rijksuniversiteit groningen
faculteit Wiskunde en Natuurwetenschappen

The Symplectic Camel

Bacheloronderzoek Wiskunde en Natuurkunde

Augustus 2013

Student: K. Frieswijk

Eerste Begeleider: Prof. dr. H. Waalkens

Tweede Begeleider: Prof. dr. D. Boer

The Symplectic Camel

K. Frieswijk
Department of Mathematical Sciences
Rijksuniversiteit Groningen

26 august 2013

Abstract

Recent advances in symplectic topology suggest that classical and quantum mechanics are much closer than might appear at first sight.
After having discussed some general properties of Hamiltonian mechanics and symplectic topology, we will show that Heisenberg's uncertainty principle has left a footprint in classical mechanics. We do this by making use of a surprising theorem called Gromov's non-squeezing theorem (1985), which states that no canonical transformation can squeeze a ball $B^{2n}(r)$ with radius r through a circular hole in a plane of conjugate coordinates x_j, p_j with smaller radius $R < r$. This theorem was lovingly nicknamed "the symplectic camel".

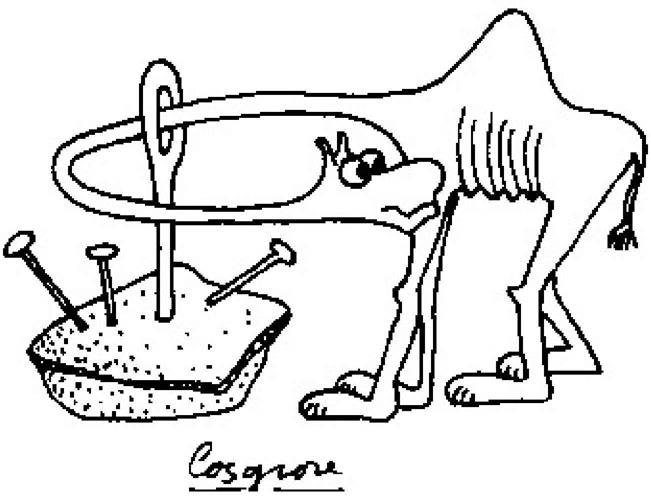

Figure 1: Gromov's symplectic camel [1]

Contents

1 Introduction

William Rowan Hamilton discovered in the nineteenth century that Newton's laws of physics have an elegant geometric interpretation if every moving particle is seen as moving in phase space. This realization was the root of a new field of study called *symplectic topology*.
In phase space the particle is described by both its position $\mathbf{x}$ and momentum $\mathbf{p}$, while in regular space the particle is only described by its position $\mathbf{x}$. Thus, a moving particle with n degrees of freedom in regular space, moves in a $2n$-dimensional phase space, i.e. $\mathbb{R}^n \times \mathbb{R}^n \equiv \mathbb{R}^{2n}$. For example, if the system consists of N moving point-like particles in 3-dimensional space, we have $n = 3N$. Thus, the phase space for this system is $\mathbb{R}^{3N} \times \mathbb{R}^{3N} \equiv \mathbb{R}^{6N}$. Symplectic manifolds are a generalization of the phase space of a closed system.
The concept of phase space and symplectic structures arose in the study of classical mechanical systems, such as an oscillating pendulum or a planet orbiting the sun. If one knows the position and the momentum of such a system at one time, then the trajectory of this system can be determined.

Classical mechanics and quantum mechanics are subfields of the branch of physics called mechanics. While quantum mechanics deals with the microscopic world, classical mechanics succesfully describes macroscopic systems. In physics, Bohr's correspondence principle states that if large quantum numbers are considered, then the laws of quantum mechanics will appear to obey the laws of classical mechanics. The transition region between quantum and classical mechanics is still a field of current research among physicists and mathematicians.

An example of a phenomenon that classical mechanics cannot account for is Heisenberg's uncertainty principle, which states that it is impossible to measure the momentum and the position of a particle precisely at the same time. For one degree of freedom, the uncertainty principle is mathematically expressed as

$$\Delta x \Delta p \geq \frac{1}{2}\hbar, \tag{1}$$

where x and p are the position and the momentum of the particle respectively.
Since it is impossible to know both the position and momentum to an arbitrary degree of accuracy, a particle should be thought of as lying in a region of the phase plane, instead of occupying a single point. One can think of these regions in phase plane as a measure of the incompatibility of position and momentum.
Hamilton realized that in phase space, Newton's laws preserve area under time evolution; i.e. if you define the region S_1 as the set of all possible initial positions and velocities for the moving particle, then at any later time its set of possible positions and velocities will form a region S_2 with the same area, although it may be highly distorted.

The field of symplectic geometry has had some recent developments, yet the symplectic side to the nature of various fields of mathematics and physics has yet to be brought out of the closet. It was stated in [1] that we are witnessing just the tip of the symplectic iceberg, but so far this so-called symplectic iceberg has not received the attention it deserves.
The recent advances in symplectic geometry and topology suggest that classical and quantum mechanics are much closer than might appear at first sight; by means of Gromov's non-squeezing theorem (1985) one can show that Heisenberg's uncertainty principle has left a footprint in classical mechanics. In this essay the analogue of the uncertainty principle in classical mechanics is explored, where chapters 3 and 4 are largely based upon an article by Maurice de Gosson [2].

2 Hamiltonian Mechanics

Lagrangian mechanics describes motion in mechanical systems by means of the configuration space. The configuration space of a mechanical system has the structure of a differentiable manifold, on which its group of diffeomorphisms acts. A lagrangian mechanical system is given by a manifold ("configuration space") and a function on its tangent bundle ("the lagrangian function").
A natural mechanical system is a particular case of a lagrangian system; the configuration space in this case is Euclidean. The lagrangian function for a natural mechanical system is given by the difference between the kinetic and potential energies; $\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}, t) = \mathcal{T} - \mathcal{U}$, where x_i are the generalized coordinates, $\dot{x}_i$ are generalized velocities, $\partial \mathcal{L}/\partial \dot{x}_i = p_i$ are generalized momenta and $\partial \mathcal{L}/\partial x_i$ are generalized forces. The Euler-Lagrange equation is given by

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i}\right) - \frac{\partial \mathcal{L}}{\partial x_i} = 0. \tag{2}$$

Lagrangian mechanics is equivalent to Hamiltonian mechanics [3].
By means of a Legendre transformation, a lagrangian system of n second-order differential equations can be converted into a remarkably symmetrical system of $2n$ first-order equations called a hamiltonian system of equations (or canonical equations).

A Legendre transformation is defined as follows. Let $y = f(x)$ be a convex function, $f''(x) > 0$. The Legendre transformation of the function f is a new function g of a new variable p, which is constructed in the following way.
Consider the straight line $y = px$ in the x, y plane, where p is a given number. We take the point $x = x(p)$ at which the curve f is farthest from the straight line $y = px$ in the vertical direction. For each p the function $px - f(x) = F(p, x)$ has a maximum with respect to x at the point $x(p)$, so the point $x(p)$ is defined by the extremal condition; $\partial F/\partial x = 0$. Since f is convex, the point $x(p)$ is unique, if it exists. Now we define the Legendre transform as $g(p) = F(p, x(p)) = px(p) - f(x(p))$.

Consider the system of Lagrange's equations $\dot{\mathbf{p}} = \partial\mathcal{L}/\partial\mathbf{x}$, where $\mathbf{p} = \partial\mathcal{L}/\partial\dot{\mathbf{x}}$, with a given lagrangian function $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, which we will assume to be convex with respect to the second argument $\dot{\mathbf{x}}$.

Theorem 1. *The system of Lagrange's equations is equivalent to the system of* $2n$ *first-order equations (Hamilton's equations)*

$$\dot{\boldsymbol{p}} = -\frac{\partial\mathcal{H}(\boldsymbol{x},\boldsymbol{p},t)}{\partial\boldsymbol{x}}, \quad \dot{\boldsymbol{x}} = \frac{\partial\mathcal{H}(\boldsymbol{x},\boldsymbol{p},t)}{\partial\boldsymbol{p}}. \tag{3}$$

Proof. By definition, the Legendre transform of $\mathcal{L}(\mathbf{x},\dot{\mathbf{x}},t)$ with respect to $\dot{\mathbf{x}}$ is the function $\mathcal{H}(\mathbf{p},\mathbf{x},t) = \mathbf{p}\dot{\mathbf{x}} - \mathcal{L}(\mathbf{x},\dot{\mathbf{x}},t)$, in which $\dot{\mathbf{x}}$ is expressed in terms of $\mathbf{p}$ by the formula $\mathbf{p} = \partial\mathcal{L}/\partial\dot{\mathbf{x}}$, and which depends on the parameters $\mathbf{x}$ and t. This function $\mathcal{H}$ is called the hamiltonian.

