^{HZ} discussed in Subsection 3.2.2 relates the notion of symplectic capacity to that of action (the relationship is obvious when $n = 1$). Since we have (formula (46))

$$c^{\text{HZ}}(\Omega) = \oint_{\gamma_{\min}} pdx$$

for every compact convex subset of phase space it is natural to conjecture that c^{HZ} (or perhaps more general symplectic capacities?) might be adequate for the description of adiabatic properties in higher dimensions. the first author has examined this possibility in [26, 31], but there is still much work needed to come to a conclusive and general theory.

7.4 Guth's "catalyst map"

In Subsection we mentioned a possible physical application of Guth's embedding result [44] for polydisks. The argument is the following. Let us consider a physical system consisting of three particles moving along a line (the x-axis); the phase space is thus here $\mathbb{R}^6$. Assume that the system is, at initial time $t = 0$ in a polydisk $\mathbb{D}(\varepsilon, 1, 1) = B_1^2(\varepsilon) \times B_2^2(1) \times B_3^2(1)$ where ε is very small. We thus have very good knowledge of the state of the first particle and medium knowledge of the states of the two other particles. Consider now the generalized polydisk $\mathbb{D}(2\varepsilon, 10\varepsilon, \infty) = B_1^2(2\varepsilon) \times B_2^2(10\varepsilon) \times \mathbb{R}^2$. If ε is small enough Guth [44] shows that there exists a canonical transformation f sending $\mathbb{D}(\varepsilon, 1, 1)$ into $\mathbb{D}(2\varepsilon, 10\varepsilon, \infty)$. This amounts to say that we can in some way evolve the system so that one loses a little knowledge about the first particle, and gains much knowledge about the second particle, but at the expense of all the knowledge about the third particle. Following Guth's terminology, the first particle has thus played the role of a catalyst: its state comes out almost unchanged, but with its help the two other particles have interacted in a way they could not have done on their own! For this reason Guth calls the embedding f the "catalyst map". It would certainly be very interesting to push the argument further, and apply it for larger system. An understanding of the catalyst map, and its generalization to higher-dimensional systems is still lacking; Proposition 18 would certainly be helpful to understand the situation provided that the constant $C(n)$ in the estimates (49) is known, at least with some precision (Guth's result is a mathematical existence theorem, and it seems hard to extract from it any numerical value for $C(n)$).

7.5 Wigner spectrum and uncertainty

We have briefly discussed in Subsection 4.1.3 the relation between the positivity of an operator $\widehat{\rho}$ and the uncertainty principle, and mentioned the KLM conditions. In [72] Narcowich introduces the notion of Wigner spectrum $WS(\widehat{\rho})$ of a trace-class self-adjoint operator $\widehat{\rho}$ *on* $L^2(\mathbb{R}^n)$. By definition, $WS(\widehat{\rho})$ is the set of all real numbers η such that $\widehat{\rho}$ is positive semidefinite, that is for which ρ_σ is of η-positive type. Thus, $\widehat{\rho}$ is a density matrix if and only if $\hbar \in WS(\widehat{\rho})$. The set $WS(\widehat{\rho})$ is closed and bounded and hence compact. As Bröcker and Werner [6] point out, pure states with positive Wigner distributions are just the usual coherent states, and it is easy to see that the Wigner spectrum of such a state contains the interval $[0, \hbar]$. It would certainly be very interesting to study the Wigner spectrum and its properties from the point of view of the "quantum blobs" we have introduced. We mention that Costa Dias and Prata [11] have obtained decisive results that certainly deserve to be studied further; for instance these authors show that the only pure states having full Wigner spectrum are

Gaussians; this property is certainly related in some way to the notion of quantum phase space $\mathrm{QPS}^{\mathrm{lin}}(2n)$ introduced in our paper.

