Uncertainty Relations for Quantum Particles

Spyridon Kechrimparis

PhD

University of York

Mathematics

December, 2015

Abstract

The focus of the present investigation is *uncertainty relations* for quantum particles, which quantify the fundamental limitations on some of their properties due to their incompatibility. In the first, longer part, we are concerned with *preparational* uncertainty relations, while in the second we touch upon *measurement* inequalities through a generalisation of a model of joint measurement of position and momentum.

Specifically, starting from a triple of canonical operators, we prove product and sum inequalities for their variances with bounds larger than those following from combining the pairwise ones. We extend these results to N observables for a quantum particle and prove uncertainty relations for the sums and products of their variances in terms of the commutators.

Furthermore, we present a general theory of preparational uncertainty relations for a quantum particle in one dimension and derive conditions for a smooth function of the second moments to assume a lower bound. The *Robertson-Schrödinger* inequality is found to be of special significance and we geometrically study the space of second moments. We prove new uncertainty relations for various functions of the variances and covariance of position and momentum of a quantum particle. Some of our findings are shown to extend to more than one spatial degree of freedom and we derive various types of inequalities.

Finally, we propose a generalisation of the Arthurs-Kelly model of joint measurement of position and momentum to incorporate the case of more than two observables and derive *joint-measurement* inequalities for the statistics of the probes. For the case of three canonical observables and suitable definitions of error and disturbance we obtain a number of *error-error-error* and *error-error-disturbance* inequalities and show that the lower bound is identical to the preparational one.

Contents

Acknowledgments

There is a long list of people that have contributed one way or the other to the realisation of this thesis and to whom I would like to express my gratitude.

First of all, I wish to thank my parents, without their unconstrained support whenever I was in need, this thesis would not be possible; I will be forever in debt for your contribution. In extension, my brother and all my family members that have been supportive and understanding throughout this time.

I also wish to express my gratitude to my supervisor, Stefan Weigert, for his guidance throughout these years, his patience, support, the numerous discussions and insightful suggestions. Moreover, I would like to thank him for granting me the freedom to explore my own ideas, while at the same time making sure I was not getting off track. It was a great experience working with him and his contribution in shaping my current thinking as a researcher is invaluable.

In addition, I acknowledge financial support from the State Scholarship Foundation (I.K.Y.) in Greece, for awarding me a scholarship covering the first three and a half years of my studies, and a grant from the WW Smith Fund in York, that covered part of my expenses of the last six months.

Last but not least, I wish to thank all my friends and fellow travellers in space-time, in York and around the globe, through the interaction with whom the last few years I have been a complete and happy person. Cheers to James, Fabian, Dan, Dave H., Leon, Dave B., Dave C., Umberto, Tom, Ben, Alejandro, Jesus, Emilio, Oliver and Eli, Pablo and Alessi, Ella, Kimi, Vasilis, Christopher, and of course Mr. John Smith. Special thanks to Dave Hunt for useful comments on parts of this thesis. Thank you all.

Declaration

I declare that the work presented in this thesis, except where otherwise stated, is based on my own research and has not been submitted previously for any degree at this or other university. Other sources are acknowledged by explicit references.

Chapters 3 and 5 of this thesis have been published in:

S. Kechrimparis and S. Weigert, *Phys. Rev. A* **90**, 062118 (2014).
S. Kechrimparis and S. Weigert, *arXiv:1509.02146* (2015) (under review).

Part of Chapter 6 is also under review while Chapters 4 and 7 contain original research that has not been published.

Chapter 1

Introduction

Quantum theory is one of the most fundamental and successful models of physical reality known to man. It has far reaching applications in technology and everyday experience; the laser and magnetic resonance imaging (nMRI) are only some examples of its influence on modern life. At a time when it was widely accepted that all major phenomena observed in nature were already described by the contemporary physical theories, Planck's solution to the black body radiation problem, one of the few remaining obstacles to the perceived complete description of nature, and Einstein's explanation of the photoelectric effect planted the seeds for the development of a new theory of microscopic phenomena. Its first coherent formulations were developed in the first half of the twentieth century and in the subsequent decades it was put on more solid mathematical grounds. Bohr, Dirac, Heisenberg, Schrödinger, Einstein, Wigner and von Neumann are some of the physicists and mathematicians who substantially contributed to the understanding of the new theory.

However, regardless of its exceptional success as a mathematical framework to make testable predictions concerning our experience of interaction with the world, it was obvious, from an early stage on, that there were a number of obstacles challenging some of the standard underlying assumptions of classical physical theories; few would question the implicitly assumed interpretation of what is happening when an apple falls on a physicist's head, when such a process is described by Newton's law of gravitation. However, quantum theory brought to the limelight a type of philosophical questions that could no longer be easily suppressed. Even today, a little less than a century from the theory's first formulations, significant foundational problems exist within quantum mechanics. This claim can be demonstrated, for example, by the disagreement of experts in the field, on the meaning of different aspects of the theory [61, 51] and the wide variety of different interpretations [24]. The famous *measurement problem*, for example, persistent in the *Copenhagen* interpretation of quantum mechanics is not a difficulty for a *QBist*, who accepts that the quantum state only relates to the agent's beliefs about the system [33, 72]. We will not dwell on interpretational questions of quantum mechanics however in this thesis since the scientific value in such an exercise is debatable [32], and we will leave such questions open to the discretion of the reader.

One of the most peculiar features of quantum mechanics is the celebrated *uncertainty principle*, which is one example of the groundbreaking implications of the theory that substantially separates it from the viewpoint of a *classical* world described by Newton's theory. Since its first appearance, it has been the focus of research, discussion and debates; even today, almost ninety years after it was first introduced, there are strong disagreements regarding its meaning and physical content.

Heisenberg in 1927 [35] was the first to introduce the idea of quantum indeterminacy, using qualitative arguments about the impossibility to measure simultaneously position and momentum with arbitrary accuracy. In Heisenberg's description of the uncertainty principle through a thought experiment, the observation of the position of a particle with a microscope would disturb the particle's momentum due to the Compton effect by an amount of the order of Planck's constant.

Soon after, Kennard [41] in 1927 and Weyl [81] in 1928, provided mathematical proofs of what is commonly referred to as *Heisenberg's uncertainty relation*, which as it is now understood, is different in spirit from the notion of uncertainty introduced by Heisenberg. Nevertheless, the two have been taken to be the same statement for years. In 1929, Robertson [56] provided an inequality for the variances of two self-adjoint operators, resulting from their non-commutativity, generalising the result of Kennard. Finally, Schrödinger in 1930 [62] further strengthened the uncertainty relation by including a "covariance" term. In the case of position and momentum, the resulting *Robertson-Schrödinger inequality* plays a special role for a quantum particle in one dimension.

It is worth mentioning that the inequality proved by Kennard and Weyl is conceptually different from the bound described by Heisenberg. His original microscope argument presented a theoretical lower bound concerning the accuracy of measuring the position of a particle and the disturbance that such a measurement would cause to its momentum; however, it is important to note that no definitions of error or disturbance were explicitly given. Moreover, no inequality appears in his paper and the only relevant relation presented was "$p_1 q_1 \sim h$", where q_1 is "*the precision with which the value q is known (q_1 is, say, the mean error of q)*" whereas p_1 is "*the precision with which the value p is determinable*" [35] ([82], p.64) . For many years following Heisenberg's seminal paper, there was confusion on the subject and the inequality proved by Kennard and Weyl was thought of as an adequate description of the limitations associated with the *error* and *disturbance* of a measurement of conjugate variables. Only relatively recently was it understood that different types of inequalities need to be formulated in order to capture the various aspects of quantum uncertainty.

The standard textbook inequality proved by Kennard, usually referred to as Heisenberg's uncertainty relation, is an expression of the impossibility to prepare a quantum system in a state where both its position and momentum are arbitrarily well localized: there is a universal lower bound equal to half of the reduced Planck constant, $\hbar/2$, for the product of the spreads of the position and momentum distributions, obeyed by all possible quantum states, where the measure of spread used is the standard deviation. Thus, the textbook uncertainty relation is a statement about *preparations* of a quantum system rather than measurements. Alternatively, it can be thought of as a bound concerning the spreads in outcomes of ideal separate measurements of position and momentum on an ensemble of identically prepared particles. In this thesis, we will be mostly dealing with *preparational inequalities* for one or more quantum particles, and only in the penultimate chapter we will discuss *measurement* uncertainty, through a proposed generalisation of a model of joint measurement of position and momentum, introduced by Arthurs and Kelly [6] in 1965.

Recently, there has been considerable interest in revisiting Heisenberg's original ideas and putting the uncertainty principle on solid mathematical grounds; these new inequalities, usually referred to as *error-disturbance* uncertainty relations, deal with the effect that a measurement of a particle's position has on momentum and vice-versa. Ozawa [52, 53] in 2003 refuted and reformulated an inequality which he attributed to Heisenberg, while experiments [58, 31] that were performed recently provided evidence of his theoretical findings. At the same time, however, a proof of Heisenberg's error-disturbance relation was given by Busch. et. al. [19], in a form resembling the inequality disproved by Ozawa, although different in mathematical content. The apparent contradiction between these works is due to the use of different measures quantifying the error and disturbance relating to the measurement. There is an ongoing debate on the merits and drawbacks of the various measures and of the different resulting formulations of Heisenberg's error-disturbance uncertainty relation. Although the ones employed by Ozawa have been claimed to be advantageous due to the fact that they are *state-dependent*, it seems that they suffer from a number of shortcomings and have been, for this reason, strongly criticised [29, 43, 16, 18]. On the other hand, *state-independent* measures have been argued to capture the intuitive ideas presented in Heisenberg's paper with more success [16]. Finally, Appleby [4] has formulated inequalities of both types.

Joint-measurement inequalities are statements concerning the statistics of a *simultaneous* measurement of the position and momentum of a particle. A now famous model describing this process was introduced by Arthurs and Kelly [6] and subsequently studied by Busch [15], Appleby [3, 5] and others [70, 22, 23, 55, 36]. In the original model, position and momentum of a quantum particle are coupled to the momenta of two quantum systems, *the probes*, constituting the measuring apparatus, through an impulsive interaction Hamiltonian. It is a consequence of the process that the standard deviations of the pointer observables conjugate to the momenta of the probes, obey a Heisenberg-type inequality with a bound twice as large as the usual value of $\hbar/2$. The common interpretation of this result is that the attempt for a joint measurement of the non-commuting observables of position and momentum introduces extra noise, attributed to the fact that the measuring device is also a quantum mechanical system. Finally, within this model and its generalisations, it is possible to derive various types of error-disturbance inequalities [3, 4].

This thesis is conceptually split into two parts: the first longer part is concerned with preparational uncertainty for continuous variables, while the second focuses on joint-measurement and error-disturbance inequalities through a proposed generalisation of the Arthurs-Kelly model. The first part arises as an answer to a number of questions: why are mathematical expressions of the intrinsic indeterminacy of a quantum system limited to the product and sum of the variances, and not extended to other functions of them? Moreover, potentially stemming from Bohr's complementarity principle, usually stated as the impossibility to access both properties of an incompatible pair of a quantum system, almost all uncertainty statements are concerned only with conjugate pairs as the one of position and momentum. It is a fundamental question to ask what are the implications, if any, in considering more than two incompatible observables and whether any resulting statements are only a consequence of the incompatibility of the pairs. The second part of this thesis, building on the question of preparational uncertainty beyond a pair of observables, examines what are the fundamental implications in attempting to jointly measure more than two observables and the equivalent of error and disturbance inequalities in that setting.

In Chapter 3, we derive lower bounds for the product, sum and other functions symmetric under the exchange of any two variables, of the variances of a triple of canonical operators. Such a triple, unique up to unitary transformations, arises from considerations of existence of mutually unbiased bases for continuous variables and appears in connection with the work of Weigert et. al. [75]. Already in 1934, Robertson [57] derived an inequality for more than two observables, which however becomes trivial for an odd number of them. In [65] an inequality is mentioned for the variances of three operators but cannot be used to infer the minimum bound for a canonical triple. The inequality we derive demonstrates that the incompatibility of three operators, as expressed by uncertainty relations, is not a straightforward consequence of the incompatibility of each pair alone.

In Chapter 4 we prove inequalities for the sums and products of the variances of N observables for a particle with one degree of freedom forming regular polygons. We show that the obtained bounds can be expressed in geometric terms, using the *area* and *circumradius* associated with the regular polygons. The results are extended to observables which are arbitrary linear combinations of position and momentum, and we express the lower bounds in terms of the commutators. Then, N observables for a system of M degrees of freedom are studied; under the assumption of non-existence of correlations between the different degrees of freedom, we find a lower bound that depends on commutators of suitably defined local operators. Finally, we discuss how some of these inequalities can be used as entanglement detection criteria.

In Chapter 5 we respond to the question of whether a smooth function of the three second moments, the two variances and the covariance, is bounded from below. After Kennard's proof of Heisenberg's uncertainty relation, Robertson's extension to arbitrary operators and Schrödinger's addition of the covariance term, few new inequalities appeared for other functions of the second moments of position and momentum, as for example the *sum* inequality for the variances of position and momentum, providing a lower bound of $\hbar$. We will show that given specific conditions, a lower bound exists and the minimum is contained in the solutions of a set of coupled equations for the second moments. Through the theory we will develop, we are able to specify the set of extremal states for the function under consideration, if any exist. Moreover, we find that among all possible *uncertainty functionals* [78], the one associated with the inequality of Robertson-Schrödinger is singled out as one of special importance. We show that its set of extremal states is *universal* in that all potential extrema of an arbitrary functional have to be a subset of it; the states in this universal set are found to be the *squeezed number states* of a harmonic oscillator of unit mass and frequency. We subsequently define a *space of second moments* and show that the universal set is described by the surfaces of a countably infinite number of nested hyperboloids. Finally, we derive a number of new preparational inequalities for a particle with one degree of freedom.

Chapter 6 extends the framework developed in Chapter 5 to the case of more than one spatial degree of freedom. We show that the extremisation of a function of the second moments leads to an equation quadratic in position and momentum operators, which can be diagonalised whenever the matrix of the first partial derivatives of the function is positive definite. We also derive a set of consistency conditions for such a function to be bounded which lead to the following characterisation for extrema: a state is an extremum of a function of the second moments if there exists a symplectic matrix that diagonalises the resulting covariance matrix, while the transpose of its inverse diagonalises the matrix of the partial derivatives of the function.

Chapter 7 is devoted to generalisations of the Arthurs-Kelly process to accommodate the joint measurement of more than two observables. A number of inequalities describing the statistics of the pointer observables after the interaction are obtained. For the generalised version of the model, we find that correlations between the probes allow for bounds lower than those of uncorrelated ones, a feature that is not present in the joint measurement of two conjugate observables. For specific measures of error and disturbance [3, 4] based on noise operators, we prove a number of *error-error-error* and *error-error-disturbance* uncertainty relations concerning the joint measurement of a canonical triple of operators.

The thesis concludes in Chapter 8, with a summary and discussion of the main results.

Chapter 2

Mathematical Preliminaries

For the investigations presented in this thesis, the following framework of quantum mechanics is sufficient. With a quantum system we associate a complex Hilbert space $\mathscr{H}$. The physical *states* of the system correspond to unit vectors in $\mathscr{H}$; such (pure) states differing by a phase factor describe the same physical state. Strictly speaking, physical states are thus identified with equivalence classes of vectors, called *rays*. *Observables* are represented by self-adjoint operators acting on $\mathscr{H}$, a subset of the class of linear transformations of the Hilbert space. Unitary transformations are also important, the set of isometries of the Hilbert space. An example of a unitary transformation is given by $\hat{U} = e^{i\hat{H}t}$, representing the time evolution operator of a quantum system, where $\hat{H}$ is the Hamiltonian operator. In this thesis, we work exclusively with infinite dimensional Hilbert spaces. There are a number of intricacies arising in the case of an infinite dimensional Hilbert space as opposed to a finite dimensional one, but it is beyond the scope of this chapter and in extension of this thesis, to develop the underlying theory in full rigour.