The total differential of the hamiltonian

$$d\mathcal{H} = \frac{\partial\mathcal{H}}{\partial\mathbf{p}}d\mathbf{p} + \frac{\partial\mathcal{H}}{\partial\mathbf{x}}d\mathbf{x} + \frac{\partial\mathcal{H}}{\partial t}dt$$

is equal to the total differential of $\mathcal{H}(\mathbf{p},\mathbf{x},t) = \mathbf{p}\dot{\mathbf{x}} - \mathcal{L}(\mathbf{x},\dot{\mathbf{x}},t)$;

$$d\mathcal{H} = \dot{\mathbf{x}}d\mathbf{p} - \frac{\partial\mathcal{L}}{\partial\mathbf{x}}d\mathbf{x} + \frac{\partial\mathcal{L}}{\partial t}dt.$$

Both equations for $d\mathcal{H}$ must be the same. Therefore,

$$\dot{\mathbf{x}} = \frac{\partial\mathcal{H}}{\partial\mathbf{p}}, \; \frac{\partial\mathcal{H}}{\partial\mathbf{x}} = -\frac{\partial L}{\partial\mathbf{x}} = -\dot{\mathbf{p}}, \; \frac{\partial\mathcal{H}}{\partial t} = -\frac{\partial L}{\partial t}. \tag{4}$$

Thus, if we are working with a time-invariant system, we obtain Hamilton's equations.

□

Hamilton's equations are equivalent to Newton's second law;

$$F = m\ddot{\mathbf{x}}. \tag{5}$$

Suppose now that we have a natural mechanical system, so the lagrangian has the usual form $\mathcal{L} = \mathcal{T} - \mathcal{U}$, where the kinetic energy $\mathcal{T} = \mathcal{T}(\mathbf{x},\dot{\mathbf{x}})$ and the potential energy $\mathcal{U} = \mathcal{U}(\mathbf{x})$.

For example, if $\mathcal{T}$ is given by $\mathcal{T} = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$ and $\dot{\mathbf{x}} = (\dot{x},\dot{y},\dot{z})$, then

$$\begin{aligned}
\mathcal{H} &= \mathbf{p}\dot{\mathbf{x}} - \mathcal{L} \\
&= \frac{\partial\mathcal{L}}{\partial\dot{\mathbf{x}}}\dot{\mathbf{x}} - \mathcal{L} \\
&= \frac{\partial\mathcal{T}}{\partial\dot{\mathbf{x}}}\dot{\mathbf{x}} - \mathcal{L} \\
&= m\dot{x}^2 + m\dot{y}^2 + m\dot{z}^2 - (\mathcal{T} - \mathcal{U}) \\
&= 2\mathcal{T} - \mathcal{T} + \mathcal{U} \\
&= \mathcal{T} + \mathcal{U}. && (6)
\end{aligned}$$

Thus, under the given assumptions, the hamiltonian $\mathcal{H}$ is the total energy function $\mathcal{H} = \mathcal{T} + \mathcal{U}$, which is a constant of motion.
In this thesis, we will confine ourselves to natural mechanical systems, so the hamiltonian is given by the sum of the kinetic and potential energy.
For the sake of convenience, we convert equation (5) to the following equation; $F(\mathbf{x}) = \ddot{\mathbf{x}}$. In this case, the velocity of the particle equals its momentum.

3 Symplectic Topology

In this section we will introduce some notation and terminology used within the field of symplectic topology; a branch of differential geometry and differential topology which studies symplectic manifolds. The structure of symplectic geometry is discussed in [4] and [5].
A symplectic manifold is a pair (M, ω), where M is a differentiable manifold and ω is a 2-form defined as

$$\omega : M \times M \to \mathbb{R}. \tag{7}$$

A symplectic manifold M is even-dimensional, smooth, and orientable. One of the most intriguing aspects of symplectic topology is its curious mixture of rigidity (structure) and lack of structure.
The symplectic form ω is:

- linear in each of its components:

$$\begin{aligned} \omega(\alpha \mathbf{z}_1 + \beta \mathbf{z}_2, \mathbf{z}') &= \alpha\omega(\mathbf{z}_1, \mathbf{z}') + \beta\omega(\mathbf{z}_2, \mathbf{z}') \\ \omega(\mathbf{z}', \alpha \mathbf{z}_1 + \beta \mathbf{z}_2) &= \alpha\omega(\mathbf{z}', \mathbf{z}_1) + \beta\omega(\mathbf{z}', \mathbf{z}_2) \end{aligned} \tag{8}$$

  for all $\alpha, \beta \in \mathbb{R}$ and all $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}' \in TM$, where TM denotes the tangent bundle of M;

- antisymmetric:

$$\omega(\mathbf{z}, \mathbf{z}') = -\omega(\mathbf{z}', \mathbf{z}) \text{ for all } \mathbf{z}, \mathbf{z}' \in TM; \tag{9}$$

- non-degenerate:

$$\omega(\mathbf{z}, \mathbf{z}') = 0 \text{ for all } \mathbf{z} \in TM \text{ if and only if } \mathbf{z}' = 0; \tag{10}$$

- closed:

$$d\omega = 0. \tag{11}$$

The fact that the symplectic form ω is non-degenerate implies that for each nonzero tangent direction $\mathbf{z}$ there is another direction $\mathbf{z}'$ such that the area $\omega(\mathbf{z}, \mathbf{z}')$ of the little parallelogram spanned by these vectors is nonzero.
According to Darboux's theorem, the closedness condition forces all symplectic structures to be locally indistinguishable; any two symplectic manifolds of the same dimension are locally symplectomorphic to one another and their only distinguishing characteristics are large-scale.

A particle moving in a system with configuration space $\mathbb{R}^n$ has n position coordinates $x_1, ..., x_n$ and n corresponding velocity coordinates $p_1 = \dot{x}_1, ..., p_n = \dot{x}_n$. The position $\mathbf{x}$ is given by $\mathbf{x} = (x_1, ..., x_n)$, and the momentum $\mathbf{p}$ is given by $\mathbf{p} = (p_1, ..., p_n)$, so $\mathbf{z} = (\mathbf{x}, \mathbf{p}) \in \mathbb{R}^{2n}$ describes the particle. Whenever matrix calculations are performed, $\mathbf{x}$, $\mathbf{p}$, and $\mathbf{z}$ are viewed as column vectors.

The symplectic form ω measures the area of 2-dimensional surfaces S in $\mathbb{R}^{2n}$ by adding the areas of the projections of S onto the (x_j, p_j)-plane, where $j = 1, ..., n$. For example, for the classical phase space $(\mathbb{R}^{2n}, \omega_0)$, the symplectic form is given by the sum of contributions for each of the n pairs of directions:

$$\omega_0 = \sum_{j=1}^{n} dx_j \wedge dp_j. \tag{12}$$

ω_0 is known as the standard symplectic form on Euclidean space and is conserved under Hamiltonian flows; i.e. the sum over the symplectic areas in the phase planes of the conjugate pairs (x_j, p_j) remains the same.

If the momentum p_j of a moving particle is treated as an imaginary position coordinate in a symplectic manifold, then the symplectic form almost gives the space the structure of a complex manifold, i.e. a manifold whose functions have a real and imaginary part. Herman Weyl realized the strong tie between complex differential geometry and symplectic geometry when he coined the word *symplectic* in 1938 by converting the Latin roots *com-* and *plex-* to their Greek equivalents.

The standard symplectic matrix is given by

$$J = \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix}. \tag{13}$$

Note that;

$$J^2 = -I_{2n} \text{ and } J^T = J^{-1} = -J. \tag{14}$$

A real $2n \times 2n$ matrix S is called symplectic if it satisfies the condition

$$S^T J S = J. \tag{15}$$

If we write the matrix S in the following block-form

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{16}$$

where A, B, C and D are $n \times n$-matrices, then (15) is equivalent to;

$$A^T C = C^T A, \; B^T D = D^T B, \; A^T D - C^T B = I_n. \tag{17}$$

The determinant of a symplectic matrix S is one, so S is nonsingular with $S^{-1} = J^T S^T J$. Furthermore, both the inverse S^{-1} and the transpose S^T are symplectic as well; after rewriting $S^T J S = J$ as $JS = (S^{-1})^T J$, it follows that $(S^{-1})^T J S^{-1} = J$. Hence S^{-1} is symplectic.