7.6 Flandrin's conjecture

According to this conjecture we would have

$$\int_{\Omega} W\psi(z)dz \leq ||\psi||$$

for all square integrable functions ψ when Ω is a convex subset of the phase plane R^2. (we have equality for $\Omega = R^2$. The conjecture of Flandrin says that over convex sets Ω the positive and negative values of the Wigner distribution $W\psi$ cancel each other to an extent that the integral over Ω does not exceed the integral over the whole plane. In particular the set of z such that $W\psi \geq 0$ can never be a convex set, unless it is the whole plane. Flandrin proved in [18] that this conjecture is true for balls of radius $\sqrt{\hbar}$. Janssen showed in [50] that this follows from the well-known fact that the Wigner distribution of Hermite functions are Laguerre polynomials. It would be very interesting to generalize Flandrin's results to the n-dimensional case; the present authors wouldn't be too surprised if the notion of quantum capacity (and the related concept of quantum blob) played a key role in this generalization.

7.7 Modulation spaces

The approach outlined in this paper has been almost purely geometrical, and we have been somewhat sloppy with issues related to functional spaces. There is however a very interesting point which deserves to be studied in depth. In our discussion of the covariance matrix in Subsection 4.1.1 we imposed the integrability condition (75) to the quasi-distribution ρ; this condition ensures that the associated covariance matrix is well-defined. Let us introduce the following generalization of this condition. Namely, for $s \geq 0$ we consider those quasi-distributions ρ such that

$$\int_{\mathbb{R}^{2n}} |\rho(z)|(1+|z|^2)^{s/2}dz < \infty; \tag{170}$$

condition (75) corresponds to the choice $s = 2$. Writing $\rho = \sum_j \alpha_j W\psi_j$ condition (170) is satisfied as soon as ψ_j is in the Feichtinger modulation space $M_s^1(\mathbb{R}^n)$ which plays an ubiquitous role in time-frequency analysis and signal theory (see Feichtinger [16, 17]; Gröchenig gives a comprehensive study of these spaces in his book [39]). The spaces $M_s^1(\mathbb{R}^n)$ are Banach spaces; they have many useful properties, for instance they are invariant under metaplectic operators. In particular, $\psi \in M_s^1(\mathbb{R}^n)$ implies $F^\hbar\psi \in M_s^1(\mathbb{R}^n)$. It is reasonable to conjecture that the modulation spaces $M_s^1(\mathbb{R}^n)$ are the "right" spaces to use when dealing with quantum mechanical states, mixed, or pure. We hope to come back to this important functional-analytical topic in a near future.

References

[1] V.I. Arnold, A characteristic class entering in quantization conditions. Funkt. Anal. i. Priloz. 1(1) (1967) 1–14 (in Russian); Functional Analysis and Applications 1 (1967) 1–14.

[2] V.I. Arnold, Mathematical Methods of Classical Mechanics. Graduate Texts in Mathematics, second edition, Springer-Verlag (1989).

[3] S. Artstein-Avidan, V. Milman, and Y. Ostrover, The M-ellipsoid, Symplectic Capacities and Volume, Commentarii Mathematici Helvetici 83(2) (2008) 359–369; Arxiv preprint math.SG/0603411 (2006).

[4] S. Artstein-Avidan, V. Milman, The Concept of Duality in Convex Analysis, and the Characterization of the Legendre Transform, Annals of Mathematics 169 (2009), 661–674.

[5] K.M. Ball, Ellipsoids of maximal volume in convex bodies, Geometrica Dedicata 41(2) (1992), 241–250.

[6] T. Bröcker and F. Werner, Mixed states with positive Wigner functions. Journal of Mathematical Physics 36(1) (1995) 62–75.

[7] N.G. de Bruijn, Inequalities, 55–71 in: O. Shisha, ed., Academic Press, NY, 1967.

[8] N.G. de Bruijn, A theory of generalized functions, with applications to Wigner distribution and Weyl correspondence, Nieuw Archiev voor Wiskunde 21 (1973) 205–280.