We will often use *Dirac's bra-ket* notation, where the *kets*, denoted by $|.\rangle$, represent the vectors in the Hilbert space, $\mathscr{H}$, while the *bras*, denoted by $\langle .|$, are living in the dual space $\mathscr{H}^*$, the space of linear functions from $\mathscr{H}$ to the complex numbers, $\mathbb{C}$; by the *Riesz representation theorem* the space of linear functionals is identical to the one of the maps $\langle\varphi| : |\psi\rangle \mapsto \langle\varphi|\psi\rangle$, where $\langle .|.\rangle : \mathscr{H} \times \mathscr{H} \to \mathbb{C}$ denotes the inner product of $\mathscr{H}$.

In this work, the main tool we will be working with is called an *uncertainty func-*

tional. Generally speaking, such a functional is a map of unit vectors to the real numbers. More specifically, it assigns a real number to combinations of the second moments of position and momentum. Uncertainty functionals can be divided into unbounded, trivially bounded (by zero) and bounded ones. To systematically study the lower bounds of the latter, we employ a calculus of variations briefly presented later in this chapter.

2.1 Fractional Fourier transform

The *fractional Fourier transform* of the function $f : \mathbb{R} \to \mathbb{C}$ is defined through [50]

$$\mathscr{F}_\alpha[f](x) = \frac{e^{\frac{i(\pi-2\alpha)}{4}}}{\sqrt{2\pi \sin\alpha}} \int_{\mathbb{R}} dy f(y) \exp\left(\frac{i}{2}\left(-(x^2+y^2)\cot\alpha + 2xy\csc\alpha\right)\right), \tag{2.1}$$

for some real number α called the angle of the transform. Whenever $\alpha = \pi/2$, we recover the standard *Fourier transform* of a function f,

$$\mathscr{F}[f](x) \equiv \mathscr{F}_{\pi/2}[f](x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dy f(y) e^{ixy}, \tag{2.2}$$

while for $\alpha = -\pi/2$, we obtain its inverse

$$\mathscr{F}^{-1}[f](x) \equiv \mathscr{F}_{-\pi/2}[f](x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dy f(y) e^{-ixy}. \tag{2.3}$$

The cases of an angle that is an integer multiple of π, i.e. $\alpha = k\pi$, with $k \in \mathbb{Z}$ in Eq. (2.1) can be understood as a limit. For example, for $\alpha = 0$, substituting α and $1/\alpha$ for $\sin\alpha$ and $\cot\alpha$ respectively and using the fact that

$$\lim_{\epsilon\to 0} \frac{1}{\sqrt{\pi i \epsilon}} e^{-x^2/i\epsilon} = \delta(x), \tag{2.4}$$

then we find

$$\mathscr{F}_0[f](x) = \int_{\mathbb{R}} dy f(y) \delta(y-x) = f(x), \tag{2.5}$$

and thus the fractional Fourier transform for $\alpha = 0$ reduces to the identity operator, as expected. In a similar manner, one can show that for $\alpha = \pi$ it becomes a parity transformation, i.e.

$$\mathscr{F}_{\pi}[f](x) = \int_{\mathbb{R}} dy f(y)\delta(y+x) = f(-x)\,. \tag{2.6}$$

We call the *order*, m, of the fractional Fourier transform, as the ratio of its angle α to the angle of the usual Fourier transform, that is

$$m = \frac{2\alpha}{\pi}\,. \tag{2.7}$$

Thus, the standard Fourier transform is of order 1, the identity is of order 0, while the parity is of order 2; negative orders correspond to inverse transforms. If the order m is a rational number, we can define the *period*, p, of the fractional Fourier as the smallest number of times the transformation needs to be applied before we obtain the original function, or equivalently $(\mathscr{F}_{\alpha})^{p} = \mathbb{I}$, where $\mathbb{I}$ denotes the identity operator. Whenever the angle is of the form $\alpha = 2\pi/n\,,\ n \in \mathbb{Z}\backslash\{0\}$, the period is given by

$$p = \frac{2\pi}{\alpha}\,. \tag{2.8}$$

Note that the usual Fourier transform corresponds to $\alpha = \pi/2$ and is of period 4, while $\alpha = \pi$ corresponds to the parity operator, which is of period 2.

2.2 Wigner function

For a pure state $|\psi\rangle$ and with $\psi(q) = \langle q|\psi\rangle$ denoting the wavefunction, one can define the *Wigner function*, or *Wigner quasi-probability distribution* as [85]

$$W(p,q) = \frac{1}{\pi\hbar}\int_{\mathbb{R}} \psi^{*}(q+y)\psi(q-y)e^{2ipy/\hbar}dy\,, \tag{2.9}$$

where $\psi^{*}(q)$ denotes the complex conjugate of the wave function. The Wigner function is a real quantity and its marginals give the position and momentum distributions,

$$|\psi(q)|^2 = \int_{\mathbb{R}} W(p,q)dp\ , \quad |\tilde{\psi}(p)|^2 = \int_{\mathbb{R}} W(p,q)dq \tag{2.10}$$

where $\tilde{\psi}(p)$ denotes the Fourier transform of the function $\psi(q)$. It cannot be interpreted as a probability distribution since it can assume negative values.

2.3 Measures of uncertainty

We define the *variance* of a self-adjoint operator $\hat{A}$ in a pure state ψ to be

$$\Delta^2 A \equiv \langle\psi|\hat{A}^2|\psi\rangle - \langle\psi|\hat{A}|\psi\rangle^2 = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2 \,. \tag{2.11}$$

Often, the *standard deviation* ΔA will be used, which is obtained by taking the square root of the variance. In the same way, the *covariance* of two non-commuting operators $\hat{A}, \hat{B}$, when the system is in the state ψ, is defined as

$$C_{AB} \equiv \frac{1}{2}\langle\psi|(\hat{A}\hat{B} + \hat{B}\hat{A})|\psi\rangle - \langle\psi|\hat{A}|\psi\rangle\langle\psi|\hat{B}|\psi\rangle \,. \tag{2.12}$$

For a quantum particle in one spatial dimension, we thus have three second moments: the variances of position and momentum and their covariance. It is often convenient to arrange them into the real and symmetric *covariance matrix* **C**,

$$\mathbf{C} = \begin{pmatrix} \Delta^2 p & C_{pq} \\ C_{pq} & \Delta^2 q \end{pmatrix} , \tag{2.13}$$

which is positive definite. A covariance matrix is physically realisable, according to quantum mechanics, if it obeys the condition [67]

$$\mathbf{C} + i\frac{\mathbf{\Omega}}{2} \geq 0 \,, \tag{2.14}$$

where $\mathbf{\Omega}$ is a skew-symmetric matrix, often called the n-symplectic form, explicitly defined in the next section.

In the case of position and momentum, one can simplify the definitions of the variances and the covariance, by employing the invariance of the second moments under phase space translations.

Let $\hat{T}_\alpha = \exp\left(\frac{i}{\hbar}\left(p_0\hat{q} - q_o\hat{p}\right)\right)$ denote the unitary operator that effects translations in phase space, where $\alpha \equiv (q_0 + ip_0)/\sqrt{2\hbar}$ and $p_0, q_0 \in \mathbb{R}$; its action on the position and momentum operators is

$$\hat{T}_\alpha^\dagger \, \hat{q} \, \hat{T}_\alpha = \hat{q} - q_0 \,, \quad \hat{T}_\alpha^\dagger \, \hat{p} \, \hat{T}_\alpha = \hat{p} - p_0 \,. \tag{2.15}$$

It is easy to show that the variances of position and momentum in the states $|\psi\rangle$ and $|\varphi\rangle = \hat{T}_\alpha |\psi\rangle$ are the same, that is

$$(\Delta q)^2_\psi = (\Delta q)^2_\varphi \,, \quad (\Delta p)^2_\psi = (\Delta p)^2_\varphi \,. \tag{2.16}$$

From the last observation it follows that if one chooses $p_0 = \langle \hat{p} \rangle_\psi \,, q_0 = \langle \hat{q} \rangle_\psi$, then the variances in the state $|\varphi\rangle$ reduce to

$$(\Delta q)^2_\varphi = \langle \varphi | \hat{q}^2 | \varphi \rangle \,, \quad (\Delta p)^2_\varphi = \langle \varphi | \hat{p}^2 | \varphi \rangle \,, \tag{2.17}$$

which we will often assume in the following chapters.

Apart from the second moments, one can define the *Shannon entropy* for a probability density ρ,

$$S_q = -\int_{\mathbb{R}} dq \rho(q) \log \rho(q). \tag{2.18}$$

If the probability density is given by $\rho = |\psi(q)|^2$, where $\psi(q)$ is the wave function in position space, then the Shannon entropy is a measure of the localisation of the probability distribution associated with the particle's position; the lower the value of the entropy, the more localised it is.

In a similar way, starting from the probability density in momentum space, $\tilde{\rho} = |\tilde{\psi}(p)|^2$, we define the quantity

$$S_p = -\int_{\mathbb{R}} dp |\tilde{\psi}(p)|^2 \log |\tilde{\psi}(p)|^2, \tag{2.19}$$

with an analogous physical interpretation.

It should be noted that both of these entropies are not bounded and the more concentrated the distribution, the more the Shannon entropy approaches negative infinity.

To see this, take the probability density to be of the form

$$\rho_\epsilon(q) = \frac{1}{2\sqrt{\pi\epsilon}} e^{-\frac{q^2}{4\epsilon}}, \tag{2.20}$$

which in the limit where ϵ approaches zero from positive values, the probability density approaches a delta function. The Shannon entropy of this probability density is found to be

$$S_{q,\epsilon} = -\int_{\mathbb{R}} dq \rho_\epsilon(q) \log \rho_\epsilon(q) = \frac{1}{2}\left(1 + \log(4\pi\epsilon)\right), \tag{2.21}$$

and in the limit of $\epsilon \to 0^+$ it approaches negative infinity, demonstrating the fact that the entropies are not bounded. However, as it will be shown in Chapter 4, the sum of the entropies of the position and momentum distributions is bounded by a positive real number.

2.4 The symplectic group and symplectic transformations

Assume we have a quantum system with n spatial degrees of freedom, each one characterised by a pair of conjugate variables. More specifically, we have a set of $2n$ self-adjoint operators acting on a Hilbert space $\mathscr{H}$, which if arranged in a column vector $\hat{\mathbf{z}}$ as

$$\hat{\mathbf{z}} = \left(\hat{p}_1, \hat{q}_1, \ldots, \hat{p}_n, \hat{q}_n\right), \tag{2.22}$$

they obey the commutation relations

$$[\hat{z}_a, \hat{z}_b] = i\hbar\omega_{ab}; \tag{2.23}$$

ω_{ab} are the entries of the $2n \times 2n$ block diagonal matrix

$$\mathbf{\Omega} = (\omega_{ab}) = \begin{pmatrix} \mathbf{\Omega_2} & & \\ & \ddots & \\ & & \mathbf{\Omega_2} \end{pmatrix}, \tag{2.24}$$

where

$$\mathbf{\Omega_2} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} . \tag{2.25}$$

A linear transformation of the operators $\hat{z}_a$ is called *canonical* if it preserves the commutation relations (2.23), that is, the transformed operators $\hat{z}'_a$ obey

$$\left[\hat{z}'_a, \hat{z}'_b\right] = i\hbar\omega_{ab} \, . \tag{2.26}$$

If the two sets of operators are related through

$$\hat{z}'_a = \sum_b S_{ab}\hat{z}_b, \tag{2.27}$$

where the S_{ab} are real numbers, then the requirement for the transformation to be canonical is equivalent to the matrix $\mathbf{S} = (S_{ab})$ being a symplectic matrix. A matrix is called symplectic, $\mathbf{S} \in Sp(2n, \mathbb{R})$, if it satisfies the condition

$$\mathbf{S}.\mathbf{\Omega}.\mathbf{S}^\top = \mathbf{\Omega} \, . \tag{2.28}$$

For any symplectic matrix $\mathbf{S}$, there is a unitary operator $\mathscr{U}(\mathbf{S})$ implementing the transformation

$$\hat{z}'_a = \sum_b S_{ab}\hat{z}_b = \mathscr{U}(\mathbf{S})^{-1}\hat{z}_a\mathscr{U}(\mathbf{S}) \, . \tag{2.29}$$

The unitary $\mathscr{U}(\mathbf{S})$ is unique up to an arbitrary constant phase factor. One can choose this arbitrary phase factor in such a way to achieve the maximum simplification for the composition rule of unitaries, which turns out to be

$$\mathscr{U}(\mathbf{S}_1)\mathscr{U}(\mathbf{S}_2) = \pm\mathscr{U}(\mathbf{S}_1\mathbf{S}_2) \, , \tag{2.30}$$

with $\mathbf{S}_1, \mathbf{S}_2 \in Sp(2n, \mathbb{R})$. This shows that we have a two-valued representation of the symplectic group. Moreover, these operators give a faithful unitary representation of the metaplectic group, $Mp(2n, \mathbb{R})$, which is the double covering of $Sp(2n, \mathbb{R})$ [8].

There are various ways one can decompose a symplectic matrix in terms of two or three factors drawn from subsets or subgroups of the symplectic group, $Sp(2n, \mathbb{R})$.

Here, we only mention the *Iwasawa* or $\mathscr{K}\mathscr{A}\mathscr{N}$ decomposition for $n = 1$, which is unique and each factor is taken from a subgroup of $Sp(2, \mathbb{R})$. According to this decomposition, a symplectic matrix S,

$$S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{2.31}$$

can be factorised as

$$S = \begin{pmatrix} 1 & \xi \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e^{-\eta/2} & 0 \\ 0 & e^{\eta/2} \end{pmatrix} \begin{pmatrix} \cos \varphi/2 & \sin \varphi/2 \\ -\sin \varphi/2 & \cos \varphi/2 \end{pmatrix}, \tag{2.32}$$

where the parameters ξ, η, φ are given by

$$\begin{aligned} \xi &= \frac{ac + bd}{c^2 + d^2} \in (-\infty, \infty), \\ \eta &= \ln(c^2 + d^2) \in (-\infty, \infty), \\ \varphi &= 2 \arg(d - ic) \in (-2\pi, 2\pi]. \end{aligned} \tag{2.33}$$

When it describes a canonical transformation, the first factor in the decomposition takes $\hat{p}$ to $\hat{p} + \xi\hat{q}$, while leaving $\hat{q}$ unchanged. The second factor corresponds to rescalings or squeezing transformations that take $\hat{p}$ to $\hat{p}/b$ and $\hat{q}$ to $b\hat{q}$, for some positive real number b. Finally, the last factor implements $SO(2)$ rotations that take $\hat{p}$ to $\hat{p} \cos \varphi/2 + \hat{q} \sin \varphi/2$ and $\hat{q}$ to $-\hat{p} \sin \varphi/2 + \hat{q} \cos \varphi/2$.

A theorem by Williamson [8, 84] states that if $\mathbf{F}$ is a $2n \times 2n$ positive or negative definite matrix, then there exists a symplectic matrix $\mathbf{S} \in Sp(2n, \mathbb{R})$ such that

$$\mathbf{F} = \mathbf{S}^\top \mathbf{D} \mathbf{S}, \tag{2.34}$$

where $\mathbf{D} = \mathrm{diag}(\lambda_1, \lambda_1, \ldots, \lambda_n, \lambda_n)$ is known as the *symplectic spectrum* of $\mathbf{F}$ [1, 2], while the n positive real numbers λ_i are its *symplectic eigenvalues*.

Although one could use Williamson's theorem to find the symplectic spectrum of a covariance matrix $\mathbf{C}$, it is often much more convenient to use

$$\mathbf{D} = \mathrm{Eig}_+(i\mathbf{C}\boldsymbol{\Omega}), \tag{2.35}$$

where $\mathrm{Eig}_+(\mathbf{A})$, denotes the diagonal matrix of positive eigenvalues of $\mathbf{A}$.

We conclude this section by mentioning that the condition for a covariance matrix to be physically realisable, as given by Eq. (2.14), can be equivalently restated in terms of the symplectic eigenvalues, as the condition

$$\lambda_i \geq \frac{1}{2}, \quad \forall i = 1, \ldots, n. \tag{2.36}$$

2.5 Gâteaux differential and variational calculus

In this subsection we define the concept of the Gâteaux differential, which generalises the notion of a directional derivative for functions between normed linear spaces. In addition, we briefly review a calculus of variations based on the Gâteaux derivative [60].