In order to see that the transpose S^T is symplectic, it suffices to take the inverse of the equality $(S^{-1})^T J S^{-1} = J$, where we make use of the fact that $(S^T)^T = S$:

$$\begin{aligned} [(S^{-1})^T J S^{-1}]^{-1} &= J^{-1} \\ (S^{-1})^{-1} J^{-1} S^T &= -J \\ -SJS^T &= -J \\ (S^T)^T J S^T &= J. \end{aligned} \tag{18}$$

Since a matrix is symplectic if and only if its transpose is, equation (15) is equivalent to

$$SJS^T = J. \tag{19}$$

If we suppose that S and S' are two symplectic matrices, then the product SS' of two symplectic matrices is also a symplectic matrix:

$$(SS')^T J(SS') = S'^T (S^T JS)S' = S'^T JS' = J. \tag{20}$$

The group of symplectic $2n \times 2n$ matrices is denoted by $\mathrm{Sp}(2n, \mathbb{R})$. The associated symplectic form of J is given by $\omega(\mathbf{z}, \mathbf{z}') = (\mathbf{z}')^T J\mathbf{z}$. A 2-form is called symplectic if $\omega(S\mathbf{z}, S\mathbf{z}') = \omega(\mathbf{z}, \mathbf{z}')$ for all vectors $\mathbf{z}, \mathbf{z}' \in TM$.

3.1 Gromov's Non-squeezing Theorem

A Hamiltonian phase flow consists of canonical transformations.

Definition 1. *If a transformation $f(\boldsymbol{x}, \boldsymbol{p}) = (\boldsymbol{x}', \boldsymbol{p}')$ of phase space $\mathbb{R}^{2n}$ is canonical, then its Jacobian matrix*

$$Df(\boldsymbol{x}, \boldsymbol{p}) = \frac{\partial(\boldsymbol{x}', \boldsymbol{p}')}{\partial(\boldsymbol{x}, \boldsymbol{p})} \tag{21}$$

calculated at any phase space point $(\boldsymbol{x}, \boldsymbol{p})$ were f is defined, is symplectic.

In 1838, Joseph Liouville proved that, for multiple degrees of freedom, a Hamiltonian flow f_t^H is volume preserving. This is known as Liouville's theorem; one of the best known results from elementary statistical mechanics. The proof for this theorem consists of the fact that a Hamiltonian flow consists of canonical transformations, so the Jacobian matrix of f_t^H is symplectic at each point and has therefore a determinant equal to one. In this thesis, the terms "canonical" and "symplectic" are used equivalently and the notions "transformations" and "maps" are used synonymously.

Gromov's non-squeezing theorem [6] is a considerable refinement of Liouville's theorem on conservation of phase space volume under canonical transformations. Gromov showed that it is impossible for a Hamiltonian phase flow to squeeze a ball into a cylinder of a smaller radius. While a volume-preserving map can do this easily, a symplectic map cannot.

Theorem 2 (Gromov's non-squeezing theorem)**.** *No canonical transformation can squeeze a ball $B^{2n}(r)$ with radius r through a circular hole in a plane of conjugate coordinates x_j, p_j with smaller radius $R < r$.*

Consider a system $\mathcal{S}$ of N particles, where the particles are very close to each other and the amount N of particles is very high. In this case, we may approximate the system $\mathcal{S}$ with a "cloud" of points in phase space $\mathbb{R}^{2n}$. If we assume that this cloud is initially spherical, then it is represented by a phase space ball $B^{2n}(r)$ with radius r and center $(\mathbf{a}, \mathbf{b})$:

$$B^{2n}(r) : |\mathbf{x} - \mathbf{a}|^2 + |\mathbf{p} - \mathbf{b}|^2 \leq r^2. \tag{22}$$

The orthogonal projection $\Delta x_j \Delta p_j$ of this ball on any plane of conjugate phase space coordinates x_j, p_j will always be defined by a circle with area πr^2. As time passes, the motion of this phase space cloud will distort the spherical shape, while the volume remains the same by Liouville's theorem. Eventually, the cloud gets very thinly spread out over huge regions of phase space, since the ball $B^{2n}(r)$ can be stretched in all directions by Hamiltonian phase flows. This especially holds for systems with a large number of degrees of freedom, since this results in a high amount of directions in which the cloud can spread. Conventional wisdom suggests that the projections on any plane x_j, p_j could thus become arbitrarily small after a certain amount of time. However, this turned out to be wrong.
In 1985, Mikhael Gromov came to the surprising conclusion that the projections of the deformed ball on any plane of conjugate phase space coordinates x_j, p_j will never decrease below its original value of πr^2. This is known as Gromov's non-squeezing theorem. By contrast, had we chosen a plane of non-conjugate coordinates (for example, x_1, p_2 or x_1, x_2), then the projection on the plane could become arbitrarily small.

Hence, it is impossible to deform a phase space ball $B^{2n}(r)$ by using canonical transformations in such a way that it can be orthogonally squeezed through a hole in a conjugate phase plane x_j, p_j, if the area of that hole is smaller than the cross-section of the ball. Therefore, Gromov's non-squeezing theorem shows that arbitrary spreading of the ball in phase space is prevented.
The same result is obtained if we replace the hole in the conjugate phase plane with the base of a symplectic cylinder $Z^{2n}(R)$ (see figure 2), where

$$Z^{2n}(R) \equiv B^2(R) \times \mathbb{R}^{2n-2}. \tag{23}$$

Gromov's non-squeezing theorem provides us with the hint that Heisenberg's uncertainty principle of quantum mechanics has left a footprint in classical mechanics. Suppose, for example, that the radius of $B^{2n}(r)$ is given by $r = \sqrt{\hbar/2\pi}$. Then the projection on a conjugate phase plane x_j, p_j is defined by

$$\Delta x_j \Delta p_j \geq \pi r^2 = \pi \left(\sqrt{\hbar/2\pi}\right)^2 = \frac{1}{2}\hbar, \tag{24}$$

which has quite some resemblance to Heisenberg's uncertainty principle.

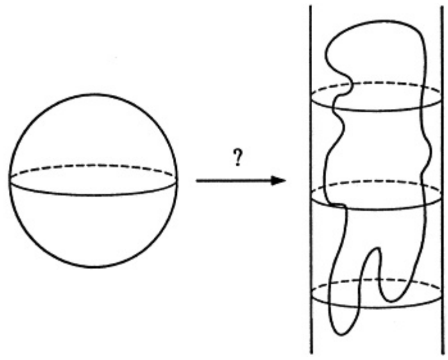

Figure 2: Gromov's non-squeezing theorem states that if there is a symplectic embedding $B^{2n}(r) \hookrightarrow Z^{2n}(R)$, then $R \geq r$. [7]

Vladimir Arnold, a well-known mathematical physicist, nicknamed Gromov's non-squeezing theorem "the symplectic camel", referring to the Biblical phrase
"*...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...*" Matthew 19[24]

In this metaphor, the eye of the needle represents the hole in the x_j, p_j plane and the camel represents the phase space ball. A camel cannot pass through the eye of a needle because its fattest cross section (through one of its humps) cannot be shrunk to the size of the eye.

3.2 Symplectic Capacity

In 1990, Ekeland and Hofer introduced the notion of symplectic capacities [8]. Using canonical transformations, an arbitrary volume Ω in phase space $\mathbb{R}^{2n}$ cannot be squeezed into a cylinder whose base $B^2(R)$ is smaller than the symplectic capacity of the volume. Ω may be bounded or unbounded, large or small.