[9] P. Busch, T. Heinonen, and P. Lahti, Heisenberg's uncertainty principle, Physics Reports (2007)

[10] K. Cieliebak, H. Hofer, J. Latschev, and F. Schlenk. Quantitative Symplectic Geometry, in: Dynamics, ergodic theory, and geometry, 1–44,

Mathematical Research Institute Publishing 54, Cambridge University Press, 2007; arXiv:math/0506191v1 [math.SG].

[11] N. Costa Dias, J. Prata, The Narcowich–Wigner spectrum of a pure state, arXiv:0812.0043v1 [math-ph]; to appear in Reports in Mathematical Physics (2009).

[12] B. Dacorogna and J. Moser, On a partial differential equation involving the Jacobian determinant, Annales de l'Institut Henri Poincaré, analyse non linéaire 7 (1990) 1–26.

[13] D.L. Donoho and P.B. Stark, Uncertainty principles and signal recovery, SIAM Journal of Applied Mathematics 49(3) (1989) 906–931.

[14] I. Ekeland and H. Hofer, Symplectic topology and Hamiltonian dynamics, I and II. Mathematische Zeitschrift 200 (1990) 355–378 and 203 (1990) 553–567.

[15] Ya. Eliashberg, S. Kim, and L. Polterovich, Geometry of contact transformations and domains: orderability versus squeezing, Geometry and Topology 10 (2006) 1635–1747.

[16] H.G. Feichtinger, On a new Segal algebra, Monatshefte für Mathematik 92 (1981) 269–289.

[17] H.G. Feichtinger, Modulation spaces on locally compact abelian groups, Technical Report, january 1983; also in: R. Radha, M. Krishna, and S. Thangavelu, editors, Proc. Internat. Conf. on Wavelets ans Applications, p. 1–56, Chennai, January 2002, 2003. New Delhi Allied Publishers.

[18] P. Flandrin, Maximum signal energy concentration in a time-frequency domain in: Acoustics, Speech, and Signal Processing (1988) Volume 4 (1988) 2176–2179, DOI 10.1109/ICASSP (1988).

[19] G.B. Folland, Harmonic Analysis in Phase space, Annals of Mathematics studies, Princeton University Press Princeton N.J (1989).

[20] G. B. Folland and A. Sitaram, The Uncertainty Principle: A Mathematical Survey, Journal of Fourier Analysis and Applications 3(3) (1997) 207–238.

[21] G. Giedke, J. Eisert, J.I. Cirac, and M.B. Plenio, Entanglement transformations of pure Gaussian states, e-print quant-ph/0301038, v1 (2003).

[22] H. Goldstein, Classical Mechanics. Addison–Wesley (1950); 2nd edition (1980) 3d edition (2002).

[23] M. de Gosson, The symplectic camel and phase space quantization, Journal of Physics A: Mathematics and General 34 (2001) 10085–10096.

[24] M. de Gosson, The Principles of Newtonian and Quantum Mechanics: the Need for Planck's Constant h; with a foreword by B. Hiley, Imperial College Press and World Scientific (2001).

[25] M. de Gosson, The "symplectic camel principle" and semiclassical mechanics, Journal of Physics A: Mathematics and General 35(32) (2002), 6825–6851.

[26] M. de Gosson, A class of symplectic ergodic adiabatic invariants,in: Foundations of probability and physics, 2 (Växjö, 2002) 151–158. Växjö Univ. Press, Växjö (2003).

[27] M. de Gosson, Phase Space Quantization and the Uncertainty Principle. Physics Letters A, 317(5-6) (2003) 365–369.

[28] M. de Gosson, The optimal pure Gaussian state canonically associated to a Gaussian quantum state. Physics Letters A 330(3–4) (2004) 161–167.

[29] M. de Gosson, Cellules quantiques symplectiques et fonctions de Husimi–Wigner, Bulletin des Sciences Mathématiques 129 (2005) 211–226.