Let $F : X \to Y$ be a function between two Banach spaces X, Y. For $x \in X$, the *Gâteaux differential* along the direction $v \in X$ is defined as

$$dF_v(x) = \lim_{\epsilon \to 0} \frac{F(x+\epsilon v) - F(x)}{\epsilon} = \frac{d}{d\epsilon} F(x+\epsilon v)\Big|_{\epsilon=0}. \tag{2.37}$$

There is a stronger notion of differentiation that one can define, called the *Fréchet derivative*. A function $F : X \to Y$ is Fréchet differentiable at $x \in X$ if there exists a bounded linear operator $K : X \to Y$ such that

$$\lim_{h \to 0} \frac{\|F(x+h) - F(x) - Kh\|_Y}{\|h\|_X} = 0. \tag{2.38}$$

Then, the Fréchet derivative is $dF(x) = K$. Whenever they both exist, they are related through $dF_v(x) = dF(x)v$. For our analysis, the notion of the Gâteaux derivative is sufficient.

We define a *functional* $J[x]$ as a mapping of elements of a Banach space to the real numbers, $J : X \to \mathbb{R}$. The *Gâteaux variation* of the functional $J[x]$ at x_0, which we denote by $\delta J[x]$ is given by

$$\delta J[v] = \frac{d}{d\epsilon} J[x_0 + \epsilon v]_{\epsilon=0} \,, \tag{2.39}$$

if it exists for all $v \in X$, where $\epsilon \in \mathbb{R}$. Moreover, it can be shown that it is unique, provided it exists.

In the following chapters, we apply a variational calculus on what we term *uncertainty functionals*, which are essentially smooth real functions of the second moments of position and momentum, as defined in Sec. (2.3).

As a simple example, consider the functional $J[\psi] = \langle \psi | A \psi \rangle$, where $\psi \in \mathscr{H}$ is a vector in a Hilbert space, $\langle .|. \rangle$ denotes the inner product defined in $\mathscr{H}$, and A a self-adjoint operator. With $\epsilon \in \mathbb{R}$ one finds

$$J[\psi + \epsilon\varphi] = \langle \psi + \epsilon\varphi | A(\psi + \epsilon\varphi) \rangle = \langle \psi | A\psi \rangle + \epsilon \left(\langle A\psi | \varphi \rangle + \langle \varphi | A\psi \rangle \right) + \epsilon^2 \langle \varphi | A\varphi \rangle, \tag{2.40}$$

and thus its first variation in terms of the Gâteaux differential is given by

$$\delta J[\varphi] = \frac{d}{d\epsilon} J[\psi + \epsilon\varphi]_{\epsilon=0} = \langle A\psi | \varphi \rangle + \langle \varphi | A\psi \rangle \,, \tag{2.41}$$

One can symbolically restate this by defining

$$D_\varphi = \langle \varphi | \frac{\delta}{\delta \langle \psi |} + \frac{\delta}{\delta | \psi \rangle} | \varphi \rangle \,, \tag{2.42}$$

in terms of which the first variation becomes

$$\delta J[\varphi] = D_\varphi J[\psi] \,. \tag{2.43}$$

Note that the Fréchet differential in the last example exists only if the operator A is bounded, which follows from the following equivalent definition of the Fréchet derivative. If $F : X \to Y$ is a function between Banach spaces, then the linear transformation $K : X \to Y$ is a Fréchet derivative if for every $\epsilon > 0$, there is a $\delta > 0$ such that

$$||F(x+h) - F(x) - Kh||_Y \leq \epsilon ||h||_X \,, \tag{2.44}$$

for all h with $\|h\|_X \leq \delta$.

Applying the last definition to $J[\psi] = \langle\psi|A\psi\rangle$, we find

$$|J(x+h) - J(x) - Kh| = |\langle x|Ax\rangle| \leq \|A\|_{\text{op}}\|h\|^2_{\mathscr{H}}\,, \tag{2.45}$$

where in the last step we used the *Cauchy-Schwarz inequality*,

$$|\langle x|y\rangle| \leq \|x\|\|y\|\,, \tag{2.46}$$

and where the *operator norm* of A is defined as

$$\|A\|_{\text{op}} = \inf\{c \geq 0 : \|Ax\| \leq c\|x\|\,,\, \forall x \in V\} \tag{2.47}$$

and V denotes the normed vector space on which A operates.

From the above considerations, it immediately follows that if A is an unbounded operator, the functional $J[\psi]$ is not Fréchet differentiable. It is however Gâteaux differentiable.

We conclude this section by mentioning the *chain rule* for Gâteaux differentiation. Let $F : X \to Y$ and $G : Y \to Z$ be Gâteaux differentiable functions between Banach spaces, then their composition obeys the chain rule,

$$d_\varphi(FG)[h] = (d_{F(\varphi)}G)[d_\varphi F[h]]. \tag{2.48}$$

For example, consider the functional $J[\psi] = f(\langle\psi|A\psi\rangle)$, where $f : \mathbb{R} \to \mathbb{R}$ is a differentiable function, then its Gâteaux derivative is given by

$$\begin{aligned} d_\varphi J[\psi] &= d_\varphi f(\langle\psi|A\psi\rangle) = f'(\langle\psi|A\psi\rangle)d_\varphi(\langle\psi|A\psi\rangle) \\ &= f'(\langle\psi|A\psi\rangle)\left(\langle A\psi|\varphi\rangle + \langle\varphi|A\psi\rangle\right)\,. \end{aligned} \tag{2.49}$$

Chapter 3

Uncertainty relations for a canonical triple

3.1 Introduction

In quantum theory, two observables $\hat{p}$ and $\hat{q}$ are canonical if they satisfy the commutation relation

$$[\hat{p}, \hat{q}] = \frac{\hbar}{i}, \tag{3.1}$$

with the momentum and position of a particle being a well-known and important example. The non-vanishing commutator expresses the incompatibility of the Schrödinger pair $(\hat{p}, \hat{q})$ of observables since it imposes a lower limit on the product of their standard deviations, namely

$$\Delta q \, \Delta p \geq \frac{\hbar}{2}. \tag{3.2}$$

In 1927, Heisenberg [35] analysed the hypothetical observation of an individual electron with photons and concluded that the product of the measurement errors should be governed by a relation of the form (3.2). His proposal inspired Kennard [41] and Weyl [81] to mathematically derive Heisenberg's uncertainty relation, thereby turning it into a constraint on measurement outcomes for an ensemble of identically prepared systems. Schrödinger's [62] generalization of (3.2) included a correlation term, and Robertson [56, 57] derived a similar relation for any two non-commuting Hermitean operators. Recently claimed violations of an uncertainty relation similar in form to (3.2) do not refer to Kennard and Weyl's *preparation* uncertainty relation but to Heisenberg's *error-disturbance* relation (cf. [52, 31, 58]). However, these claims have been criticized strongly [19, 29].

Uncertainty relations are now understood to provide fundamental limits on what can be said about the properties of quantum systems. Imagine measuring the standard deviations Δp and Δq separately on two ensembles prepared in the same quantum state. Then, the bound (3.2) does not allow one to simultaneously attribute definite values to the observables $\hat{p}$ and $\hat{q}$.

In this chapter, we will consider a *Schrödinger triple* $(\hat{p}, \hat{q}, \hat{r})$ consisting of *three* pairwise canonical observables [75], i.e.

$$[\hat{p}, \hat{q}] = [\hat{q}, \hat{r}] = [\hat{r}, \hat{p}] = \frac{\hbar}{i}, \tag{3.3}$$

and derive a *triple* uncertainty relation. In a system of units where both $\hat{p}$ and $\hat{q}$ carry physical dimensions of $\sqrt{\hbar}$, the observable $\hat{r}$ is given by

$$\hat{r} = -\hat{q} - \hat{p} \tag{3.4}$$

which corresponds to a suitably *rotated* and *rescaled* position operator $\hat{q}$. It is important to point out that any Schrödinger triple for a quantum system with one degree of freedom is unitarily equivalent to $(\hat{p}, \hat{q}, \hat{r})$; furthermore, any such triple is maximal in the sense that there are no four observables that equi-commute to $\hbar/i$ [77]. Therefore, the algebraic structure defined by a Schrödinger triple $(\hat{p}, \hat{q}, \hat{r})$ is unique up to unitary transformations.

Given that (3.1) implies Heisenberg's uncertainty relation (3.2), we wish to determine the consequences of the commutation relations (3.3) on the product of the *three* uncertainties associated with a Schrödinger triple $(\hat{p}, \hat{q}, \hat{r})$.

3.2 Results

We will establish the *triple uncertainty relation*

$$\Delta p \, \Delta q \, \Delta r \ \geq \left(\tau \, \frac{\hbar}{2} \right)^{3/2} , \tag{3.5}$$

where the number τ is the *triple constant* with value

$$\tau = \csc \left(\frac{2\pi}{3} \right) \equiv \sqrt{\frac{4}{3}} \simeq 1.16 \, . \tag{3.6}$$

The bound (3.5) is found to be *tight*; the state of minimal triple uncertainty is found to be a generalized squeezed state,

$$|\Xi_0\rangle = \hat{S}_{\frac{i}{4} \ln 3} \, |0\rangle \, , \tag{3.7}$$

being *unique except for rigid translations in phase space.* The operator $\hat{S}_{\frac{i}{4}\ln 3}$, defined in Eq. (3.22) is a generalized squeezing operator: it generates the state $|\Xi_0\rangle$ by *contracting* the standard coherent state $|0\rangle$ (i.e., the ground state of a harmonic oscillator with unit mass and unit frequency) along the main diagonal in phase space by an amount characterized by $\ln \sqrt[4]{3} < 1$, at the expense of a *dilation* along the minor diagonal.

To visualize this result, let us determine the Wigner function of the state $|\Xi_0\rangle$ with position representation (cf. [48])

$$\langle q | \Xi_0 \rangle = \frac{1}{\sqrt[4]{\tau \pi \hbar}} \exp \left(- \frac{1}{2\hbar} e^{-i \frac{\pi}{6}} q^2 \right) \, . \tag{3.8}$$

Thus, its Wigner function associated with the state $|\Xi_0\rangle$ minimizing the triple uncertainty relation is found to be

$$W_{\Xi_0}(q, p) = \frac{1}{\pi \hbar} \exp \left(- \frac{\tau}{\hbar} \left(q^2 + p^2 + pq \right) \right) , \tag{3.9}$$

which is positive. Its phase-space contour line enclosing an area of size $\hbar$, shown in Fig. 3.1, confirms that we deal with a squeezed state aligned with the minor diagonal.

To appreciate the bound (3.5), let us evaluate the triple uncertainty $\Delta p \Delta q \Delta r$ in two

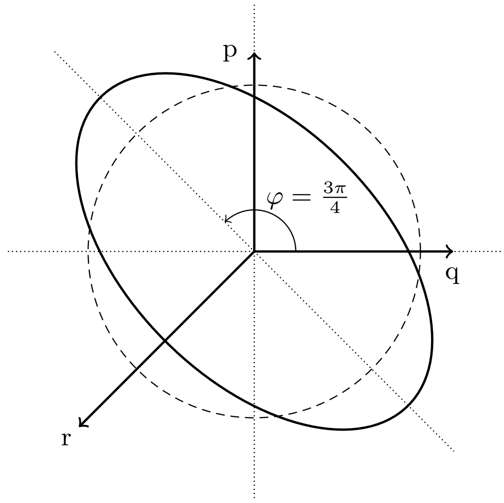

Figure 3.1: Phase-space contour lines of the Wigner functions associated with the states $|\Xi_0\rangle$ (full line) and a standard coherent state $|0\rangle$ (dashed), respectively; both lines enclose the same area.

instructive cases. (i) Since the pairs $(\hat{p}, \hat{q})$, $(\hat{q}, \hat{r})$, and $(\hat{r}, \hat{p})$ are canonical, the inequality (3.2)—as well as its generalization due to Robertson and Schrödinger—applies to each of them implying the lower bound

$$\Delta p \, \Delta q \, \Delta r \geq \left(\frac{\hbar}{2}\right)^{3/2} . \tag{3.10}$$

However, it remains open whether there is a state in which the triple uncertainty saturates this bound. Our main result (3.5) reveals that *no* such state exists. (ii) In the vacuum $|0\rangle$, a coherent state with minimal pair uncertainty, the *triple* uncertainty takes the value

$$\Delta p \, \Delta q \, \Delta r = \sqrt{2} \left(\frac{\hbar}{2}\right)^{3/2} . \tag{3.11}$$

The factor of $\sqrt{2}$ in comparison with (3.10) has an intuitive explanation: while the vacuum state $|0\rangle$ successfully minimizes the product $\Delta p \, \Delta q$, it does not simultaneously minimize the uncertainty associated with the pairs $(\hat{q}, \hat{r})$ and $(\hat{r}, \hat{p})$. Thus, the minimum of the inequality (3.5) cannot be achieved by coherent states.

The observations (i) and (ii) suggest that the bound (3.5) on the triple uncertainty is not an immediate consequence of Heisenberg's inequality for canonical *pairs*, Eq. (3.2). Furthermore, the invariance groups of the triple uncertainty relation, of Heisenberg's

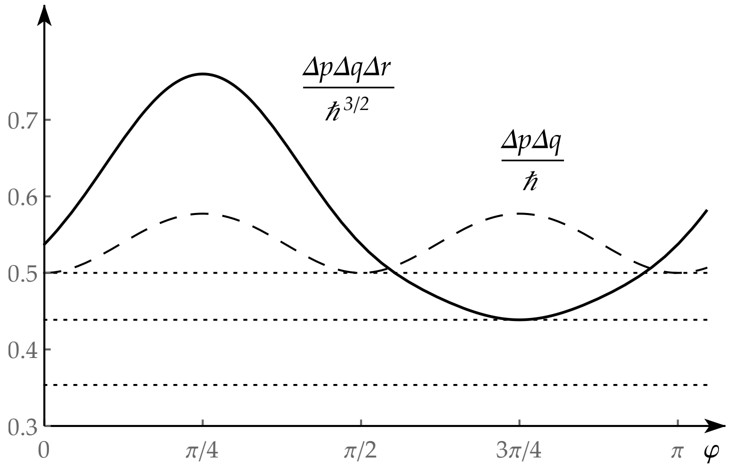

Figure 3.2: Dimensionless pair and triple uncertainties for squeezed states with $\gamma = \ln \sqrt[4]{3}$, rotated away from the position axis by an angle $\varphi \in [0, \pi]$. The pair uncertainty $\Delta p \Delta q$ starts out at its minimum value of $1/2$ which is achieved again for $\varphi = \pi/2$ and $\varphi = \pi$ (dashed line). The triple uncertainty has period π, reaching its minimum for $\varphi = 3\pi/4$ for the state $|\Xi_0\rangle$ (full line). The dotted lines (top to bottom) represent the bounds (3.2), (3.5), and (3.10), with values $1/2$, $(\tau/2)^{3/2}$, and $(1/2)^{3/2}$.

uncertainty relation, and of the inequality by Schrödinger and Robertson are different, because they depend on two, three and four (cf. [73]) continuous parameters, respectively.

3.3 Symmetry of the triple

The commutation relations (3.3) are invariant under the cyclic shift $\hat{p} \to \hat{q} \to \hat{r} \to \hat{p}$, implemented by a unitary operator $\hat{Z}$,

$$\hat{Z}\hat{p}\hat{Z}^\dagger = \hat{q}\,, \quad \hat{Z}\hat{q}\hat{Z}^\dagger = \hat{r}\,, \quad \hat{Z}\hat{r}\hat{Z}^\dagger = \hat{p}\,. \tag{3.12}$$

Note that the third equation follows from the other two equations. The third power of $\hat{Z}$ obviously commutes with both $\hat{p}$ and $\hat{q}$ so it must be a scalar multiple of the identity, $\hat{Z}^3 \propto \mathbb{I}$.