The symplectic capacity is any function that associates to Ω a non-negative number $c(\Omega)$, or $+\infty$, and for which the properties i), ii), iii) and iv) are verified [9]:

i) A symplectic capacity is a symplectic invariant;

$$c(f(\Omega)) = c(\Omega) \text{ if } f \text{ is canonical.} \tag{25}$$

ii) It is also monotone:

$$c(\Omega) \leq c(\Omega') \text{ if } \Omega \subseteq \Omega' \tag{26}$$

iii) and 2-homogeneous under phase space dilations:

$$c(k\Omega) = k^2 c(\Omega) \text{ for all } k \in \mathbb{R}, \tag{27}$$

where $k\Omega$ consists of all points $k\mathbf{z}$ such that $\mathbf{z}$ is in Ω.

iv) Furthermore,

$$c(B^{2n}(R)) = \pi R^2 = c(Z_j(R)), \tag{28}$$

where $Z_j(R)$ denotes the phase space cylinder consisting of all phase space points whose j-th position and momentum coordinate satisfy $x_j^2 + p_j^2 \leq R^2$.

There is an infinite number of symplectic capacities, but every symplectic capacity c of an arbitrary volume Ω in $\mathbb{R}^{2n}$ lies between a minimal and a maximal symplectic capacity;

$$c_{\text{min}}(\Omega) \leq c(\Omega) \leq c_{\text{max}}(\Omega). \tag{29}$$

The minimal capacity c_{min} is also known as the Gromov capacity [6] and is calculated as follows. If there does not exist a canonical transformation which sends the phase space ball $B^{2n}(r)$ inside Ω, then $c_{\text{min}}(\Omega) = 0$. If, on the other hand, there do exist such canonical transformations, then the minimal symplectic radius of Ω is defined as the supremum R of all the radii r for which this is possible. No canonical transformation will send a phase space ball with radius $r > R$ inside Ω, but one can find canonical transformations sending $B^{2n}(r)$ inside Ω for all $r \leq R$.
We define the minimal capacity of Ω by

$$c_{\text{min}}(\Omega) = \pi R^2. \tag{30}$$

We will now show that the minimal capacity satisfies formula (28).
The equality $c_{\text{min}}(Z_j(R)) = \pi R^2$ is just a reformulation of Gromov's non-squeezing theorem. Since it is impossible to squeeze a ball $B^{2n}(r)$ with radius $r > R$ into the cylinder $Z_j(R)$, it follows that $r \leq R$. Thus, the orthogonal projection of the deformed ball on any plane of conjugate phase space coordinates x_j, p_j would be less or equal to πR^2. Since c_{min} is defined as the supremum of the possible radii, $c_{\text{min}}(Z_j(R)) = \pi R^2$. The equality $c_{\text{min}}(B^{2n}(R)) = \pi R^2$ is trivial, since it follows directly from the definition of the minimal symplectic capacity.

We can calculate the maximal capacity of Ω, denoted by $c_{\text{max}}(\Omega)$, in the following way. If there does not exist a canonical transformation which sends Ω inside a cylinder $Z_j(r)$, no matter how large we choose the radius r of the cylinder, then $c_{\text{max}} = +\infty$. If, on the other hand, there do exist such canonical transformations, then the maximal symplectic radius of Ω is defined as the infimum R of all the radii r for which this is possible.
The maximal symplectic capacity of Ω is defined by

$$c_{\text{max}}(\Omega) = \pi R^2. \tag{31}$$

By the formulas of the minimal and maximal capacity, we can see that c_{max} and c_{max} both have the dimension of an area. Property (27) plus the fact that $c(B^{2n}(R)) = \pi R^2$ suggest that symplectic capacities in general have something to do with the notion of area.
If an arbitrary volume Ω in $\mathbb{R}^{2n}$ is connected, then $c_{\text{min}}(\Omega)$ defines the area and if Ω is disconnected, $c_{\text{min}}(\Omega)$ does not.

If we assume that Ω is simply connected; then $c_{\max}(\Omega)$ defines the area. And if Ω is not simply connected, $c_{\max}(\Omega)$ is not the respective area. For example, suppose that Ω is an annulus. The existence of the hole in the domain of Ω is the reason that it is impossible to squeeze Ω into the cylinder $Z_j(r)$, so $c_{\max}(\Omega)$ does not define the area.
Since all symplectic capacities c lie between the minimal and the maximal capacity by formula (29), $c(\Omega)$ coincides with the area for all connected and simply connected domains.

We will now state a theorem which will be important later on, namely Williamson's diagonalization theorem [10], which was stated in 1963. According to this theorem, every symmetric, positive definite matrix M can be diagonalized by using a symplectic matrix S.
In linear algebra, a symmetric $2n \times 2n$ real matrix M is said to be positive definite if $\mathbf{z}^T M \mathbf{z}$ is positive for any non-zero column vector $\mathbf{z}$ of $2n$ real numbers, where $\mathbf{z}^T$ denotes the transpose of $\mathbf{z}$. Furthermore, all the eigenvalues of M are positive numbers.

Theorem 3 (Williamson's theorem)**.** *Let M be a symmetric positive definite real $2n \times 2n$ matrix. Then there exists a matrix $S \in Sp(2n)$ such that*

$$S^T M S = \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda \end{pmatrix}, \tag{32}$$

where $\Lambda = diag(\lambda_1, ..., \lambda_n)$ is a $n \times n$ matrix whose non-zero entries are the moduli λ_j of the eigenvalues $\pm i\lambda_j$ of JM (where $\lambda_j > 0$).
The diagonalizing symplectic matrix S is not unique and the sequence $\lambda_1, ..., \lambda_n$ does not depend, up to a reordering of its terms, on the choice of S diagonalizing M.

A very nice property of phase space ellipsoids is defined in the following lemma.

Lemma 4. *All symplectic capacities agree for phase space ellipsoids Ω_{ell}.*

Proof. A solid phase space ellipsoid is defined by the formula

$$(\mathbf{z} - \bar{\mathbf{z}})^T M^{-1} (\mathbf{z} - \bar{\mathbf{z}}) \leq 1, \tag{33}$$

where $\bar{\mathbf{z}}$ is the mean and M is a symmetric positive definite $2n \times 2n$ matrix. The eigenvalues of M are given by the squares of the semi-principal axes of the ellipsoid.
Since symplectic capacities are invariant under phase space translations, we may assume that Ω_{ell} is an ellipsoid centered at $\bar{\mathbf{z}} = \mathbf{0}$. In this case, equation (33) becomes

$$\mathbf{z}^T M^{-1} \mathbf{z} \leq 1. \tag{34}$$

Let R_j denote the length of the jth semi-principal axis of the ellipsoid. Another way to define the phase space ellipsoid centered at $\bar{\mathbf{z}} = \mathbf{0}$ is

$$\Omega_{\text{ell}} : \sum_{j=1}^{n} \frac{1}{R_j^2} (x_j^2 + p_j^2) \leq 1. \tag{35}$$

Now note that if M is a symmetric positive definite matrix, then M^{-1} is one as well. So according to Williamson's diagonalization theorem, M^{-1} can be diagonalized by means of a symplectic matrix S;

$$S^T M^{-1} S = \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix}, \tag{36}$$

where $\Sigma = \text{diag}(1/R_1^2, ..., 1/R_n^2)$. Thus, the eigenvalues of JM^{-1} are given by $\pm i\lambda_j = \pm i/R_j^2$, for $j = 1, ..., n$. Since $\lambda_j = 1/R_j^2 > 0$, we see that $R_j^2 = 1/\lambda_j$.

We now claim that

$$c(\Omega_{\text{ell}}) = \frac{\pi}{\lambda_{\text{max}}} \tag{37}$$

for every symplectic capacity c, where λ_{max} denotes the largest of all the positive numbers λ_j.
Symplectic capacities are invariant by canonical transformations, so

$$c(\Omega_{\text{ell}}) = c(S(\Omega_{\text{ell}})) \tag{38}$$

and proving formula (37) is equivalent to proving that $c(S(\Omega_{\text{ell}})) = \pi/\lambda_{\text{max}}$.