[30] M. de Gosson, Uncertainty principle, phase space ellipsoids, and Weyl calculus, in: Operator Theory: Advances and applications Vol. 164, 121–132, Birkhäuser Verlag, Basel (2006).

[31] M. de Gosson, The adiabatic limit for multi-dimensional Hamiltonian systems, Journal of Geometry and Symmetry in Physics 4 (2005–2006) 19–44.

[32] M. de Gosson, Symplectic Geometry and Quantum Mechanics, series "Operator Theory: Advances and Applications" Vol. 166 (Birkhäuser, Basel, 2006).

[33] M. de Gosson, Uncertainty Principle, Phase Space Ellipsoids and Weyl calculus. Operator Theory: Advances and applications. Vol. 164, 121–132 (Birkhäuser Verlag Basel, 2006).

[34] M. de Gosson, Metaplectic Representation, Conley–Zehnder Index, and Weyl Calculus on Phase Space. Reviews in Mathematical Physics 19(8) (2007) 1149–1188.

[35] M. de Gosson, The Symplectic Camel and the Uncertainty Principle: The Tip of an Iceberg? Foundations of Physics 99 (2009) 194–214.

[36] M. de Gosson and F. Luef, Remarks on the fact that the uncertainty principle does not characterize the quantum state. Physics Letters A. 364 (2007) 453–457.

[37] M. de Gosson and F. Luef, Quantum States and Hardy's Formulation of the Uncertainty Principle: a Symplectic Approach. Letters in Mathematical Physics 80 (2007) 69–82.

[38] M. de Gosson and F. Luef, Principe d'Incertitude et Positivité des Opérateurs à Trace ; Applications aux Opérateurs Densité. Annales de l'Institut Henri Poincaré 9(2) (2008) 329–346.

[39] K. Gröchenig, Foundations of Time-Frequency Analysis. Applied Numerical Harmonic Analysis (Birkhäuser, Boston, 2001).

[40] K. Gröchenig and G. Zimmermann, Hardy's theorem and the short-time Fourier transform of Schwartz functions, Journal of the London Mathematical Society 2(63) (2001) 205–214.

[41] M. Gromov, Pseudoholomorphic curves in symplectic manifolds. Inventiones Mathematica 82 (1985) 307–347.

[42] A. Grossmann, G. Loupias, and E.M. Stein, An algebra of pseudo-differential operators and quantum mechanics in phase space, Annales de l' Institut Fourier (Grenoble) 18(2) (1968) 343–368.

[43] M.J.W. Hall and M. Reginatto, Schrödinger equation from an exact uncertainty principle, Journal of Physics A: Mathematics and General 35 (2002) 3289–3303.

[44] L. Guth. Symplectic embeddings of polydisks, Inventiones Mathematica 172(3) (2008) 477–489.

[45] G.H. Hardy, A theorem concerning Fourier transforms, Journal of the London Mathematical Society 8 (1933) 227–231.

[46] M. Hillery, R.F. O'Connell, M.O. Scully, and E.P. Wigner, Distribution Functions In Physics: Fundamentals, Physics Reports (Review Section of Physics Letters) 40 123–167 (1984).

[47] H. Hofer and E. Zehnder, Symplectic Invariants and Hamiltonian Dynamics, Birkhäuser Advanced texts (Basler Lehrbücher), Birkhäuser Verlag (1994).

[48] F.-Y. Hsiao and D.J. Scheeres, Fundamental Constraints on Uncertainty Relations in Hamiltonian Systems, IEEE Transactions In Automatic Control 52(4) (2007) 686–691.

[49] K. Husimi. Some Formal Properties of the Density Matrix, Proceedings of the Physico-Mathematical Society of Japan 22 (1940) 264.

[50] A.J.E.M. Janssen, Positivity and spread of bilinear time-frequency distributions. In: The Wigner distribution, pages 1–58. Elsevier, Amsterdam, (1997).