To determine the operator $\hat{Z}$ we first note that its action displayed in (3.12) is achieved by a clockwise rotation by $\pi/2$ in phase space followed by a gauge transformation in the position basis:

$$\hat{Z} = \exp\left(-\frac{i}{2\hbar}\hat{q}^2\right)\exp\left(-\frac{i\pi}{4\hbar}\left(\hat{p}^2+\hat{q}^2\right)\right) . \tag{3.13}$$

A Baker-Campbell-Hausdorff calculation re-expresses this product in terms of a single exponential:

$$\hat{Z} = \exp\left(-i\frac{\pi}{3\hbar\sqrt{3}}\left(\hat{p}^2+\hat{q}^2+\hat{r}^2\right)\right) . \tag{3.14}$$

The operator $\hat{Z}$ cycles the elements of the Schrödinger triple $(\hat{p}, \hat{q}, \hat{r})$ just as a Fourier transform operator swaps position and momentum of the Schrödinger pair $(\hat{p}, \hat{q})$ (apart from a sign). If one introduces a unitarily equivalent symmetric form of the Schrödinger triple with operators $(\hat{P}, \hat{Q}, \hat{R})$ associated with an equilateral triangle in phase space, the metaplectic operator $\hat{Z}$ simply acts as a rotation by $2\pi/3$, i.e., as a fractional Fourier transform.

Furthermore, denoting the factors of $\hat{Z}$ in (3.13) by $\hat{A}$ and $\hat{B}$ (with suitably chosen phase factors), respectively, we find that $\hat{B}^2 = \mathbb{I}$ and $(\hat{A}\hat{B})^3 \equiv \hat{Z}^3 = \mathbb{I}$. These relations establish a direct link between the threefold symmetry of the Schrödinger triple $(\hat{p}, \hat{q}, \hat{r})$ and the *modular group* $SL_2(\mathbb{Z})/\{\pm 1\}$ which $\hat{A}$ and $\hat{B}$ generate [64].

3.4 Experiments

To experimentally confirm the triple uncertainty relation (3.5), we propose an approach based on optical homodyne detection. We exploit the fact that the state $|\Xi_0\rangle$ is a *squeezed state*, also known as a *correlated coherent state* [27]: such a state is obtained by squeezing the vacuum state $|0\rangle$ along the momentum axis followed by a suitable rotation in phase space.

The basic scheme for homodyne detection consists of a beam splitter, photodetectors and a reference beam, called the *local oscillator*, with which the signal is mixed; by adjusting the phase of the local oscillator one can probe different directions in phase space. If θ is the phase of the local oscillator, a homodyne detector measures the probability distribution of the observable

$$\hat{x}(\theta) = \frac{1}{\sqrt{2}} \left(a^{\dagger} e^{i\theta} + a e^{-i\theta} \right) = \hat{q} \cos\theta + \hat{p} \sin\theta \tag{3.15}$$

along a line in phase space defined by the angle θ; here $\hat{q}$ and $\hat{p}$ denote the quadratures of the photon field while the operators $a^{\dagger}$ and a create and annihilate single photons [80]; note that $\hat{r} \equiv \sqrt{2}\,\hat{x}(5\pi/4)$.

The probability distributions of the observables $\hat{q}, \hat{p}$ and $\hat{r}$, corresponding to the angles $\theta = 0$, $\pi/2$, and $5\pi/4$, can be measured upon preparing a large ensemble of the state $|\Xi_0\rangle$. The resulting product of their variances may then be compared with the value of the tight bound given in Eq. (3.5). Under rigid phase-space rotations of the triple $(\hat{q}, \hat{p}, \hat{r})$ by an angle φ the triple uncertainty will vary as predicted in Fig. 3.2 (solid line). A related experiment has been carried out successfully in order to directly verify other Heisenberg- and Schrödinger-Robertson-type uncertainty relations [47, 13].

3.5 Minimal triple uncertainty

To determine the states which minimize the left-hand-side of Eq. (3.5), we need to evaluate it for all normalized states $|\psi\rangle \in \mathscr{H}$ of a quantum particle. To this end we introduce the *uncertainty functional* (cf. [38]),

$$J_{\lambda}[\psi] = \Delta_p[\psi]\,\Delta_q[\psi]\,\Delta_r[\psi] - \lambda(\langle\psi|\psi\rangle - 1)\,, \tag{3.16}$$

using the standard deviations $\Delta_x[\psi] \equiv \Delta x \equiv \left(\langle\psi|\hat{x}^2|\psi\rangle - \langle\psi|\hat{x}|\psi\rangle^2\right)^{1/2}$, $x = p, q, r$, while the term with Lagrange multiplier λ takes care of normalization. In a first step, we determine the extremals of the functional $J_{\lambda}[\psi]$. Changing its argument from $|\psi\rangle$ to the state $|\psi\rangle + |\varepsilon\rangle$, where $|\varepsilon\rangle = \varepsilon|e\rangle$, with a normalized state $|e\rangle \in \mathscr{H}$ and a real parameter $\varepsilon \ll 1$, leads to

$$J_{\lambda}[\psi + \varepsilon] = J_{\lambda}[\psi] + \varepsilon J_{\lambda}^{(1)}[\psi] + \mathscr{O}(\varepsilon^2)\,. \tag{3.17}$$

The first-order variation $J_\lambda^{(1)}[\psi]$ only vanishes if $|\psi\rangle$ is an extremum of the functional $J_\lambda[\psi]$ or, equivalently, if

$$\frac{1}{3}\left(\frac{(\hat{p}-\langle\hat{p}\rangle)^2}{\Delta_p^2}+\frac{(\hat{q}-\langle\hat{q}\rangle)^2}{\Delta_q^2}+\frac{(\hat{r}-\langle\hat{r}\rangle)^2}{\Delta_r^2}\right)|\psi\rangle=|\psi\rangle \tag{3.18}$$

holds, which follows from generalizing a direct computation which had been carried out in [78] to determine the extremals of the product $\Delta p\,\Delta q$.

Eq. (3.18) is non-linear in the unknown state $|\psi\rangle$ due to the expectation values $\langle\hat{p}\rangle, \Delta_p^2$, etc. Its solutions can be found by initially treating these expectation values as constants to be determined only later in a self-consistent way. The unitary operator $\hat{U}_{\alpha,b,\gamma}=\hat{T}_\alpha\hat{G}_b\hat{S}_\gamma$ transforms the left-hand side of (3.18), which is quadratic in $\hat{p}$ and $\hat{q}$, into a standard harmonic-oscillator Hamiltonian,

$$\frac{1}{2}\left(\hat{p}^2+\hat{q}^2\right)|\psi_{\alpha,b,\gamma}\rangle=\frac{3}{2c}|\psi_{\alpha,b,\gamma}\rangle\,, \tag{3.19}$$

where $|\psi_{\alpha,b,\gamma}\rangle\equiv\hat{U}^\dagger_{\alpha,b,\gamma}|\psi\rangle$, and c is a real constant. The unitary $\hat{U}_{\alpha,b,\gamma}$ consists of a *rigid phase-space translation* by $\alpha\equiv(q_0+ip_0)/\sqrt{2\hbar}\in\mathbb{C}$,

$$\hat{T}_\alpha=\exp\left[i\left(p_0\hat{q}-q_0\hat{p}\right)/\hbar\right]\,, \tag{3.20}$$

followed by a *gauge transformation* in the momentum basis

$$\hat{G}_b=\exp\left(ib\hat{p}^2/2\hbar\right)\,,\quad b\in\mathbb{R}\,, \tag{3.21}$$

and a *squeezing transformation*,

$$\hat{S}_\gamma\equiv\exp[i\gamma(\hat{q}\hat{p}+\hat{p}\hat{q})/2\hbar]\,,\quad \gamma\in\mathbb{R}\,. \tag{3.22}$$

According to (3.19), the states $|\psi_{\alpha,b,\gamma}\rangle$ coincide with the eigenstates $|n\rangle, n\in\mathbb{N}_0$, of a harmonic oscillator with unit mass and frequency,

$$|n;\alpha,b,\gamma\rangle\equiv\hat{T}_\alpha\hat{G}_b\hat{S}_\gamma|n\rangle\,,\qquad n\in\mathbb{N}_0\,, \tag{3.23}$$

where we have suppressed irrelevant constant phase factors; for consistency, the quantity $3/2c$ in (3.19) may only take the values $\hbar(n+1/2)$ for $n \in \mathbb{N}_0$, as a direct but lengthy calculation confirms. The parameters b and γ must take specific values for (3.19) to hold, namely

$$b = \frac{1}{2} \qquad \text{and} \qquad \gamma = \frac{1}{2} \ln \tau \, ; \tag{3.24}$$

we will denote the restricted set of states obtained from Eq. (3.23) by $|n;\alpha\rangle$. There are no constraints on the parameter α, which means that we are free to displace the states $|n\rangle$ in phase space without affecting the values of the variances. The variances of the observables $\hat{p}, \hat{q}$, and $\hat{r}$ are found to be equal, taking the value

$$\Delta_x^2[n;\alpha] = \tau\hbar \left(n + \frac{1}{2} \right) , \quad x = p, q, r \, , \tag{3.25}$$

with the triple constant τ introduced in (3.6). Inserting these results into Eq. (3.18) we find that

$$\frac{1}{3} \left(\hat{p}^2 + \hat{q}^2 + \hat{r}^2 \right) |n;\alpha\rangle = \tau\hbar \left(n + \frac{1}{2} \right) |n;\alpha\rangle \, , \tag{3.26}$$

where

$$|n;\alpha\rangle = \hat{T}_\alpha \hat{G}_{\frac{1}{2}} \hat{S}_{\frac{1}{2}\ln\tau} |n\rangle \, , \quad n \in \mathbb{N}_0 \, , \alpha \in \mathbb{C} \, . \tag{3.27}$$

For each value of α, the *extremal states* of the uncertainty functional (3.16) form a complete set of orthonormal states,

$$\sum_{n=0}^{\infty} |n;\alpha\rangle\langle n;\alpha| = \mathbb{I} \, , \tag{3.28}$$

since the set of states $\{|n\rangle\}$ has this property.

At its extremals the uncertainty functional (3.16) takes the values

$$J_\lambda[n;\alpha] = \left[\tau \, \hbar \left(n + \frac{1}{2} \right) \right]^{3/2} , \quad n \in \mathbb{N}_0 \, , \tag{3.29}$$

according to Eq. (3.25), with the minimum occurring for $n = 0$. Thus, the two-parameter family of states $|0;\alpha\rangle, \alpha \in \mathbb{C}$, which we will denote by

$$|\Xi_\alpha\rangle = \hat{T}_\alpha \left(\hat{G}_{\frac{1}{2}} \hat{S}_{\frac{1}{2}\ln\tau} |0\rangle \right) , \tag{3.30}$$

minimize the triple uncertainty relation (3.5).

The states $|\Xi_\alpha\rangle$ are *displaced generalized squeezed* states, with a squeezing direction along a line different from the position or momentum axes. To show this, it is sufficient to consider the state $|\Xi_0\rangle$, which satisfies (3.26) with $n \equiv 0$ and $\alpha \equiv 0$. The product of unitaries in (3.30) acting on the vacuum $|0\rangle$ is easily understood if one rewrites it using the identity

$$\hat{G}_b \hat{S}_\gamma = \hat{S}_\xi \hat{R}_\varphi \,, \tag{3.31}$$

where the unitary $\hat{R}_\varphi = \exp(i\varphi a^\dagger a)$ is a counterclockwise rotation by φ in phase space while the operator

$$\hat{S}_\xi = \exp\left[\left(\bar{\xi} a^2 - \xi a^{\dagger 2}\right)/2\right], \quad \xi = \gamma e^{i\theta}, \gamma > 0 \,, \tag{3.32}$$

generalizes $\hat{S}_\gamma$ in (3.22) by allowing for squeezing along a line with inclination $\theta/2$; the annihilation operator and its adjoint $a^\dagger$ are defined by $a = (\hat{q} + i\hat{p})/\sqrt{2\hbar}$. Another standard Baker-Campbell-Hausdorff calculation (using result from Sec. 6 of [83]) reveals that the values $\xi = (i/4)\ln 3$ and $\varphi = -\pi/12$ turn Eq. (3.31) into an identity for the values of b and γ given in (3.24). This confirms that the state of minimal triple uncertainty is the generalized squeezed state given in (3.7).

3.6 Discussion

We have established a tight inequality (3.5) for the triple uncertainty associated with a Schrödinger triple $(\hat{p}, \hat{q}, \hat{r})$ of pairwise canonical observables. Ignoring rigid translations in phase space, there is only one state $|\Xi_0\rangle$ which minimizes the triple uncertainty, shown in Eq. (3.30). The state $|\Xi_0\rangle$ is an eigenstate of the operator $\hat{Z}$ in (3.14) which describes the fundamental threefold cyclic symmetry of the Schrödinger triple $(\hat{p}, \hat{q}, \hat{r})$. Conceptually, the triple uncertainty and the one derived by Schrödinger and Robertson are linked because both incorporate the operator $(\hat{p}\hat{q} + \hat{q}\hat{p})/2$, be it explicitly or indirectly via the expression $\hat{r}^2$.

The smallest possible value of the product $\Delta p \Delta q \Delta r$ is noticeably *larger* than the unachievable value $(\hbar/2)^{3/2}$, which follows from inequality (3.2) applied to each of the Schrödinger pairs $(\hat{p}, \hat{q})$, $(\hat{q}, \hat{r})$, and $(\hat{r}, \hat{p})$. At the same time, the true minimum *undercuts* the value of the triple uncertainty in the vacuum state $|0\rangle$ by more than 10% [cf. Eq. (3.11)].

The results obtained in this chapter add another dimension to the problem of earlier attempts to obtain uncertainty relations for more than two observables. In 1934, Robertson studied constraints which follow from the positive semi-definiteness of the covariance matrix for N observables [57] but the resulting inequality trivializes for an odd number of observables. Shirokov obtained another inequality [65] which contains little information about the canonical triple considered here.

The result for a Schrödinger triple obtained here suggests conceptually important generalizations. A tight bound for an *additive* uncertainty relation associated with the operators $(\hat{p}, \hat{q}, \hat{r})$ is easily established by a similar approach: the inequality

$$(\Delta p)^2 + (\Delta q)^2 + (\Delta r)^2 \geq \tau \frac{3\hbar}{2} \tag{3.33}$$

is saturated only by the state $|\Xi_0\rangle$ in (3.30), ignoring irrelevant rigid phase-space translations. This observation contrasts the relation between the *additive* and the *multiplicative* uncertainty relations for Schrödinger *pairs* $(\hat{p}, \hat{q})$. According to [73] the states saturating the inequality $(\Delta p)^2 + (\Delta q)^2 \geq \hbar$ are a proper subset of those minimizing Heisenberg's *product* inequality (3.2).

An uncertainty relation for pairs of canonical observables also exists for the Shannon entropies S_p and S_q of their probability distributions [37, 14]. We conjecture that the relation $S_p + S_q + S_r \geq (3/2)\ln(\tau e \pi)$ holds for the Schrödinger triple $(\hat{p}, \hat{q}, \hat{r})$, the minimum being achieved by the state $|\Xi_0\rangle$. This bound is tighter than $(3/2)\ln(e\pi)$, the value which follows from applying the bound $\ln(e\pi)$ for pairwise entropies to the triple.

It is a known fact that the maximum Shannon entropy of a distribution with a given variance is that of a Gaussian [14]. For the position distribution, for example, maximising

$$J = -\langle \ln \rho \rangle - \lambda(\langle 1 \rangle - 1) - \mu(\Delta^2 q - r_0^2)\,, \tag{3.34}$$

where

$$-\langle \ln \rho \rangle \equiv S_q = -\int \rho(q) \ln \rho(q) dq\,, \tag{3.35}$$

denotes the Shannon entropy, leads to the inequality

$$\Delta^2 q \geq \frac{1}{2e\pi} e^{2S_q}, \tag{3.36}$$

with equality only for Gaussians, and similarly for the distributions of the other two. Combining the three inequalities, we obtain

$$\Delta^2 p \cdot \Delta^2 q \cdot \Delta^2 r \geq \frac{1}{(2e\pi)^3} e^{2(S_p+S_q+S_r)}. \tag{3.37}$$

For the state $|\Xi\rangle$, that minimises the variances, the sum of the entropies is equal to $\ln(\tau e\pi)^{3/2}$ and thus for all Gaussian states it holds that

$$S_p + S_q + S_r \geq \ln(\tau e\pi)^{3/2}, \tag{3.38}$$

which we conjecture for all quantum states.