We will first show that $c_{\text{min}}(\Omega_{\text{ell}}) = \pi/\lambda_{\text{max}}$.
Suppose that there exist canonical transformations which send the phase space ball $B^{2n}(r)$ inside Ω_{ell}. We now look for the smallest semi-principal axis R_{min} of the ellipsoid Ω_{ell}, because this semi-principal axis determines the supremum of the radii of the ball $B^{2n}(r)$ for which a canonical transformation into the ellipsoid exists.
Since $R_j^2 = 1/\lambda_j$, we can easily see that we obtain the smallest axis R_j, when λ_j is largest. So $R_{\text{min}}^2 = 1/\lambda_{\text{max}}$. And thus,

$$\begin{aligned} c_{\text{min}}(\Omega_{\text{ell}}) &= \pi R_{\text{min}}^2 \\ &= \frac{\pi}{\lambda_{\text{max}}}. \end{aligned} \tag{39}$$

Now we will show that $c_{\text{max}}(\Omega_{\text{ell}}) = \pi/\lambda_{\text{max}}$.
Assume that there exist canonical transformations which send the ellipsoid Ω_{ell} into a cylinder $Z_j(r)$. This time, we are looking for the infimum of the radii of the cylinder for which a canonical transformation into the cylinder is possible, which is again determined by the smallest semi-principal axis R_{min} of the ellipsoid. Hence,

$$c_{\text{max}}(\Omega_{\text{ell}}) = \pi R_{\text{min}}^2 = \frac{\pi}{\lambda_{\text{max}}}. \tag{40}$$

Since all symplectic capacities lie between c_{min} and c_{max} by formula (29), all symplectic capacities must agree on phase space ellipsoids and are defined by equation (37).

□

4 The Uncertainty Principle

Heisenberg's uncertainty principle (1) is a particular case of the more general Schrödinger-Robertson inequality

$$\Delta x^2 \Delta p^2 \geq \mathrm{Cov}(x,p)^2 + \frac{1}{4}\hbar^2. \tag{41}$$

This is a general formula that comes from considering arbitrary random variables. In quantum mechanical systems, $\mathrm{Cov}(x,p) = 0$. So for these systems, above formula reduces to Heisenberg's uncertainty principle (1).

For n degrees of freedom, formula (41) can be generalized to

$$(\Delta x_j)^2(\Delta p_j)^2 \geq \mathrm{Cov}(x_j,p_j)^2 + \frac{1}{4}\hbar^2, \text{ for } j = 1, ..., n, \tag{42}$$

where the co-variances are expressed in terms of measure errors; $\Delta(x_j, p_j)$. $\mathrm{Cov}(x_j, p_j)$ is a measure of how much the two variables x_j, p_j are correlated. The covariance between two jointly distributed real-valued random variables x and p is defined as

$$\mathrm{Cov}(x,p) = E[xp] - E[x]E[p], \tag{43}$$

where $E[x]$ is the expectation value of x, also known as the mean of x. For n degrees of freedom, the expectation value of a variable can be expressed in integrals (in $\mathbb{R}^{2n}$);

$$E[\mathbf{x}] = \int_{-\infty}^{\infty} \mathbf{x}\rho(\mathbf{x})d^n x, \; E[\mathbf{xp}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{xp}\rho(\mathbf{x},\mathbf{p})d^n x d^n p, \tag{44}$$

where $d^n x = dx_1 \dots dx_n$, $d^n p = dp_1 \dots dp_n$ and ρ is some (here undefined) phase space probability density. For a normal distribution, ρ is defined as

$$\rho(z) = \left(\frac{1}{2\pi}\right)^n (\det\Sigma)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{z}-\bar{\mathbf{z}})^T \Sigma^{-1}(\mathbf{z}-\bar{\mathbf{z}})\right]. \tag{45}$$

If we assume, as before, that $\bar{\mathbf{z}} = (\bar{\mathbf{x}}, \bar{\mathbf{p}}) = \mathbf{0}$, then $E[\mathbf{x}] = E[\mathbf{p}] = \mathbf{0}$ and consequently $\mathrm{Cov}(\mathbf{x},\mathbf{p}) = E[\mathbf{xp}]$.
Thus, we get the following equations;

$$\begin{aligned}
\mathrm{Cov}(x_j,x_k) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_j x_k \rho(\mathbf{x},\mathbf{p})d^n x d^n p \\
\mathrm{Cov}(x_j,p_k) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_j p_k \rho(\mathbf{x},\mathbf{p})d^n x d^n p \\
\mathrm{Cov}(p_j,p_k) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_j p_k \rho(\mathbf{x},\mathbf{p})d^n x d^n p.
\end{aligned} \tag{46}$$

We now consider a cloud Ω of $K \gg 1$ points $\mathbf{z}_1 = (\mathbf{x}_1, \mathbf{p}_1), , ..., \mathbf{z}_K = (\mathbf{x}_K, \mathbf{p}_K)$ lying in phase space, where each of the points corresponds to a joint position/momentum measure.

In statistical analysis, it is standard procedure to down-weigh outliers, i.e. observations that do not follow the pattern of the majority of the data. One can do this by discarding these outliers or by using a statistical method that is robust to outliers, such as the minimum volume ellipsoid (MVE) method [11], [12]. This method geometrically amounts to finding the minimal volume ellipsoid circumscribing a set of points.

The set $\{\mathbf{z}_1, ..., \mathbf{z}_K\}$ of all retained points in Ω determines a convex polyhedron $\mathcal{S}$ in $\mathbb{R}^{2n}$. The convex hull of $\mathcal{S}$ is defined by the intersection of all convex sets in $\mathbb{R}^{2n}$ which contain $\mathcal{S}$, and is denoted $\tilde{\mathcal{S}}$. In 1948, Fritz John proved that there exists a unique minimal volume ellipsoid $\mathcal{J}$ in $\mathbb{R}^{2n}$ containing $\tilde{\mathcal{S}}$ [13], which is known as the John-Löwner ellipsoid. By means of this ellipsoid one can identify outliers quickly, because the outliers are essentially points on the boundary of the minimum volume ellipsoid.
The John-Löwner ellipsoid is a useful tool in a variety of different application areas, such as convex optimization and computational geometry, and has been studied for over 50 years.

4.1 System with a single degree of freedom

We will first consider the case where the system of particles has only one degree of freedom. Consider a cloud Ψ of $K \gg 1$ points $z_1 = (x_1, p_1), ..., z_K = (x_K, p_K)$ lying in the phase plane.
For one degree of freedom, Ψ is replaced by the John-Löwner ellipse $\mathcal{J}$ containing Ψ. The center of that ellipse is then identified with the mean and the shape of the ellipse determines the covariance. An ellipse is defined by the formula

$$(z - \bar{z})^T M^{-1} (z - \bar{z}) \leq m^2, \tag{47}$$

where $\bar{z}$ is the mean and M is a positive definite matrix.

We now choose an adequate value m_0^2, which determines the shape of the ellipse and thus also the covariance matrix. If the sample of phase space points z_j can reasonably assumed to be normally distributed, then a standard choice would be to determine m_0^2 by means of the chi-square distribution table. For $n = 1$ and a probability value of 0.5, $m_0^2 = \chi^2_{0.5}(2n) = \chi^2_{0.5}(2) \approx 1.39$.

For one degree of freedom, the covariance matrix Σ is defined as

$$\Sigma = \begin{pmatrix} \Delta x^2 & \mathrm{Cov}(x,p) \\ \mathrm{Cov}(p,x) & \Delta p^2 \end{pmatrix}, \tag{48}$$

which is a real symmetric positive-definite matrix. Σ can be visualized by the error ellipse

$$\mathcal{J} : (z - \bar{z})^T \Sigma^{-1} (z - \bar{z}) \leq m_0^2, \tag{49}$$

which is known as the John-Löwner ellipse.

Subsequently, we associate to $\mathcal{J}$ a covariance ellipse $\mathcal{C}$;

$$\mathcal{C} : \frac{1}{2}(z - \bar{z})^T \Sigma^{-1}(z - \bar{z}) \leq 1. \tag{50}$$

One can easily see that the ellipses $\mathcal{C}$ and $\mathcal{J}$ are homothetic, i.e. by multiplying the area of $\mathcal{J}$ with a certain value, we get the area of $\mathcal{C}$;

$$\text{Area}(\mathcal{C}) = \frac{2}{m_0^2}\text{Area}(\mathcal{J}). \tag{51}$$

Let a and b define the length of the semi-major and semi-minor axes of the ellipse respectively. Then the eigenvalues of the covariance matrix Σ are given by a^2 and b^2. Because the determinant of a matrix is given by the product of its eigenvalues, it follows that $\det(\Sigma) = a^2b^2$. Thus,

$$\begin{aligned}\text{Area}(\mathcal{C}) &= 2\pi ab;\\ &= 2\pi(\det\Sigma)^{1/2};\\ &= 2\pi[\Delta x^2 \Delta p^2 - \text{Cov}(x,p)^2]^{1/2},\end{aligned} \tag{52}$$

where the input of the value 2 in the equation of (52) comes from the fact that the value $\frac{1}{2}$ in the formula of the covariance ellipse (50) has to be neutralized.