[51] F. John, Extremum problems with inequalities as subsidiary conditions, Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, (1948), Interscience Publishers, Inc., New York N.Y. 187–204 (1948).

[52] D. Kastler, The C^*-Algebras of a Free Boson Field, Communication in Mathematical Physics 1 (1965) 14–48.

[53] A. Katok, Ergodic perturbations of degenerate integrable Hamiltonian systems. Izv. Akad. Nauk. SSSR Ser. Mat. 37 539–576, (1973) [Russian]. English translation: Math. USSR izvestija 7 (1973) 535–571.

[54] Y. Katznelson, An Introduction to Harmonic Analysis (Dover, New York, 1976).

[55] J.B. Keller, Corrected Bohr-Sommerfeld Quantum Conditions for Nonseparable Systems. Annals of Physics 4 (1958) 180–188.

[56] E.H. Kennard, Zur Quantenmechanik einfacher Bewegungstypen, Zeitschrift für Physik A Hadrons and Nuclei 44(4–5) (1927) 326–352.

[57] L.D. Landau and E.M. Lifshitz, Mechanics, 3rd. ed., Pergamon Press (1976).

[58] J. Leray, Lagrangian Analysis and Quantum Mechanics, a mathematical structure related to asymptotic expansions and the Maslov index (the MIT Press, Cambridge, Mass, 1981); translated from Analyse Lagrangienne RCP 25, Strasbourg Collège de France, (1976–1977).

[59] J. Leray, The meaning of Maslov's asymptotic method the need of Planck's constant in mathematics. In Bulletin of the American Mathematical Socity, Symposium on the Mathematical Heritage of Henri Poincaré, (1980).

[60] G. Lindblad, Gaussian quantum stochastic processes on the CCR algebra, Journal of Mathematical Physics 20 (1979) 2081–2087.

[61] R.G. Littlejohn, The semiclassical evolution of wave packets. Physics Reports 138(4–5) (1986) 193–291.

[62] G. Loupias and S. Miracle-Sole, C^*-Algèbres des systèmes canoniques, I. Communication on Mathematical Physics 2 (1966) 31–48.

[63] G. Loupias and S. Miracle-Sole, C^*-Algèbres des systèmes canoniques, II. Annales de l'Institut Henri Poincaré 6(1) (1967) 39–58.

[64] F. Luef and Yu. Manin, Quantum theta functions and Gabor frames for modulation spaces. Letters in Mathematical Physics (2009) (to appear).

[65] S. Luo, Heisenberg uncertainty relation for mixed states, Physical Reviews A 72 (2005) 042110–042112.

[66] S.L. Luo, Quantum versus classical uncertainty, Theoretical and Mathematical Physics 143(2) (2005) 681–688.

[67] S. Luo and Z. Zhang, An Informational Characterization of Schrödinger's Uncertainty relations, Journal of Statistical Physics, 114(516) (2004) 1557–1576.

[68] J.M. Maruskin, D.J. Scheeres, and A.M. Bloch, Dynamics of Symplectic Subvolumes, SIAM Journal of Applied Dynamical Systems 8(1) (2009) 180–201.

[69] V.P. Maslov, Théorie des Perturbations et Méthodes Asymptotiques. Dunod, Paris, 1972; translated from Russian [original Russian edition 1965].

[70] V.P. Maslov and M.V. Fedoriuk, Semi-Classical Approximations in Quantum Mechanics. Reidel, Boston, 1981.

[71] D. McDuff and D. Salamon, Symplectic Topology, Oxford Science Publications, (1998).

[72] F.J. Narcowich, Conditions for the convolution of two Wigner distributions to be itself a Wigner distribution. Journal of Mathematical Physics 29(9) (1988) 2036–2041.

[73] F.J. Narcowich, Distributions of $\hbar$-positive type and applications. Journal of Mathematical Physics 30(11) (1989), 2565–2573.