Chapter 4

Uncertainty relations for canonical structures and beyond

4.1 Introduction

In Chapter 3 we derived bounds for the product and sum of the variances of three canonical operators, for a particle in one spatial dimension. The triple consisted of the position operator of the particle, its momentum and an operator defined by $\hat{r} = -\hat{p} - \hat{q}$, for which we showed that the product of the variances can never be less than the constant $\left(\tau\hbar/2\right)^{3/2}$, Eq. (3.5). Representing the three operators in the triple as vectors in $\mathbb{R}^2$ and connecting their tips, they form an isosceles triangle. However, in accordance to what was mentioned in the previous chapter, one can go to a symmetric triple of operators corresponding to an equilateral triangle. In that case, each operator is obtained from the previous by a rotation by $2\pi/3$; this symmetric triple associated with the regular triangle can be obtained from the non-symmetric one by means of a unitary transformation.

Naturally, one may ask whether we can go beyond a canonical triple, or geometrically speaking, why restrict to a regular triangle and not consider other *regular polygons* in phase space? In this chapter we respond to this question and we build on the ideas presented for the triple. It should be noted that there is one difference when considering higher regular polygons in comparison to the triangle: there are only three operators that pairwise commute to $\pm i\hbar$ [77]. However, relaxing the assumptions that all operators are pairwise canonical and only demanding that *neighbouring* ones are, we can extend the results of the last chapter to the case of more than three observables.

This type of generalisation will be the focus of next section, where uncertainty relations for the variances of operators corresponding to other regular canonical structures, e.g. squares, pentagons and so on, will be derived. As we will show, the lower bounds of those inequalities allow for a geometric interpretation in terms of the associated polygon.

The assumption that the considered operators form regular polygons will be subsequently lifted, and we will derive inequalities for N observables that are arbitrary linear combinations of position and momentum; we will show that the lower bound can be expressed in terms of all commutators. Moreover, it can be seen as the area of the parallelogram spanned by two vectors in $\mathbb{R}^N$, formed by the coefficients of position and momentum in each observable.

Moving away from inequalities concerning discrete sums of the variances of operators, we next prove an uncertainty relation for the integral of variances of rotated observables through an angle φ. For the values $\varphi = \pi\,,\ 2\pi\,$, we find that the lower bound agrees with the area of the half and full unit circle, respectively, but for all other values, this ceases to be the case. We show that there is only one squeezed state reaching the lower bound, for each value of the angle.

In the last two sections, we state inequalities for the case of observables in one or more spatial degrees of freedom and provide examples of how they could be used for entanglement detection.

4.2 Regular canonical polygons

Let us represent the operators of position and momentum by orthogonal unit vectors in $\mathbb{R}^2$ and any other combination of them, $\hat{r} = a\hat{p} + b\hat{q}$ with $a, b \in \mathbb{R}$, by the vector with coordinates (a, b). Using this definition, the triple $(\hat{p}, \hat{q}, \hat{r})$ defined in (3.3), for example, corresponds to three vectors whose tips are the vertices of an isosceles triangle. Let us call the "*length*" of the observable $\hat{r}$, the magnitude of the associated vector. Any observable of the form of $\hat{r}$ can be obtained from position or momentum by a suitable symplectic transformation effected by a unitary (metaplectic) operator. The Iwasawa

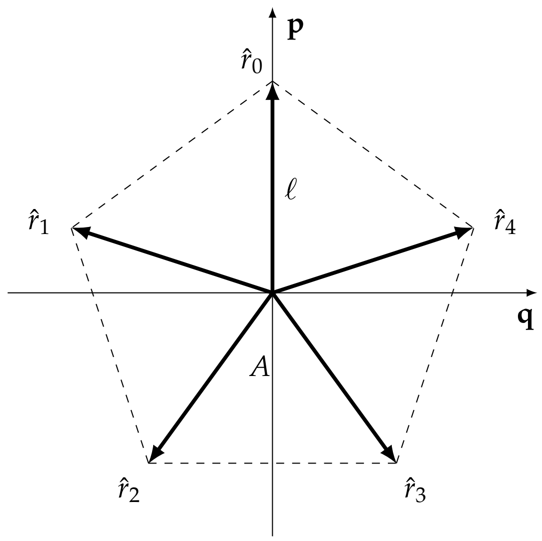

Figure 4.1: A regular pentagon of circumradius ℓ and area A, associated with five canonical operators in phase space.

decomposition of $Sp(2, \mathbb{R})$ (see Chapter 2) suggests that the three primary symplectic transformations are squeezing, gauge and rotations, since any general transformation can uniquely be expressed as a product of the three. In the case of observables of equal length, however, which will be the focus of this section, rotations suffice to transform one to the other.

Let us now consider N observables forming a regular polygon in the two dimensional space of their coefficients, such that neighbouring ones make canonical pairs; this last constraint completely determines their length. We choose the first to be of the form $\hat{r}_0 = \ell\hat{p}$ and the remaining $N-1$ operators are obtained by rotating anticlockwise by angles $2\pi k/N$, $k = 1, \ldots, N-1$, with $N \geq 3$. For example, for $N = 5$ we have a *canonical pentagon* shown in Fig. (4.1). The arbitrary choice of selecting a multiple of momentum as the first is of no importance, since we can always rotate the whole configuration by an arbitrary angle. The requirement that neighbouring ones are canonical pairs, i.e.

$$[\hat{r}_j, \hat{r}_{j+1}] = i\hbar \,, \tag{4.1}$$

gives their length, which turns out to be $\ell = \left(\sin\left(\frac{2\pi}{N}\right)\right)^{-1/2}$. Thus the N observables take the explicit form

$$\hat{r}_j = \ell \left(\hat{p} \cos\left(\frac{2\pi j}{N}\right) - \hat{q} \sin\left(\frac{2\pi j}{N}\right) \right), \quad j = 0, \dots, N-1, \tag{4.2}$$

with variances equal to

$$\Delta^2 \hat{r}_j = \ell^2 \left(\Delta^2 \hat{p} \cos^2\left(\frac{2\pi j}{N}\right) + \Delta^2 \hat{q} \sin^2\left(\frac{2\pi j}{N}\right) - C_{pq} \sin\left(\frac{4\pi j}{N}\right) \right). \tag{4.3}$$

We first focus on the sum of the variances, a case which is quite trivial. Using the trigonometric identities

$$\sum_{j=0}^{N-1} \cos^2\left(\frac{2\pi j}{N}\right) = \sum_{j=0}^{N-1} \sin^2\left(\frac{2\pi j}{N}\right) = \frac{N}{2}, \qquad \sum_{j=0}^{N-1} \sin\left(\frac{4\pi j}{N}\right) = 0, \tag{4.4}$$

one finds that

$$\sum_{j=0}^{N-1} \Delta^2 r_j = \frac{N\ell^2}{2} \left(\Delta^2 p + \Delta^2 q \right) \geq \frac{N\ell^2}{2} \hbar = \frac{N\hbar}{2 \sin\left(\frac{2\pi}{N}\right)}, \tag{4.5}$$

by using the sum inequality for the variances of position and momentum, $\Delta^2 q + \Delta^2 p \geq \hbar$. Observing that the *circumradius* of the regular polygon is equal to ℓ and its *area* is given by $A = \frac{N\ell^2}{2} \sin \frac{2\pi}{N}$, the sum inequalities can be rewritten in the geometrical form

$$\sum_{j=0}^{N-1} \Delta^2 r_j \geq \frac{N\ell^2 \hbar}{2} = \frac{A\hbar}{\sin\left(\frac{2\pi}{N}\right)}. \tag{4.6}$$

We see that the lower bound is equivalent to the area of $N/2$ squares of side ℓ, or equivalently to the sum of the areas of N triangles associated with a conjugate pair like position and momentum. Alternatively, we see that the lower bound is equal to the area of the regular polygon divided by the sine of the angle $2\pi/N$; this is to be expected since the area of the regular polygon corresponds only to neighbouring commutators and ignores the rest.

The corresponding inequality for the product of N pairwise canonical observables is found to be

$$\prod_{j=0}^{N-1} \Delta^2 r_j \geq \frac{\hbar^N}{2^N \sin\left(\frac{2\pi}{N}\right)^N}, \tag{4.7}$$

which, in accordance with Eq. (4.6) can be expressed in geometric terms as

$$\prod_{j=0}^{N-1} \Delta^2 r_j \geq \left(\frac{\hbar \ell^2}{2}\right)^N . \tag{4.8}$$

They both can be obtained from the sum inequality, Eq. (4.5), by an optimisation method. More specifically, one has to look for the minimum of the product of the N variances under the constraint that they obey the sum inequality, Eq. (4.5). This is then an optimisation problem with an inequality constraint and can be dealt with the aid of the Karush-Kuhn-Tucker (KKT) conditions [59]. In more detail, looking for the local minima of a functional $J(\vec{x})$, subject to the inequality constraints $g(\vec{x}) \leq 0$ is equivalent to the conditions:

$$\begin{aligned} \frac{\partial J}{\partial \vec{x}} + \mu^\top \frac{\partial g}{\partial \vec{x}} &= 0\,, \\ \mu^\top g &= 0\,. \end{aligned} \tag{4.9}$$

Let us first set $x_i = \Delta^2 r_i$ for notational simplicity and rewrite Eq. (4.5) as $\sum x_i \geq \frac{N}{2a}$, where $a = \sin\left(\frac{2\pi}{N}\right)/\hbar$. We define the functional

$$J = \prod_j x_j + \mu \left(\frac{N}{2a} - \sum_j x_j\right) , \tag{4.10}$$

and we look for its minimum through the KKT equations, (4.9), whose solution is

$$x_j = x_k = \frac{1}{2a}\,, \quad \mu = \left(\frac{1}{2a}\right)^{N-1} , \quad \forall k, j\,. \tag{4.11}$$

Thus, the minimum is found to be $\left(\frac{1}{2a}\right)^N$, which leads to Eq. (4.7).

The states that saturate such inequalities are the usual coherent states $|\alpha\rangle = \hat{T}_\alpha|0\rangle$ (with the exception of the case $N = 4$), which is to be expected since by fixing all the N directions of equal length, we break scale invariance. The case $N = 4$ is special since the four observables are $\hat{q}, \hat{p}, -\hat{q}, -\hat{p}$ and the sum inequality reads $\Delta^2 q + \Delta^2 p \geq \hbar$. Using the KKT conditions to find the minimum for the product, we do not get the full class of states that attain the minimum but only the restricted class of coherent states where the variances are equal. This is due to the fact that we considered as a constraint only a special case of the general sum inequality $\Delta^2 q/\lambda + \lambda\Delta^2 p \geq \hbar\,,\ \ \lambda > 0\,.$ Had we used the latter, we would have found the full class of squeezed states with real squeezing parameter that are known to saturate Heisenberg's inequality. This distinction only holds for the case $N = 4$, and for all other $N \geq 3$, the sum and product inequalities have the same class of minimising states. However, all even cases are in a certain sense trivial, since the considered set of operators can be split into two, one obtained from the other by a change of signs. The variances of the corresponding probability distributions for each pair differing only by a sign are the same and are counted twice.

4.3 Symmetry of N canonical operators and entropic inequalities

In analogy to the canonical triple of Chapter 3, there is a discrete symmetry associated with the N canonical operators. In the symmetric arrangement we considered in the last section, which is unique up to unitary transformations, the N operators are associated with a regular polygon. In this symmetric case, the commutation relations are preserved under rotations by angles $2\pi k/N$, with $k = 1, \ldots, N$, which only cyclically permute the operators. The unitary effecting (anticlockwise) rotations by an angle φ is given by the exponential of the harmonic oscillator Hamiltonian of unit mass and frequency:

$$\hat{R}(\varphi) = \exp\left(-\frac{i\varphi}{2\hbar}\left(\hat{p}^2 + \hat{q}^2\right)\right). \tag{4.12}$$

Thus, in relation to the N canonical operators case, the operator that effects cyclic permutations, $\hat{Z}_N$, is given by

$$\hat{Z}_N = \hat{R}\left(\frac{2\pi}{N}\right), \tag{4.13}$$

which is effectively just a fractional Fourier transform. To see this, let $|x_q\rangle$ denote the position basis, i.e. $\psi_q(x) = \langle x|\psi\rangle$, where $\psi_q(x)$ denotes the wavefunction and let

$$\begin{aligned}\psi_{q_\alpha}(x) = \langle x_{q_\alpha}|\psi\rangle = \langle x_q|\hat{R}^\dagger(\alpha)|\psi\rangle &= \int dy \langle x_q|\hat{R}^\dagger(\alpha)|y_q\rangle\langle y_q|\psi\rangle \\ &= \int dy \langle x_q|\hat{R}^\dagger(\alpha)|y_q\rangle \psi_q(x)\,,\end{aligned} \tag{4.14}$$

where we have introduced an identity. Using the results in [8], one can explicitly determine the "Huygens kernel" , $\langle x_q|\hat{R}^\dagger(\alpha)|y_q\rangle$, which is found to be

$$\langle x_q|\hat{R}^\dagger(\alpha)|y_q\rangle = \frac{e^{-\frac{i\pi}{4}}}{\sqrt{2\pi\hbar|\sin\alpha|}} \exp\left(\frac{i}{2\hbar}\left(-(x^2+y^2)\cot\alpha + 2xy\csc\alpha\right)\right). \tag{4.15}$$

This is of the same form as the definition of the fractional Fourier transform in Chapter 2, up to a constant phase factor and a redefinition of α, and thus $\psi_{q_\alpha}(x) = \mathscr{F}_\alpha[\psi(x)]$.

As mentioned in Chapter 2, the fractional Fourier transform is of *order* $m = 2\alpha/\pi$, which means that its *period* is equal to $p = 2\pi/\alpha$, i.e. applying the fractional Fourier transform p times we obtain the original function or equivalently $(\mathscr{F}_\alpha)^p = \mathbb{1}$. In addition, the usual Fourier transform, which corresponds to $\alpha = \pi/2$ is of period 4, while $\alpha = \pi$ corresponds to the parity operator. The operator that cyclically shifts through the elements of the triple of Chapter 3 is of period 3, while in the case of N canonical observables it is N. Listing the corresponding fractional Fourier transforms for all N with increasing periods, we see that the triple is unique in that it is the only possible N-tuple with (integral) period less than that of the position-momentum case, which are related through the usual Fourier transform.

Furthermore, observe that for $N = 3$ or in general the cases where N is a prime number are different from the rest, since the only symmetry present is that of period N. On the other hand, whenever N is not a prime number, every divisor of N leads to a discrete symmetry implemented with some fractional Fourier transform, for subsets of operators. For example, for $N = 9$ both $\mathscr{F}_{2\pi/9}$ and $\mathscr{F}_{2\pi/3}$ are present; the first leaves invariant the whole set of operators, while the second sets of three of them.

We conclude this section with a conjecture related to the fractional Fourier transforms and the connection with entropic uncertainty relations. First, recall that the relation

$$S_p + S_q \geq \ln(e\pi)\,, \tag{4.16}$$

can be proven from the (p,q)-*norm* of the Fourier transform [14], i.e. the best estimate $k(p,q)$ such that

$$\|\tilde{\psi}\|_q \leq k(p,q)\,\|\psi\|_p\,, \tag{4.17}$$

where the p-norm is given by

$$\|\psi\|_p = \left(\int dx |\psi|^p \right)^{1/p} . \tag{4.18}$$

The following condition must hold

$$\frac{1}{p} + \frac{1}{q} = 1 , \tag{4.19}$$

with $q \geq 2$ and $k(p,q)$ is explicitly given by [9, 11]

$$k(p,q) = \left(\frac{2\pi}{q} \right)^{1/2q} \left(\frac{2\pi}{p} \right)^{-1/2p} . \tag{4.20}$$

Defining $W(q) = k(p,q) \|\psi\|_p - \|\tilde{\psi}\|_q \geq 0$ and noting that $W(2) = 0$ due to the Parseval-Plancherel theorem, it follows that the derivative of $W(q)$ at $q = 2$ must be non-negative, from which one obtains the desired inequality on entropies. Thus, we see that the entropic uncertainty relation, Eq. (4.16), is a consequence of the fact that the distributions of position and momentum are related by a Fourier transform. We expect that the case of N observables which are related by a fractional Fourier transform will lead to an equivalent statement, which we show for Gaussians and conjecture to hold for all states.