If we assume that

$$\text{Area}(\mathcal{J}) \geq \frac{1}{4}m_0^2 h, \tag{53}$$

then it follows by equation (51) that

$$\text{Area}(\mathcal{C}) = 2\pi[\Delta x^2 \Delta p^2 - \text{Cov}(x,p)^2]^{1/2} \geq \frac{1}{2}h. \tag{54}$$

and (54) is strictly equivalent to the Schrödinger-Robertson inequality (41), where $\hbar = h/2\pi$ and h denotes a constant (possibly Planck's constant).

4.2 System with multiple degrees of freedom

Now, we will consider the case where the system $\mathcal{S}$ of particles has n degrees of freedom, and thus lies in the phase space $\mathbb{R}^{2n}$. Consider again the cloud Ω of $K \gg 1$ points $\mathbf{z}_1 = (\mathbf{x}_1, \mathbf{p}_1), , ..., \mathbf{z}_K = (\mathbf{x}_K, \mathbf{p}_K)$ lying in phase space. We associate this cloud with a domain of $\mathbb{R}^{2n}$, where we suppose that Ω is not contained by any subspace with dimension less than $2n$.
We assume that the phase space evolution of the system is controlled by Hamilton's equations. Thus, the system consists of $2n$ differential equations, that determine a phase space flow f_t^H consisting of canonical transformations. We will show later on that the inequalities (42) are conserved in time under a linear phase space flow f_t^H.

For many degrees of freedom the John-Löwner ellipse becomes an ellipsoid $\mathcal{J}$ in $\mathbb{R}^{2n}$. To that ellipsoid one can again associate a covariance matrix Σ,

determined by the shape of $\mathcal{J}$, and a covariance ellipsoid $\mathcal{C}$.
Equations (49) and (50) also apply for multiple degrees of freedom, where "area" and "ellipse" are replaced by "volume" and "ellipsoid", respectively.
An adequate value m_0^2 is again chosen to down-weigh the outliers.
When the points $\mathbf{z}_j$ are normally distributed $\mathcal{C}$ will be smaller than $\mathcal{J}$ as soon as $n > 1$, since $m_0^2 = \chi^2_{0.5}(2n)$ goes to infinity with n.

The covariance matrix Σ is defined in the block-matrix form

$$\Sigma = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XP} \\ \Sigma_{PX} & \Sigma_{PP} \end{pmatrix}, \tag{55}$$

where the blocks Σ_{XX}, Σ_{XP}, Σ_{PX}, and Σ_{PP} are $n \times n$ matrices, defined as

$$\begin{aligned} \Sigma_{XX} &= (\mathrm{Cov}(x_j, x_k))_{j,k}, \\ \Sigma_{PP} &= (\mathrm{Cov}(p_j, p_k))_{j,k}, \\ \Sigma_{XP} &= (\mathrm{Cov}(x_j, p_k))_{j,k}, \\ \Sigma_{PX} &= (\mathrm{Cov}(p_j, x_k))_{j,k}. \end{aligned} \tag{56}$$

Σ is symmetric, so $\Sigma_{XX} = \Sigma_{XX}^T$, $\Sigma_{PP} = \Sigma_{PP}^T$, and $\Sigma_{XP} = \Sigma_{PX}^T$.

It is also possible to denote the variances as

$$\begin{aligned} \mathrm{Cov}(x_j, x_j) &= (\Delta x_j)^2 \\ \mathrm{Cov}(p_j, p_j) &= (\Delta p_j)^2, \end{aligned} \tag{57}$$

so for one degree of freedom the covariance matrix becomes the matrix defined by equation (48).

We will now look for the condition which should be imposed on $\mathcal{C}$ in order to derive the Schrödinger-Robertson inequality for multiple degrees of freedom.
In analogy to the system with a single degree of freedom, where the condition $\mathrm{Area}(\mathcal{C}) \geq \frac{1}{2}h$ must hold if one wants to derive the Schrödinger-Robertson inequality, one would suspect that the condition for n degrees of freedom would be

$$\mathrm{Volume}(\mathcal{C}) \geq (\frac{1}{2}h)^n, \tag{58}$$

but this is not generally true. The Schrödinger-Robertson inequalities are not expressed in terms of volume, but again in terms of area.
In order to derive the inequalities (42), the symplectic capacity of the covariance ellipsoid $\mathcal{C}$ should be at least $\frac{1}{2}h$, i.e.

$$c(\mathcal{C}) \geq \frac{1}{2}h. \tag{59}$$

We will show later on why this must be the case. A key to the argument is the following property of the covariance matrix;

$$\Sigma + \frac{i\hbar}{2}J \geq 0, \tag{60}$$

where "≥ 0" is synonymously used for "is semi-definite positive".

Theorem 5. *Let Σ be a real symmetric $2n \times 2n$ matrix.*
Now assume that $\Sigma + \frac{i\hbar}{2}J \geq 0$. Then:

I) *The matrix Σ must be positive definite* [14];

II) The inequalities

$$(\Delta x_j)^2(\Delta p_j)^2 \geq \mathrm{Cov}(x_j, p_j)^2 + \frac{\hbar^2}{4} \tag{61}$$

hold for $j = 1, ..., n$.

Proof. **I)**
The covariance matrix Σ is a real symmetric matrix and $J^T = -J$, so the matrix $\Sigma + \frac{i\hbar}{2}J$ is obviously Hermitian;

$$(\Sigma + \frac{i\hbar}{2}J)^\dagger = \Sigma - \frac{i\hbar}{2}J^T = \Sigma + \frac{i\hbar}{2}J. \tag{62}$$

It follows in particular that all the eigenvalues of $\Sigma + \frac{i\hbar}{2}J$ are real.
We will show that Σ must be positive definite if $\Sigma + \frac{i\hbar}{2}J \geq 0$.
First, we will prove that the covariance matrix Σ is non-negative. Suppose, by contradiction, that Σ has a negative eigenvalue $\lambda < 0$. Σ is real and symmetric, so there exists a real eigenvector z_λ corresponding to λ.
Because z_λ is real, it follows that $z_\lambda^T J z_\lambda = \omega(z_\lambda, z_\lambda) = 0$.

Hence, we get

$$\begin{aligned} z_\lambda^T(\Sigma + \frac{i\hbar}{2}J)z_\lambda &= z_\lambda^T \Sigma z_\lambda + \frac{i\hbar}{2} z_\lambda^T J z_\lambda \\ &= \lambda \|z_\lambda\|^2 < 0. \end{aligned} \tag{63}$$

This is a contradiction to the assumption that $\Sigma + \frac{i\hbar}{2}J$ is non-negative, so Σ has no negative eigenvalues and is therefore also a non-negative matrix.

Subsequently, we show that zero cannot be an eigenvalue of Σ.
Suppose, by contradiction, that zero is an eigenvalue, and let z_0 be the real eigenvector corresponding to zero. Since z_0 is real, it follows again that $z_0^T J z_0 = \omega(z_0, z_0) = 0$. Note also that $\Sigma z_0 = 0$ and $z_0^T \Sigma = 0$.
Now define the vector $z_\epsilon \equiv (I + i\epsilon J)z_0$, where $\epsilon \in \mathbb{R}$. This allows us to perform the following calculation, making use of the fact that $J^2 = -I$ and $J^T = -J$:

$$
\begin{aligned}
z_\epsilon^T(\Sigma + i\frac{\hbar}{2}J)z_\epsilon &= z_0^T(I+i\epsilon J)^\dagger(\Sigma + i\frac{\hbar}{2}J)(I+i\epsilon J)z_0 \\
&= z_0^T(I+i\epsilon J)(\Sigma + i\frac{\hbar}{2}J)(I+i\epsilon J)z_0 \\
&= z_0^T(\Sigma + i\frac{\hbar}{2}J + i\epsilon J\Sigma - \epsilon\frac{\hbar}{2}J^2)(I+i\epsilon J)z_0 \\
&= z_0^T(\Sigma + i\frac{\hbar}{2}J + i\epsilon J\Sigma + \epsilon\frac{\hbar}{2}I)(I+i\epsilon J)z_0 \\
&= z_0^T(\Sigma + i\epsilon\Sigma J + i\frac{\hbar}{2}J - \epsilon\frac{\hbar}{2}J^2 + i\epsilon J\Sigma - \epsilon^2 J\Sigma J + \epsilon\frac{\hbar}{2}I + i\epsilon^2\frac{\hbar}{2}J)z_0 \\
&= z_0^T(\Sigma + i\epsilon\Sigma J + i\frac{\hbar}{2}J + \epsilon\frac{\hbar}{2}I + i\epsilon J\Sigma - \epsilon^2 J\Sigma J + \epsilon\frac{\hbar}{2}I + i\epsilon^2\frac{\hbar}{2}J)z_0 \\
&= z_0^T(\Sigma + i\epsilon\Sigma J + i\epsilon J\Sigma - \epsilon^2 J\Sigma J)z_0 + i\frac{\hbar}{2}z_0^T J(I + 2i\epsilon J + \epsilon^2 I)z_0 \\
&= -\epsilon^2 z_0^T J\Sigma J z_0 - \hbar\epsilon z_0^T J^2 z_0 \\
&= \epsilon^2 z_0^T J^T\Sigma J z_0 + \hbar\epsilon z_0^T I z_0 \\
&= \epsilon^2(Jz_0)^T\Sigma(Jz_0) + \epsilon\hbar\|z_0\|^2. \qquad (64)
\end{aligned}
$$

If we choose $\epsilon < 0$ and let ϵ be small enough, we get

$$
z_\epsilon^T(\Sigma + i\frac{\hbar}{2}J)z_\epsilon < 0, \qquad (65)
$$

which again contradicts the assumption that $\Sigma + i\frac{\hbar}{2}J$ is semi-positive definite. Since Σ cannot have negative or zero eigenvalues, it must be positive definite.

□

Proof. **II)**
We can express the non-negativity of the Hermitian matrix $\Sigma + \frac{i\hbar}{2}J$ in terms of the submatrices

$$
\Sigma_{ij} = \begin{pmatrix} (\Delta x_j)^2 & \Delta(x_j, p_j) + \frac{i\hbar}{2} \\ \Delta(p_j, x_j) - \frac{i\hbar}{2} & (\Delta p_j)^2 \end{pmatrix}, \qquad (66)
$$

which are non-negative and Hermitian provided that $\Sigma + \frac{i\hbar}{2}J$ is as well.

The trace of the matrix Σ_{ij} is non-negative, so $\Sigma_{ij} \geq 0$ if and only if

$$
\det(\Sigma_{ij}) = (\Delta x_j)^2(\Delta p_j)^2 - \Delta(x_j, p_j)^2 - \frac{\hbar^2}{4} \geq 0. \qquad (67)
$$

If we replace $\Delta(x_j, p_j)$ by $\mathrm{Cov}(x_j, p_j)$, then above equation is equivalent to the Schrödinger-Robertson inequalities (42).

□

It is important to note that, for multiple degrees of freedom, the condition $\Lambda + \frac{i\hbar}{2}J \geq 0$ is not equivalent to the uncertainty inequalities (42); one cannot generally derive the property $\Lambda + \frac{i\hbar}{2}J \geq 0$ if we assume that the Schrödinger-Robertson inequalities hold for $j = 1, ..., n$.

Theorem 6. *The condition* $\Sigma + \frac{i\hbar}{2}J \geq 0$ *is equivalent to* $c(\mathcal{C}) \geq \frac{1}{2}h$.

Proof. Consider a cloud of points Ω in phase space $\mathbb{R}^{2n}$. Assume from this point on that the convex hull $\tilde{\mathcal{S}}$ of the set $\mathcal{S} = \{\mathbf{z}_1, ..., \mathbf{z}_K\}$ of reliable points satisfies

$$c_0(\tilde{\mathcal{S}}) \geq \frac{1}{4}m_0^2 h \tag{68}$$

for some symplectic capacity c_0.
The convex hull $\tilde{\mathcal{S}}$ is contained inside the John-Löwner ellipsoid $\mathcal{J}$, i.e. $\tilde{\mathcal{S}} \subseteq \mathcal{J}$. Thus by property (26), $\mathcal{J}$ satisfies

$$c(\mathcal{J}) \geq \frac{1}{4}m_0^2 h \tag{69}$$

for every symplectic capacity c, since all symplectic capacities agree for phase space ellipsoids, by Lemma 4.
$\mathcal{J}$ is homothetic to the covariance ellipsoid $\mathcal{C}$ by equation (51), so this is equivalent to

$$c(\mathcal{C}) \geq \frac{1}{2}h. \tag{70}$$

By using Williamson's diagonalization theorem, we will now show that the conditon $\Sigma + \frac{i\hbar}{2}J \geq 0$ is equivalent to $c(\mathcal{C}) \geq \frac{1}{2}h$.

The covariance matrix Σ can be diagonalized by means of a symplectic matrix S:

$$\begin{aligned} D &= S^T \Sigma S \\ &= \begin{pmatrix} \Gamma & 0 \\ 0 & \Gamma \end{pmatrix}, \end{aligned} \tag{71}$$

where $\Gamma = \text{diag}(R_1^2, ..., R_n^2)$. Thus, the eigenvalues of $J\Sigma$ are given by $\pm i\mu_j = \pm iR_j^2$ (where $\mu_j = R_j^2 > 0$). R_j denotes the length of the jth semi-principal axis of the covariance ellipsoid $\mathcal{C}$, for $j = 1, ..., n$.

We can now state that

$$\begin{aligned} c(\mathcal{C}) &= 2\pi R_{\text{min}}^2 \\ &= 2\pi \mu_{\text{min}}, \end{aligned} \tag{72}$$

where μ_{min} is the modulus of the smallest eigenvalue of the matrix $J\Sigma$ and the input of the value 2 in above equation of (52) is again necessary to neutralize the value $\frac{1}{2}$ in the formula of the covariance ellipse (50).

So if $\mu_{\min} \geq \frac{1}{2}\hbar$, then it follows that

$$c(\mathcal{C}) = 2\pi\mu_{\min} \geq 2\pi(\frac{1}{2}\hbar) = \frac{1}{2}h. \tag{73}$$

Since $\Sigma^{-1/2}$ is positive definite, the assumption $\Sigma + \frac{i\hbar}{2}J \geq 0$ is equivalent to the assumption that

$$\Sigma^{-1/2}(\Sigma + \frac{i\hbar}{2}J)\Sigma^{-1/2} = I + \frac{i\hbar}{2}\Sigma^{-1/2}J\Sigma^{-1/2} \geq 0. \tag{74}$$

It is easy to show that $J\Sigma^{-1}$ and $\Sigma^{-1/2}J\Sigma^{-1/2}$ have the same set of eigenvalues by using the characteristic equation;

$$\begin{aligned} 0 &= \det(J\Sigma^{-1} - \lambda I) \\ &= \det(\Sigma^{-1/2})\det(J\Sigma - \lambda I)\det(\Sigma^{1/2}) \\ &= \det(\Sigma^{-1/2}(J\Sigma^{-1} - \lambda I)\Sigma^{1/2}) \\ &= \det(\Sigma^{-1/2}J\Sigma^{-1/2} - \lambda\Sigma^{-1/2}\Sigma^{1/2}) \\ &= \det(\Sigma^{-1/2}J\Sigma^{-1/2} - \lambda I). \end{aligned} \tag{75}$$

The eigenvalues of $J\Sigma^{-1}$ are given by $\pm i\lambda_j = \pm i/R_j^2$, for $j = 1, ..., n$.

$I + \frac{i\hbar}{2}\Sigma^{-1/2}J\Sigma^{-1/2} \geq 0$ is equivalent to $I + \frac{i\hbar}{2}D^{-1/2}JD^{-1/2} \geq 0$, because we can convert the conditon $I + \frac{i\hbar}{2}\Sigma^{-1/2}J\Sigma^{-1/2} \geq 0$ into the condition $I + \frac{i\hbar}{2}D^{-1/2}JD^{-1/2} \geq 0$ through conjugation by a symplectic matrix.