[74] F.J. Narcowich, Geometry and uncertainty, Journal of Mathematical Physics 31(2) (1990) 354–364.

[75] F.J. Narcowich and R.F. O'Connell, Necessary and sufficient conditions for a phase-space function to be a Wigner distribution, Physical Reviews A 34(1) (1986) 1–6.

[76] L. Polterovich, The Geometry of the Group of Symplectic Diffeomorphisms, Lectures in Mathematics, Birkhäuser (2001).

[77] M. Reed and B. Simon, Methods of Modern Mathematical Physics. Academic Press, New York, 1972.

[78] E. Samuel Reich, How camels could explain quantum uncertainty, New Scientist 2697 (2009), 12.

[79] H.P. Robertson, The uncertainty principle. Physical Reviews 34 (1929) 163–164.

[80] D.J. Scheeres, F.-Y. Hsiao, R.S. Park, B.F. Villac, and J.M. Maruskin, Fundamental Limits on Spacecraft Orbit Uncertainty and Distribution Propagation, Journal of the Astronautical Sciences 54 (2006) 505–523.

[81] F. Schlenk, Symplectic embeddings of ellipsoids, Israel Journal of Mathematics 138 (2003) 215–252.

[82] F. Schlenk, Embedding problems in symplectic geometry. De Gruyter Expositions in Mathematics, vol. 40, Berlin (2005).

[83] E. Schrödinger, Zum Heisenbergschen Unschärfeprinzip, Berliner Berichte (1930), 296–303 [English translation: A. Angelow, M.C. Batoni, About Heisenberg Uncertainty Relation. Bulgarian Journal of Physics 26 5/6 (1999) 193–203, and http://arxiv.org/abs/quant-ph/9903100.

[84] D. Shale, Linear symmetries of free Boson fields. Transactions of the American Mathematical Socity 103 (1962) 149–167.

[85] M.A. Shubin, Pseudodifferential Operators and Spectral Theory (Springer 1987) [original Russian edition in Nauka, Moskva (1978)].

[86] K.F. Siburg, Symplectic capacities in two dimensions, Manuscripta Mathematica 78 (1993) 149–163.

[87] J.-C. Sikorav, Quelques propriétés des plongements lagrangiens (Some properties of Lagrangian embeddings), Mémoires de la Société Mathématique de France (Nouv. Sér.) 46 151 (1991).

[88] R. Simon, N. Mukunda, and N. Dutta, Quantum Noise Matrix for Multimode Systems: U(n)-invariance, squeezing and normal forms. Physical Reviews A 49 (1994) 1567–1583.

[89] R. Simon, E.C.G. Sudarshan, N. Mukunda, Gaussian–Wigner distributions in quantum mechanics and optics. Physical Reviews A 36(8) (1987) 3868–3880.

[90] E.C.G. Sudarshan, Charles B. Chiu, and G. Bhamati, Generalized uncertainty relations and characteristic invariant for the multimode states. Physical Reviews A 52(1) (1995) 43–54.

[91] I. Stewart, The Symplectic Camel, Nature 329 (1987) 17–18.

[92] K. Weierstrass, Mathematische Werke, Berlin, Band I: 233–46, Band II:19–44, Nachtrag: 139–48 (1858).

[93] D. Wick, The Infamous Boundary: Seven Decades of Heresy in Quantum Mechanics, Birkhäuser, Boston (1995).

[94] E. Wigner, On the quantum correction for thermodynamic equilibrium, Physical Revues. 40 (1932) 799–755.

[95] J. Williamson, On the algebraic problem concerning the normal forms of linear dynamical systems. American Journal of Mathematics 58 (1936) 141–163.

[96] H.P. Yuen, Multimode two-photon coherent states, in International Symposium on Spacetime Symmetries, edited by Y.S. Kim and W.W. Zachary (North-Holland, Amsterdam, 1989).