Using the inequality

$$\Delta^2 p_\theta \geq \frac{1}{2e\pi} e^{2S_{p_\theta}} , \tag{4.21}$$

from Chapter 3, which holds as an equality if the given state is Gaussian, and combining it with (4.7) of last section, we find that

$$\sum_{k=0}^{N-1} S_{p_k} \geq \ln \left(\tau_N e\pi \right)^{N/2} , \tag{4.22}$$

where $\tau_N = \csc\left(\frac{2\pi}{N}\right)$ and S_{p_k} denotes the Shannon entropies of the distribution associated with the operator $\hat{p}_k$. Inequality (4.22) holds for all Gaussian states, with one of them achieving the minimum value, and we conjecture that this holds for all quantum states.

4.4 *N* observables of arbitrary "length"

The inequalities derived in the previous section can be further generalised by relaxing the requirement that they form a regular polygon. We consider N arbitrary observables, defined by

$$\hat{r}_i = \alpha_i \hat{p} + \beta_i \hat{q}\,, \quad i = 0, \ldots, N-1\,, \tag{4.23}$$

with $\alpha_i, \beta_i \in \mathbb{R}$. Their variances are given by

$$\Delta^2 r_i = \alpha_i^2 \Delta^2 p + \beta_i^2 \Delta^2 q + 2\alpha_i \beta_i C_{pq}\,, \tag{4.24}$$

and by taking their sum, we obtain

$$\begin{aligned} \sum_i \Delta^2 r_i &= \left(\sum_i \alpha_i^2\right) \Delta^2 p + \left(\sum_i \beta_i^2\right) \Delta^2 q + 2\left(\sum_i \alpha_i \beta_i\right) C_{pq} \\ &= Ax + By + 2\Gamma z\,. \end{aligned} \tag{4.25}$$

Employing an inequality that is proved in the next chapter, Eq. (5.74), we find

$$\sum_i \Delta^2 r_i \geq \hbar\sqrt{AB - \Gamma^2} = \hbar\sqrt{\left(\sum_i \alpha_i^2\right)\left(\sum_i \beta_i^2\right) - \left(\sum_i \alpha_i \beta_i\right)^2}\,. \tag{4.26}$$

Interestingly, one can equivalently rewrite this in terms of their commutators as

$$\sum_i \Delta^2 r_i \geq \sqrt{\sum_{i>j} \left|\langle [\hat{r}_i, \hat{r}_j] \rangle\right|^2}\,. \tag{4.27}$$

Thus, the sum of the variances of N operators for one degree freedom is bounded from below by the square root of the sum of the squares of all possible commutators between the observables.

To see this, first note that

$$\begin{aligned}
\sum_{i>j} \left| \langle [\hat{r}_i, \hat{r}_j] \rangle \right|^2 &= \hbar^2 \sum_{i>j} \left| (\alpha_i \beta_j - \alpha_j \beta_i) \right|^2 = \hbar^2 \sum_{i>j} \left(\alpha_i \beta_j - \alpha_j \beta_i \right)^2 \\
&= \hbar^2 \sum_{i>j} \left(\alpha_i^2 \beta_j^2 + \alpha_j^2 \beta_i^2 - 2\alpha_i \alpha_j \beta_i \beta_j \right) \\
&= \hbar^2 \frac{1}{2} \sum_i^N \sum_j^N \left(2\alpha_i^2 \beta_j^2 - 2\alpha_i \alpha_j \beta_i \beta_j \right) \\
&= \hbar^2 \left(\left(\sum_i \alpha_i^2 \right) \left(\sum_i \beta_i^2 \right) - \left(\sum_i \alpha_i \beta_i \right)^2 \right) ,
\end{aligned} \tag{4.28}$$

which is the square of $\hbar\sqrt{AB - \Gamma^2}$ of (4.26).

Finally, the product can be obtained from the sum inequality through the KKT conditions as in the last section, and is found to be

$$\prod_i \Delta^2 r_i \geq \left(\sum_{i>j} \left| \langle [\hat{r}_i, \hat{r}_j] \rangle \right|^2 \right)^{N/2} . \tag{4.29}$$

It is worth noting that the bound on the right hand side of Eq. (4.27) is equivalent to the area (multiplied by $\hbar$) of the parallelogram spanned by the two N-dimensional vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^N$, given by

$$\mathbf{a} = (\alpha_1, \ldots, \alpha_N) , \quad \mathbf{b} = (\beta_1, \ldots, \beta_N) , \tag{4.30}$$

i.e. $\mathbf{a}$ being a vector in $\mathbb{R}^N$ formed from the coefficients of momentum of each observable, while $\mathbf{b}$ being the one formed by the coefficients of position. This becomes apparent by first noting that the right hand side of Eq. (4.27) for the case of three observables reduces to the magnitude of the vector $\mathbf{c} = \mathbf{a} \times \mathbf{b}$,

$$\Delta^2 r_0 + \Delta^2 r_1 + \Delta^2 r_2 \geq \hbar \left| \mathbf{a} \times \mathbf{b} \right| . \tag{4.31}$$

In general, we can express inequality (4.27) in terms of the vectors of the coefficients of position and momentum according to

$$\sum_i \Delta^2 r_i \geq \hbar \sqrt{\|\mathbf{a}\|^2 \|\mathbf{b}\|^2 - (\mathbf{a} \cdot \mathbf{b})^2} = \hbar \sqrt{\|\mathbf{a}\|^2 \|\mathbf{b}\|^2 - \|\mathbf{a}\|^2 \|\mathbf{b}\|^2 \cos^2 \theta}\,, \tag{4.32}$$

or finally

$$\sum_i \Delta^2 r_i \geq \hbar \|\mathbf{a}\| \|\mathbf{b}\| \sin \theta\,, \tag{4.33}$$

where θ is the angle between the vectors $\mathbf{a}, \mathbf{b}$.; note that this can also be expressed in terms of the wedge product, $\mathbf{a} \wedge \mathbf{b}$, which generalises the vector product for arbitrary dimensional vectors. We observe that the formulation for three observables (or in general N) resembles in structure the inequality for the sum of standard deviations of two ± 1-valued observables [20]: $\Delta A + \Delta B \geq |\mathbf{a} \times \mathbf{b}|$, where $A = \mathbf{a} \cdot \boldsymbol{\sigma}$ and $B = \mathbf{b} \cdot \boldsymbol{\sigma}$, with $\mathbf{a}, \mathbf{b}$ unit vectors and $\boldsymbol{\sigma} = \left(\sigma_x, \sigma_y, \sigma_z\right)^\top$ is the Pauli matrix vector.

It follows from the last equation that the bound is trivial only if $\mathbf{a}, \mathbf{b}$ are parallel, which equivalently means that all the original $\hat{r}_i$ commute. Conversely, the bound is maximal whenever the two vectors are orthogonal, or equivalently, the $\hat{r}_i$ are *maximally incompatible*. As we will demonstrate, when they are of equal length, this means that they have to lay along the maximally incompatible directions, of which one specific constellation forms a regular polygon, and all others can be obtained from it by sign flips.

Let us examine the case of three observables. The assumption that their associated vectors are of equal length is mathematically expressed through

$$a_i^2 + b_i^2 = \ell^2\,, \quad i = 0, 1, 2\,. \tag{4.34}$$

Due to this constraint, it is convenient to switch to polar coordinates

$$a_i = \ell \cos \theta_i\,, \quad b_i = -\ell \sin \theta_i\,, \tag{4.35}$$

and the square of the right hand side of (4.5) becomes

$$
\begin{aligned}
I &= \sum_{i>j}(a_i b_j - a_j b_i)^2 = \ell^2 \sum_{i>j}(\sin\theta_i \cos\theta_j - \sin\theta_j \cos\theta_i)^2 = \ell^2 \sum_{i>j} \sin^2\left(\theta_i - \theta_j\right) \\
&= \frac{3\ell^2}{2} - \ell^2 \sum_{i>j} \cos\left(2\left(\theta_i - \theta_j\right)\right) = \ell^2 \left(\frac{3}{2} - \cos x - \cos y - \cos\left(x+y\right)\right) = \ell^2 f(x,y)\,,
\end{aligned}
\tag{4.36}
$$

where $x = 2(\theta_1 - \theta_2), y = 2(\theta_2 - \theta_3)$ and we have also used the trigonometric identity $\sin^2\theta = (1-\cos 2\theta)/2$. Looking for the maxima of $f(x,y)$, one finds that they occur for $x = \pm 2\pi/3 + 2\pi k_1$ and $y = \pm 2\pi/3 + 2\pi k_2$, where k_1, k_2 are integers. This equivalently leads to the condition

$$
(\theta_i - \theta_{i+1}) = \pm\frac{\pi}{3} - \pi k_{i+1}\,, \quad i = 0, 1\,,
\tag{4.37}
$$

for maximal incompatibility. Due to the invariance in rotating the whole configuration, we select the first observable to be momentum and since we measure angles from its vector, this implies $\theta_0 = 0$. All solutions of the last equation are visualised in Fig. 4.2 up to relabellings of $\hat{r}_1$ and $\hat{r}_2$. For each maximal triple, the resulting $\mathbf{a}, \mathbf{b}$ vectors are perpendicular.

The centre graphic of Fig. 4.2 depicts the directions of maximal incompatibility: selecting the first observables, all maximal triples are constructed by selecting the remaining two so that there does not exist a pair on the same line. If we select momentum as the first, then the other two can be selected in four ways in total, according to Fig. 4.2. However, note that (A,B,D) are related just by a rotation (and relabelling), while the regular structure associated with a (symmetric) triple depicted in (C) is different. The corresponding vectors $\mathbf{a}, \mathbf{b}$ defined in Eq. (4.30) are shown in Fig. 4.3.

The physical significance of this result can be expressed intuitively in the following way. Assume we have an experiment in which we can only probe three arbitrary directions in phase space with the aim of approximately reconstructing the state of the given quantum system. After gathering the approximate probability distributions associated with the three chosen directions, one could apply some standard state reconstruction algorithm, e.g. "maximum likelihood" [26], to obtain an approximation

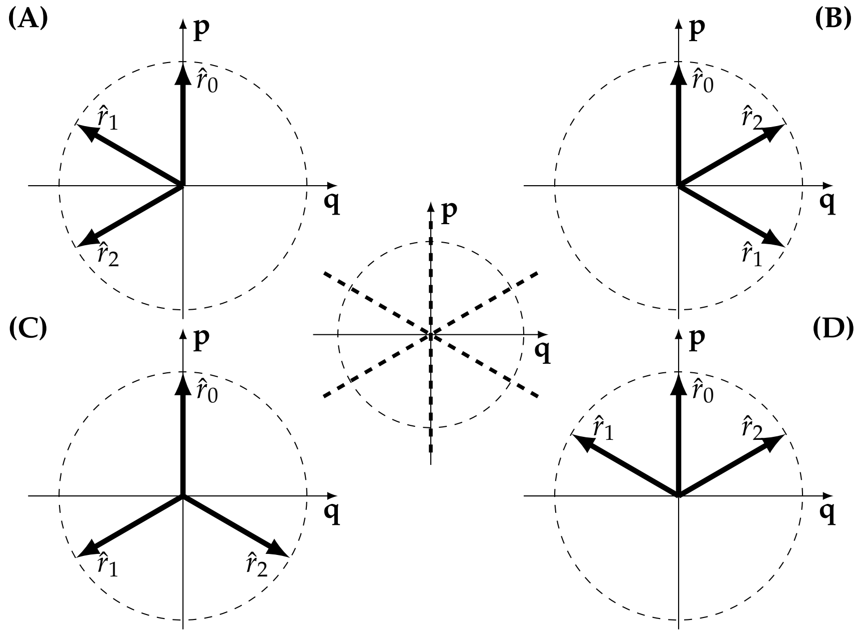

Figure 4.2: The four graphics on the corners depict the four triples in phase space (up to relabelling $\hat{r}_1, \hat{r}_2$ and arbitrary rotations of the whole configuration) that maximize the right hand side of the sum inequality (4.5), given that we arbitrarily select the first observable to be momentum. Circles are of radius ℓ to demonstrate that all observables are of equal length (the one in the centre is slightly scaled). The dashed lines of the graph in the centre are the directions where potential observables that are maximally incompatible reside.

of the state of the system. The question is how to decide which directions to measure in order to guarantee a maximal minimum bound on the information gained.

In the worst case scenario, we would perform measurements along the same direction in all three ensembles and would not obtain any new information from the latter two. At the other end, we would ideally measure the most incompatible directions so that the information gain is maximal. This incompatibility is directly connected to and expressed through the uncertainty relations and thus one would have to make measurements on the directions dictated by one of the triples of Fig. 4.2. Effectively, this way one would obtain the distributions associated with the dashed lines of the central graphic of Fig. 4.2.

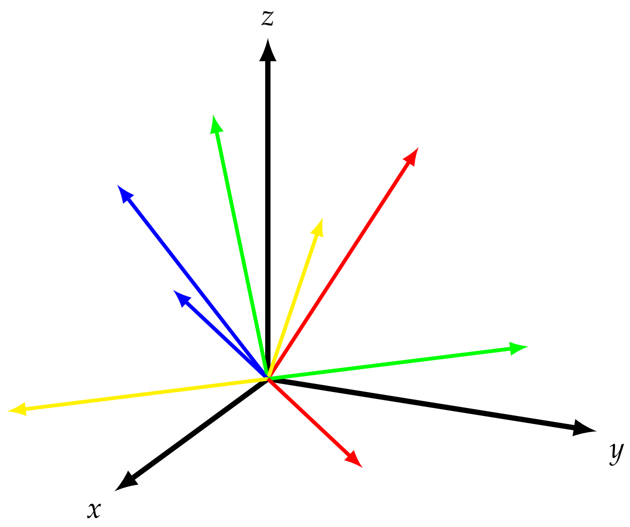

Figure 4.3: The structure of the $\mathbf{a}, \mathbf{b}$ vectors, embedded in $\mathbb{R}^3$, of the coefficients of position and momentum of all possible maximal triples; vectors of the same colour are orthogonal. Vectors on the $x - y$ plane correspond to $\mathbf{b}$ and the remaining to $\mathbf{a}$. The red pair corresponds to (A), while red to (B), green to (C) and yellow to (D) of Fig. 4.2. Reflection with respect to the $x - y$ plane gives all triples, had we selected $\hat{r}_0 = -\hat{p}$.

4.5 An inequality for the integral of variances of rotated observables

So far, we have derived inequalities for the sums and products of N discrete observables but one can use a result from the next chapter, to derive a more "exotic"-looking inequality.

In this section, we present a lower bound for the integral of the variances of rotated observables through an angle φ. More specifically, we will prove the inequality:

$$\int_0^{\varphi} \Delta^2 r_\theta \, d\theta \geq \frac{\hbar\sqrt{2\varphi^2 - 1 + \cos(2\varphi)}}{2\sqrt{2}} \,. \tag{4.38}$$

The proof of this statement follows in a straightforward way from the inequality for the linear combination of second moments, Eq. (5.74), that will be derived in Chapter 5.