The form of the matrix $D^{-1/2}JD^{-1/2}$ is given by

$$\begin{aligned} \frac{1}{2}D^{-1/2}JD^{-1/2} &= \begin{pmatrix} \Gamma^{-1/2} & 0 \\ 0 & \Gamma^{-1/2} \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} \Gamma^{-1/2} & 0 \\ 0 & \Gamma^{-1/2} \end{pmatrix} \\ &= \begin{pmatrix} 0 & \Gamma^{-1/2} \\ -\Gamma^{-1/2} & 0 \end{pmatrix} \begin{pmatrix} \Gamma^{-1/2} & 0 \\ 0 & \Gamma^{-1/2} \end{pmatrix} \\ &= \begin{pmatrix} 0 & \Gamma^{-1} \\ -\Gamma^{-1} & 0 \end{pmatrix}. \end{aligned} \tag{76}$$

And hence $I + \frac{i\hbar}{2}\Sigma^{-1/2}J\Sigma^{-1/2} \geq 0$ is equivalent to

$$\begin{pmatrix} I & \frac{i\hbar}{2}\Gamma^{-1} \\ -\frac{i\hbar}{2}\Gamma^{-1} & I \end{pmatrix} \geq 0. \tag{77}$$

The characteristic polynomial $\mathcal{P}(t)$ of $I + \frac{i\hbar}{2}\Sigma^{-1/2}J\Sigma^{-1/2}$ is given by $\mathcal{P}(t) = \mathcal{P}_1(t) \dots \mathcal{P}_n(t)$, where

$$\mathcal{P}_j(t) = t^2 - 2t + 1 - \frac{\hbar^2}{4\mu_j^2}, \text{ for } j = 1, ..., n. \tag{78}$$

So the eigenvalues t_j of $I + \frac{i\hbar}{2}\Sigma^{-1/2}J\Sigma^{-1/2}$ are given by

$$\begin{aligned} t_j &= \frac{2 \pm \sqrt{4 - 4(1 - \hbar^2/4\mu_j^2)}}{2} \\ &= \frac{2 \pm \sqrt{\hbar^2/\mu_j^2}}{2} \\ &= 1 \pm \frac{\hbar}{2\mu_j} \end{aligned} \tag{79}$$

The eigenvalues t_j are non-negative if and only if $1 \pm \frac{\hbar}{2\mu_j} \geq 0$, or equivalently if

$$|\mu_j| \geq \frac{\hbar}{2}, \text{ for } j = 1, ..., n. \tag{80}$$

Thus, $\mu_{\text{min}} \geq \frac{\hbar}{2}$, and therefore $c(\mathcal{C}) \geq \frac{1}{2}h$.

□

Hence, the assumption $\Sigma + \frac{i\hbar}{2}J \geq 0$ is equivalent to assuming that $c(\mathcal{C}) \geq \frac{1}{2}h$ and thus the Schrödinger-Robertson inequalities (42) can be derived, by Theorem 5.

4.3 Time-variant Hamiltonian Systems

The inequalities (42) are conserved in time under a linear Hamiltonian evolution. In other words, if the inequalities

$$(\Delta x_j)^2(\Delta p_j)^2 \geq \text{Cov}(x_j, p_j)^2 + \frac{1}{4}\hbar^2 \tag{81}$$

hold at time $t = 0$, then the inequalities

$$(\Delta x_{j,t})^2(\Delta p_{j,t})^2 \geq \text{Cov}(x_{j,t}, p_{j,t})^2 + \frac{1}{4}\hbar^2, \tag{82}$$

will hold for all times t, past and future, for $j = 1, ..., n$.

Consider a linear Hamiltonian flow f_t^H on the previously defined phase space cloud Ω. It was stated in [15] that the linearized flow of a Hamiltonian is still symplectic, so it consists of linear canonical transformations.
We assume that $c(\Omega) \geq \frac{1}{2}h$. As time passes, f_t^H will deform Ω into a new cloud of points $\Omega_t = f_t^H(\Omega)$, which has the same symplectic capacity, since symplectic capacities are invariant under canonical transformations.
Hence,

$$c(\Omega_t) \geq \frac{1}{2}h. \tag{83}$$

The convex hull of Ω_t is denoted by $\tilde{\Omega}_t$, and is contained by the John-Löwner ellipsoid $\mathcal{J}_{\Omega_t}$. Because $c(\tilde{\Omega}_t) \geq \frac{1}{2}h$ and $\tilde{\Omega}_t \subseteq \mathcal{J}_{\Omega_t}$, it follows by (26) that

$$c(\mathcal{J}_{\Omega_t}) \geq \frac{1}{2}h. \tag{84}$$

This condition turns out to be equivalent to the inequalities (82), where $\Delta x_{j,t}$, etc. are defined in terms of the covariance matrix for a time-variant system

$$\Sigma_t = \begin{pmatrix} \Sigma_{XX,t} & \Sigma_{XP,t} \\ \Sigma_{PX,t} & \Sigma_{PP,t} \end{pmatrix}. \tag{85}$$

When we try to generalize this result to arbitrary Hamiltonian flows, we experience some difficulties, mainly caused by the fact that a generic Hamiltonian flow does not preserve the convexity of a structure [16]. So one cannot in general associate a John-Löwner ellipsoid to the deformed cloud $f_t(\Omega)$. Several advances to the current state of art are necessary, before we can obtain practical results for the nonlinear case.

5 Discussion and Concluding Remarks

In this thesis, we have shown that for large statistic ensembles, the uncertainty principle of quantum mechanics is already in some form present in classical mechanics. In hindsight, this fact is not very surprising, since Schrödinger's formulation of quantum mechanics was from the very beginning modelled on classical Hamiltonian mechanics: the operator $\mathcal{H}$ appearing in the Schrödinger equation is obtained by "quantization" of the Hamiltonian function.

In 1939, Hermann Weyl stated that "*the angel of topology and the devil of abstract algebra fight for the soul of each individual mathematical domain.*" In the field of quantum mechanics, algebra dominated the scene since the very beginning. However, we are now witnessing a slow but steady emergence of geometric ideas. The fresh breath of air, provided by symplectic topology, should inspire physicists and mathematicians to look at the field of mechanics from a new perspective, and hence gain new insights about the grey area between quantum and classical mechanics.

Acknowledgements

A special thanks to prof. dr. H. Waalkens and prof. dr. D. Boer, who provided me with assistance during the summer and were willing to help me out on short notice.

6 References

[1] I. Stewart. The Symplectic Camel. In *Nature*, Vol. 329, Issue 6134 (1987), pp. 17-18.

[2] M. de Gosson. The Symplectic Camel and the Uncertainty principle: The Tip of an Iceberg? In *Foundation of Physics*, Vol. 39, Issue 2 (February 2009), pp. 194214.

[3] V. I. Arnold. Mathematical Methods of Classical Mechanics. 2nd edition, 1989.

[4] D. McDuff. What is Symplectic Geometry? March, 2009.

[5] A. C. da Silva. Lectures on Symplectic Geometry. January, 2006.

[6] M. Gromov. Pseudo holomorphic curves in symplectic manifolds. In *Inventiones Mathematicae*, Vol. 82 (1985), pp. 307-347.

[7] M. S. El Naschie. New elementary particles as a possible product of a disintegrating symplictic vacuum. In *Chaos, Solitons & Fractals*, Vol. 20, Issue 4 (May 2004), pp. 905-913.

[8] I. Ekeland, H. Hofer. Symplectic topology and Hamiltonian dynamics. In *Mathematische Zeitschrift*, Vol. 200, Issue 3 (1989), pp. 355-378.

[9] K. Cieliebak, H. Hofer, J. Latschev, F. Schlenk. Quantitative symplectic geometry. June, 2005.

[10] J. Williamson. On an algebraic problem, concerning the normal forms of linear dynamical systems. Amer. J. Math. Vol. 58 (1936), pp. 141-163.

[11] M. Henk. Löwner-John Ellipsoids. 2010.

[12] P. Sun, R. M. Freund. Computation of Minimum-Volume Covering Ellipsoids. In *Operations Research*, Vol. 52, Issue 5 (September, 2003), pp. 690-706.

[13] F. John. Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays Presented to R. Courant on his 60th Birthday*, Vol. 52, Issue 5 (January, 1948), pp. 187204.

[14] F. J. Narcowich. Geometry and uncertainty. In *Journal of Mathematical Physics*, Vol. 31, Issue 2 (1990).

[15] F. Y. Hsiao, D. J. Scheeres. Fundamental Constraints on Uncertainty Evolution in Hamiltonian Systems. In *IEEE Transactions on Automatic Control*, Vol. 52, Issue 4 (April 2007), pp. 686-691.

[16] M. de Gosson. How classical is the quantum universe? August, 2008.