Let us start by considering a rotated momentum operator $\hat{p}$ by an angle $\theta \in [0, 2\pi)$, that is measured from the vector associated with momentum. Explicitly, it is given by

$\hat{r}_\theta = \hat{p}\cos\theta - \hat{q}\sin\theta$, and by squaring and taking the average in an arbitrary quantum state, we obtain its variance

$$\Delta^2 r_\theta = \cos^2\theta\,\Delta^2 p + \sin^2\theta\,\Delta^2 q - \sin 2\theta\, C_{pq}\,. \tag{4.39}$$

Integrating from zero to an angle φ, we find

$$\int_0^\varphi \Delta^2 r_\theta\, d\theta = \mu\Delta^2 p + \nu\Delta^2 q + 2\lambda C_{pq}\,, \tag{4.40}$$

where the constants μ, ν and λ are defined through

$$\begin{aligned}
\mu &= \int_0^\varphi \cos^2\theta\, d\theta = \frac{1}{2}\left(\varphi + \cos\varphi\,\sin\varphi\right)\,,\\
\nu &= \int_0^\varphi \sin^2\theta\, d\theta = \frac{1}{2}\left(\varphi - \cos\varphi\,\sin\varphi\right)\,,\\
\lambda &= \int_0^\varphi \sin(2\theta)\, d\theta = -\frac{1}{2}\sin^2\varphi\,.
\end{aligned} \tag{4.41}$$

Substituting the values of μ, ν, λ into the linear uncertainty relation, Eq. (5.74), and simplifying we readily obtain inequality (4.38).

Although straightforward to prove, there are a few interesting observations associated with inequality (4.38) that we now list. For an angle $\varphi = 2\pi$ the bound reduces to $\pi\hbar$, agreeing with the area of the unit circle , which is also double the bound for $\varphi = \pi$. The case $\varphi = \pi$ could have been obtained by taking a pair of canonical observables and adding the contributions while rotating from zero to $\pi/2$, thus spanning half of the circle. However, for an angle $\varphi = \pi/2$, the intuitive expectation is not met and the lower bound turns out to be equal to $\hbar\sqrt{\pi^2 - 4}/4$, which is less than a quarter of the area of the unit circle, multiplied by $\hbar$. In general, for any angle $\varphi \in (0, \pi)$ the bound is less than the area of the corresponding arc. These results are shown in Fig. 4.4.

Only one state (up to phase space translations) achieves the lower bound of inequality (4.38), explicitly given in Eq. (5.78) of Chapter 5. They correspond to (possibly rotated) squeezed states, which are visualised in Fig. 4.5. There is a simple intuitive explanation of the minimising states, which we will now give. First of all, note that if we change the lower limit of integration from zero to φ_0, then the bound on the right hand side of (4.38) changes according to the substitution $\varphi \to \varphi - \varphi_0$. Moreover, if

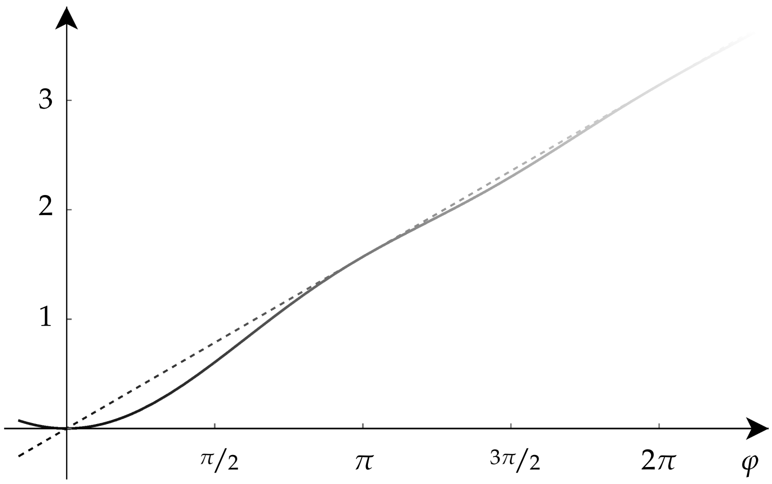

Figure 4.4: The solid line corresponds to the bound on the right hand side of inequality (4.38), divided by $\hbar$, while the dashed line gives the area of an arc of a unit circle subtending an angle φ. The two lines intersect only at the values 0, π, and 2π.

we symmetrically integrate through an angle bisected by the direction of momentum, that is, $\varphi_0 = -\varphi = \chi/2$, the inequality is changed to

$$\int_{-\frac{\chi}{2}}^{\frac{\chi}{2}} \Delta^2 r_\theta \, d\theta \geq \frac{\hbar}{2}\sqrt{\chi^2 - \sin^2\chi}\,, \tag{4.42}$$

where now $\chi \in [0, \pi]$. Due to symmetry, this inequality is saturated by a Gaussian of zero covariance and thus this state must be a squeezed state in which the variances of the position and momentum are equal to $\hbar/2\sigma$ and $\hbar\sigma/2$, say, with $\sigma > 0$. The variance of a rotated observable in such a squeezed state is equal to

$$\Delta^2 r_\theta = \frac{\hbar\sigma}{2}\cos^2\theta + \frac{\hbar}{2\sigma}\sin^2\theta\,, \tag{4.43}$$

which when integrated from $-\chi/2$ to $\chi/2$ gives

$$v(\sigma) = \frac{\hbar\sigma}{2}a_\chi + \frac{\hbar}{2\sigma}b_\chi\,, \tag{4.44}$$

with obvious definitions for a_χ and b_χ. The minimum is obtained for

$$\sigma_\chi = \sqrt{\frac{b_\chi}{a_\chi}}\,. \tag{4.45}$$

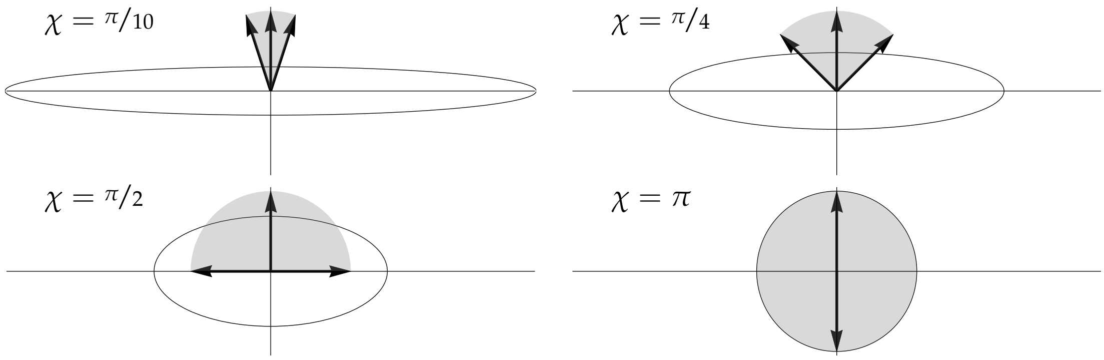

Figure 4.5: The ellipses corresponding to the unique squeezed states that saturate the integral inequality for various values of χ.

For an angle χ very close to zero, the Gaussian is "stretched" along the position axis in order to minimise the contribution from momentum around which we are integrating; at the limit this tends to a zero variance in momentum at the expense of an infinite variance in position. On the opposite end, as we integrate through angles approaching $\chi = \pi$, the Gaussian approaches a circle so that it balances the contributions from position and momentum (and everything in-between). In all cases, contour lines of the Wigner function corresponding to the considered squeezed state, define ellipses in phase space with semi-axes proportional to $\sqrt{\sigma_\chi}$ and $1/\sqrt{\sigma_\chi}$. The deformations of the state when the angle in the integral is changed, are performed in a way that preserve the area of the ellipses, as is required by the symplectic nature of the squeezing transformations. These results are shown in Fig. 4.5.

4.6 Three operators in more than one degree of freedom

In this section we consider observables in more than one degree of freedom and derive an inequality for a triple of operators. We first derive a general inequality for the variances of any three operators $\hat{A}, \hat{B}, \hat{C}$. Denoting $\hat{D} = \hat{A} - \langle \hat{A} \rangle$, $\hat{E} = \hat{B} - \langle \hat{B} \rangle$ and $\hat{F} = \hat{C} - \langle \hat{C} \rangle$ and starting from an inequality similar to ones used in [46, 69], i.e.

$$\left\| \left(\hat{D} + e^{i\varphi_1} \hat{E} + e^{i\varphi_2} \hat{F} \right) |\psi\rangle \right\|^2 \geq 0 \,, \tag{4.46}$$

we find that it is equivalent to

$$\Delta^2 A + \Delta^2 B + \Delta^2 C \geq -2C_{AB}\cos\varphi_1 - 2C_{CA}\cos\varphi_2 - 2C_{BC}\cos(\varphi_2-\varphi_1) \\ - i\sin\varphi_1\langle[\hat{A},\hat{B}]\rangle + i\sin\varphi_2\langle[\hat{C},\hat{A}]\rangle - i\sin(\varphi_2-\varphi_1)\langle[\hat{B},\hat{C}]\rangle\,, \tag{4.47}$$

where C_{AB} denotes the covariance between operators $\hat{A}$ and $\hat{B}$ and similarly for the rest. Setting $\cos\varphi_2 = \cos\varphi_2 = \cos(\varphi_2-\varphi_1)$, we find that the only solutions for $\varphi_1, \varphi_2 \in [0, 2\pi]$ are given by $\varphi_1 = 2\pi/3$ and $\varphi_2 = 4\pi/3$ (up to relabelling them), with $\cos(2\pi/3) = -1/2$. This last inequality becomes

$$\Delta^2 A + \Delta^2 B + \Delta^2 C \geq C_{AB} + C_{CA} + C_{BC} - i\frac{\sqrt{3}}{2}\left(\langle[\hat{A},\hat{B}]\rangle + \langle[\hat{B},\hat{C}]\rangle + \langle[\hat{C},\hat{A}]\rangle\right)\,, \tag{4.48}$$

Observing, in addition, that the variance of the sum of the three operators is equal to

$$\Delta^2(A+B+C) = \Delta^2 A + \Delta^2 B + \Delta^2 C + 2\left(C_{AB} + C_{BC} + C_{CA}\right)\,, \tag{4.49}$$

solving for the sum of the covariances and substituting in (4.48), we obtain the inequality

$$\Delta^2 A + \Delta^2 B + \Delta^2 C \geq \frac{1}{3}\Delta^2(A+B+C) - \frac{i}{\sqrt{3}}\left(\langle[\hat{A},\hat{B}]\rangle + \langle[\hat{B},\hat{C}]\rangle + \langle[\hat{C},\hat{A}]\rangle\right)\,. \tag{4.50}$$

Now, note the following: by relabelling $A \leftrightarrow C$ while keeping B the same, the second term flips an overall sign; moreover, substituting two of the operators by their negative, only one of the commutators in the second term flips a sign. Since all these inequalities hold, among all possibilities there is one where all three terms are positive. As a result, by dropping the first positive term of the variance of their sum (up to signs), we can express the inequality with the largest second term as

$$\Delta^2 A + \Delta^2 B + \Delta^2 C \geq \frac{1}{\sqrt{3}}\left(|\langle[\hat{A},\hat{B}]\rangle| + |\langle[\hat{B},\hat{C}]\rangle| + |\langle[\hat{C},\hat{A}]\rangle|\right)\,. \tag{4.51}$$

Consider now three observables in M degrees of freedom of the form

$$\hat{u}_i = \sum_{j=0}^{M-1} \left(a_{ij} \hat{P}_j + b_{ij} \hat{Q}_j \right) \;, \quad i = 0, 1, 2 \,, \tag{4.52}$$

where the operators $\hat{Q}_j, \hat{P}_j$ obey the commutation relations $[\hat{Q}_j, \hat{P}_k] = i\hbar\delta_{jk}$. In this case the last inequality becomes

$$\begin{aligned} \Delta^2 u_0 + \Delta^2 u_1 + \Delta^2 u_2 &\geq \frac{1}{\sqrt{3}} \left(\left| \langle \sum_{j,k} \left[a_{0j} \hat{P}_j + b_{0j} \hat{Q}_j, a_{1k} \hat{P}_k + b_{1k} \hat{Q}_k \right] \rangle \right| + \dots \right) \\ &= \frac{1}{\sqrt{3}} \left(\left| \sum_j \left(a_{0j} b_{1j} - a_{1j} b_{0j} \right) \right| + \dots \right) \,. \end{aligned} \tag{4.53}$$

If $\hat{u}_0, \hat{u}_1, \hat{u}_2$ form a triple with all pairwise commutators equal to $\pm i\hbar$, one can easily check that last inequality reduces to

$$\Delta^2 u_0 + \Delta^2 u_1 + \Delta^2 u_2 \geq \sqrt{3}\, \hbar \,, \tag{4.54}$$

which coincides with the one degree of freedom result. It follows from the fact that the commutators impose the conditions (e.g. for $\hat{u}_1, \hat{u}_2$)

$$[\hat{u}_1, \hat{u}_2] = i\hbar \quad \Rightarrow \quad \sum_j \left(a_{2j} b_{1j} - a_{1j} b_{2j} \right) = 1 \tag{4.55}$$

4.7 N observables with M degrees of freedom in a product state

In this section we will derive inequalities for the variances of N operators in M degrees of freedom where absence of correlations between different degrees of freedom is assumed.

Let the N operators, $\hat{u}_i$, be arbitrary linear combinations of all the possible positions and momenta of the M canonical pairs, $\hat{P}_i$ and $\hat{Q}_i$, one pair for each degree of freedom, or explicitly

$$\hat{u}_i = \sum_{j=0}^{M-1} \left(a_{ij} \hat{P}_j + b_{ij} \hat{Q}_j \right) \,. \tag{4.56}$$

In the absence of correlations between the degrees of freedom, their variances are given by

$$\Delta^2 u_i = \sum_{j=0}^{M-1} \left(a_{ij}^2 \Delta^2 P_j + b_{ij}^2 \Delta^2 Q_j + 2 a_{ij} b_{ij} C_{P_j Q_j} \right) . \tag{4.57}$$

The sum of their variances is then equal to

$$\begin{aligned} \sum_{i=0}^{N-1} \Delta^2 u_i &= \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \left(a_{ij}^2 \Delta^2 P_j + b_{ij}^2 \Delta^2 Q_j + 2 a_{ij} b_{ij} C_{P_j Q_j} \right) \\ &= \sum_{j=0}^{M-1} \left(A_j \Delta^2 P_j + B_j \Delta^2 Q_j + 2 \Gamma_j C_{P_j Q_j} \right) , \end{aligned} \tag{4.58}$$

where we defined

$$A_j = \sum_{i=0}^{N-1} a_{ij}^2 , \quad B_j = \sum_{i=0}^{N-1} b_{ij}^2 \quad \text{and} \quad \Gamma_j = \sum_{i=0}^{N-1} a_{ij} b_{ij} . \tag{4.59}$$

Introducing the *local* operators $\hat{r}_{ij} = a_{ij} \hat{P}_i + b_{ij} \hat{Q}_i$ for each j, and using once again the linear inequality of Chapter 2, Eq. (5.74), the bound for N observables in M degrees of freedom can be expressed in terms of the commutators of the $\hat{r}_{ij}$ as

$$\sum_{i=0}^{N-1} \Delta^2 u_i \geq \sum_{k=0}^{M-1} \sqrt{\sum_{i>j}^{N-1} \left| \langle [\hat{r}_{ik}, \hat{r}_{jk}] \rangle \right|^2 } . \tag{4.60}$$

Let us now, compare this bound for the case of the triple of last section with $[\hat{u}_i, \hat{u}_{i+1}] = \pm i\hbar$, with the indices being cyclic modulo 3 and $i = 0, 1, 2$. With $N = 3$, inequality (4.60) becomes

$$\begin{aligned} \sum_{i=0}^{2} \Delta^2 u_i &\geq \sum_{k=0}^{M-1} \sqrt{\sum_{i>j}^{2} \left| \langle [\hat{r}_{ik}, \hat{r}_{jk}] \rangle \right|^2 } \geq \frac{1}{\sqrt{3}} \sum_{k=0}^{M-1} \sum_{i>j}^{2} \left| \langle [\hat{r}_{ik}, \hat{r}_{jk}] \rangle \right| \\ &\geq \sum_{k=0}^{M-1} \left| \sum_{i>j}^{2} \langle [\hat{r}_{ik}, \hat{r}_{jk}] \rangle \right| = \sqrt{3}\, \hbar , \end{aligned} \tag{4.61}$$

where we have used the inequality

$$\sqrt{x_1^2 + \ldots + x_n^2} \geq \left(|x_1| + \ldots + |x_n| \right) / \sqrt{n} , \tag{4.62}$$

and the triangle inequality; equality in the former holds only if all $|x_i|$ are equal while in the latter if all terms are non negative. Inequality (4.62) follows from the Cauchy-Schwarz inequality written in the form,

$$\sum_{i=1}^{n} y_i^2 \sum_{i=1}^{n} z_i^2 \geq \left(\sum_{i=1}^{n} y_i z_i \right)^2 , \tag{4.63}$$

by letting $y_i = 1/\sqrt{n}$ and $z_i = |x_i|/\sqrt{n}$.

Although the lowest possible bound is no less than the one of last section, it is in general higher because of the use of the triangle inequality. As a result, this difference in the lower bounds can in certain cases be exploited for entanglement detection, as in e.g. [30].

A simple example is given by the following triple

$$\hat{u}_0 = \hat{P}_1 + \hat{Q}_2 , \quad \hat{u}_1 = \hat{P}_2 + \hat{Q}_3 , \quad \hat{u}_2 = \hat{P}_3 + \hat{Q}_1 , \tag{4.64}$$

for which inequality (4.60) reduces to

$$\Delta^2 u_0 + \Delta^2 u_1 + \Delta^2 u_2 \geq 3\,\hbar , \tag{4.65}$$

while from last section the lowest possible bound is $\sqrt{3}\,\hbar$, which can be achieved only by entangled states. Thus violation of inequality (4.60) in a given state is an indication of entanglement. The difference between the bounds can be further widened by selecting, for example, a canonical triple for each degree of freedom and then considering the triple

$$\hat{u}_0 = \hat{P}_1 - \hat{Q}_2 + \hat{R}_3 , \quad \hat{u}_1 = \hat{P}_2 - \hat{Q}_3 + \hat{R}_1 , \quad \hat{u}_2 = \hat{P}_3 - \hat{Q}_1 + \hat{R}_2 , \tag{4.66}$$

for which, inequality (4.60) now gives

$$\Delta^2 u_0 + \Delta^2 u_1 + \Delta^2 u_2 \geq 3\sqrt{3}\,\hbar , \tag{4.67}$$

while the general inequality (4.54) gives again the bound $\sqrt{3}\,\hbar$.

Although in these examples the inequalities for product states can be violated by e.g. states of the form $|\Psi\rangle = |\psi_{12}\rangle \otimes |\psi_3\rangle$, it would be interesting to ask whether one can produce inequalities that detect genuine tripartite entanglement. It turns out that this question can be answered in the affirmative.

Consider, for example, a system with three degrees of freedom and the three operators

$$\hat{u}_0 = \hat{Q}_1 + \hat{P}_2 + \hat{R}_3\,, \quad \hat{u}_1 = \hat{Q}_2 - \hat{P}_3 + \hat{R}_3\,, \quad \hat{u}_2 = -\hat{Q}_3 - \hat{P}_1 + \hat{R}_2\,, \tag{4.68}$$

Three possibilities arise: (i) the state is completely separable, $|\psi_1\rangle \otimes |\psi_2\rangle \otimes |\psi_3\rangle$, (ii) it is *bi-separable*, i.e. of the form $|\psi_1\rangle \otimes |\psi_{23}\rangle$ (and permutation of the indices), or (iii) none of the above. Then, using the inequality (4.51) repeatedly, it can be shown that the lower bounds for the sum of the variances of the $\hat{u}_i$ in each of the above cases are different. The results are shown in Table 4.1.

Table 4.1: The different lower bounds

Type of state $\|\Psi\rangle$	Lower bound of $\sum_i \Delta^2 u_i$
$\|\psi_1\rangle \otimes \|\psi_2\rangle \otimes \|\psi_3\rangle$	$\frac{9\hbar}{\sqrt{3}}$
$\|\psi_1\rangle \otimes \|\psi_{23}\rangle$ $\|\psi_2\rangle \otimes \|\psi_{31}\rangle$ $\|\psi_3\rangle \otimes \|\psi_{12}\rangle$	$\frac{5\hbar}{\sqrt{3}}$
$\|\psi_{123}\rangle$	$\frac{3\hbar}{\sqrt{3}}$

Whenever the sum of the variances of the global operators $\hat{u}_i$ attains a value less than $\frac{5\hbar}{\sqrt{3}}$, this signals the existence of genuine tripartite entanglement, while if it is in the range $[\frac{5\hbar}{\sqrt{3}}, \frac{9\hbar}{\sqrt{3}}]$, it signals the presence of either bipartite or tripartite entanglement. However, the criterion becomes inconclusive if the sum of the three variances is larger than $\frac{9\hbar}{\sqrt{3}}$.

The criterion for detecting tripartite entanglement presented in this section was based on a particular choice of operators but a systematic study of other choices that lead to similar criteria is possible, following from the requirement that the three different types of states lead to different lower bounds. Finally, a similar analysis can be applied to the case of m degrees of freedom in order to detect genuine m-partite entanglement.

4.8 Summary

In this chapter we derived a number of inequalities for observables that are linear combinations of position and momentum. Whenever their coefficients form a regular polygon in $\mathbb{R}^2$, the lower bounds of such inequalities can be re-expressed geometrically, in terms of the associated polygons. In every case, the lower bound can be alternatively seen as the area of the parallelogram spanned by the two vectors in $\mathbb{R}^N$, formed by the coefficients of the positions and momenta of all observables. An inequality for the integral of the variances of rotated observables was given, and the lower bound was found to agree with the area of the half or full circle for $\varphi = \pi$ and $\varphi = 2\pi$, respectively; for all other values the bound is lower than the one predicted by the area of the arc spanned, multiplied by $\hbar$.

In the penultimate section, we derived an inequality for the sum of variances of three arbitrary operators. Whenever each pairwise commutator is equal to $\pm i\hbar$, the inequality is tight but this is not the case in general. It is, however, in many cases tighter than the one obtained by combining the pairwise inequalities for the sum: $\Delta^2 A + \Delta^2 B \geq |\langle [A, B] \rangle|$.

In the last section we obtained inequalities for the variances of observables in more than one spatial degree of freedom but valid for non-entangled states only. By restricting to a triple and using the inequality of the penultimate section, we described how this could be used to produce entanglement detection criteria.

Chapter 5

Uncertainty relations for a single particle in one dimension

5.1 Introduction and main result

Inspired by Heisenberg's analysis [35] of Compton scattering, Kennard [41] proved the preparational uncertainty relation

$$\Delta p \, \Delta q \geq \frac{\hbar}{2}, \tag{5.1}$$

for the standard deviations Δp and Δq of momentum and position of a quantum particle with a single spatial degree of freedom. Experimentally, they are determined by measurements performed on an ensemble of systems prepared in a specific state $|\psi\rangle$. The states saturating the bound (5.1) are *squeezed states* with a *real* squeezing parameter [54, 49, 71] (we follow the review [28] regarding the naming of squeezed states). Squeezed states are conceptually important since they achieve the best possible localization of a quantum particle in a phase-space, and they are easily visualized by "uncertainty ellipses". Each squeezed state may be displaced rigidly in phase space without affecting the value of the variances, resulting in a three-parameter family of states saturating the lower bound (5.1).

Other uncertainty relations are known. The *sum* of the position and momentum variances (throughout this paper, we use a system of units where the physical dimensions of both position and momentum equal $\sqrt{\hbar}$) is bounded [73, 19] according to the relation

$$\Delta^2 p + \Delta^2 q \geq \hbar\,; \tag{5.2}$$

only the ground state of a suitable harmonic oscillator saturates the inequality if rigid displacements in phase-space are ignored. The Robertson-Schrödinger (RS) inequality [56, 62],

$$\Delta^2 p\, \Delta^2 q - C_{pq}^2 \geq \frac{\hbar^2}{4}\,, \tag{5.3}$$

sharpens Heisenberg's inequality (5.1) by including the covariance C_{pq} defined in Eq. (5.8). Eq. (5.3) is saturated by the two-parameter family of *squeezed states*, with a *complex* squeezing parameter [28], again ignoring phase-space displacements. The second parameter describes the phase-space orientation of the uncertainty ellipse which, in the previous case, was aligned with the position and momentum axes.

By introducing the observable $\hat{r} = -\hat{p} - \hat{q}$, which satisfies the commutation relations $[\hat{q}, \hat{r}] = [\hat{r}, \hat{p}] = \hbar/i$, one obtains a bound on the product of the variances of three pairwise canonical observables,

$$\Delta^2 p\, \Delta^2 q\, \Delta^2 r \geq \left(\tau \frac{\hbar}{2}\right)^3, \qquad \tau = \csc\left(\frac{2\pi}{3}\right) \equiv \sqrt{\frac{4}{3}}\,. \tag{5.4}$$

This *triple product uncertainty relation* found only recently [39] was the focus of Chapter 3. Since the variance of $\hat{r}$ is given by

$$\Delta^2 r = \Delta^2 p + \Delta^2 q + 2C_{pq}\,, \tag{5.5}$$

the left-hand-side of (5.4) can also be considered as a function of the three second moments. Ignoring phase-space translations, only one state exists which achieves the minimum.

Inequalities (5.1) to (5.4) and the search for their minima arise from one single mathematical problem:

> *Does a given smooth function of the second moments have a lower bound?*
> *If so, which states will saturate the inequality if a minimum exists?*

In this chapter, we answer these questions for a quantum particle with a single spatial degree of freedom by presenting a systematic approach to studying uncertainty relations derived from smooth functions $f(\Delta^2 p, \Delta^2 q, C_{pq})$. Proceeding in three steps we

1. identify a *universal* set of states $\mathscr{E}$ which can possibly minimize a given functional $f(\Delta^2 p, \Delta^2 q, C_{pq})$;
2. spell out conditions which determine the *extrema* of the functional f as a subset of the universal set, $\mathscr{E}(f) \subseteq \mathscr{E}$; if no admissible extrema exist, the functional has no lower bound;
3. determine the set of states $\mathscr{M}(f) \subseteq \mathscr{E}(f)$ which *minimize* the functional f, leading to an uncertainty relation in terms of the second moments.

If the considered functions f satisfy a number of additional properties, then they can be considered as a measure of the overall uncertainty associated with a number of observables and a pure state.

Noting that the covariance can be re-expressed as a linear combination of the variances of position and momentum and a third operator (e.g. $\hat{r} = -\hat{p} - \hat{q}$ of Chapter 3), then a function of the variances of a given number of observables is a good measure of the overall uncertainty if it has the following properties:

1. $f(x, y, \dots) \geq 0$ for all $x,\ y, \dots \geq 0$.
2. f increases (decreases) when one of its argument increases (decreases), while the rest are kept the same.
3. $f(x, y \dots) = 0$, if and only if x=y=…=0.

These three properties are physically justified for a measure of an overall uncertainty: it is a non-negative function that is zero only if all observables can attain arbitrarily precise values, and it increases or decreases if the uncertainty of one observable increases or decreases. Inequalities for such functions can be viewed as approximate descriptions of the *uncertainty region* of n observables (in the case considered here 3), i.e. the collection of all n-tuples of variances that arise from physical states. As we will show, for the case of position and momentum that we consider here, the Robertson-Schrödinger inequality, along with the positivity of the variances completely determines the uncertainty region for a quantum particle. All other inequalities give, in general, only approximations in the sense that the region they define may contain points that do not arise from physical states.

The inequalities will be *preparational* in spirit, i.e. they apply to scenarios in which the quantum state of the particle $|\psi\rangle$ is fixed during the three separate runs of the measurements required to determine the numerical values of the second moments. These inequalities do not describe the limitations of measuring non-commuting observables *simultaneously*.

The chapter is divided into two major sections. In Sec. 5.2 we introduce uncertainty functionals and explain how to determine their extrema and minima; to illustrate our method we re-derive a number of known inequalities. In Sec. 5.3 we derive new (families of) uncertainty relations and the states minimizing them. We conclude the chapter with a summary and discuss further applications.

5.2 Minimising uncertainty

Key to our approach is the *uncertainty functional*

$$J[\psi] = f\left(\Delta^2 p, \Delta^2 q, C_{pq}\right) - \lambda\left(\langle\psi|\psi\rangle - 1\right), \tag{5.6}$$

which sends each normalizable element $|\psi\rangle$ of the one-particle Hilbert space $\mathscr{H}$ to a real number determined by the real function $f(x_1, x_2, x_3)$ of three variables. A lower bound on a functional $J[\psi]$ of the form (5.6) will result in an uncertainty relation associated with the function $f(x_1, x_2, x_3)$.

The variances of position and momentum are defined by

$$\Delta^2 p = \langle\psi|\hat{p}^2|\psi\rangle - \langle\psi|\hat{p}|\psi\rangle^2, \tag{5.7}$$

etc., and the correlation between position and momentum reads

$$C_{pq} = \frac{1}{2}\langle\psi|\left(\hat{p}\hat{q}+\hat{q}\hat{p}\right)|\psi\rangle - \langle\psi|\hat{p}|\psi\rangle\langle\psi|\hat{q}|\psi\rangle\,, \tag{5.8}$$

which we will often arrange in to the real, symmetric *covariance* matrix, defined in Chapter 2. For notational simplicity we will denote the second moments as

$$x = \Delta^2 p\,, \quad y = \Delta^2 q\,, \quad w = C_{pq}\,, \tag{5.9}$$

but dependence of the matrix elements on the state $|\psi\rangle$ i.e. $x \equiv x(\psi)$ etc., should be kept in mind. The variances of position and momentum are strictly positive, $x, y > 0$, since a quantum particle has no normalizable position or momentum eigenstates while the correlation may take any finite real value, $w \in \mathbb{R}$.

To determine the extrema of the functional $J[\psi]$ with our method, its first-order Gâteaux derivative must exist which requires the function $f(x_1, x_2, x_3)$ to be differentiable (e.g. see example of Sec. 2.5). The Lagrange multiplier λ in (5.6) ensures that only normalized states are taken into account. It will be convenient to work with states in which the expectation values of both momentum and position vanish, $\langle\psi|\hat{q}|\psi\rangle = \langle\psi|\hat{p}|\psi\rangle = 0$. This can be achieved by rigidly displacing the observables using the unitary operator

$$\hat{T}_\alpha = \exp\left[i\left(p_0\hat{q} - q_0\hat{p}\right)/\hbar\right]\,, \qquad \alpha = \frac{1}{\sqrt{2\hbar}}\left(q_0 + ip_0\right)\,, \tag{5.10}$$

where $p_0 = \langle\psi|\hat{p}|\psi\rangle$, etc. This transformation leaves invariant the values of the second moments (5.7) and (5.8) and has thus no impact on the minimization of the functional $J[\psi]$.

To establish an uncertainty relation stemming from a given function $f(x, y, w)$ we follow the three-step procedure mention earlier, inspired by methods presented in [38, 78, 39] (see also [21]). We first apply a variational technique to derive an eigenvalue equation the solutions of which are extrema of the functional $J[\psi]$, and we derive a set of consistency conditions in Sec. 5.2.3; we then introduce a “space of moments” to visualize these results (Sec. 5.2.4). Finally, we determine the minimizing states whenever the functional is guaranteed to be bounded from below (see Sec. 5.2.6 and Sec. 5.3).

5.2.1 Extrema of uncertainty functionals

When comparing the values of the functional J at the points $|\psi\rangle$ and $|\psi+\varepsilon\rangle \equiv |\psi\rangle + \varepsilon|e\rangle$, for any unit vector $|e\rangle$ and a small parameter ε, we find to first order that

$$J[\psi+\varepsilon] - J[\psi] = \varepsilon D_\varepsilon J[\psi] + O\left(\varepsilon^2\right), \tag{5.11}$$

where the expression

$$D_\varepsilon = \langle e|\frac{\delta}{\delta\langle\psi|} + \frac{\delta}{\delta|\psi\rangle}|e\rangle, \tag{5.12}$$

denotes a Gâteaux derivative defined in Chapter 2. If the functional $J[\psi]$ does not change under this variation,

$$D_\varepsilon J[\psi] = \langle e|\left(\frac{\delta}{\delta\langle\psi|}f\left(x,y,w\right) - \lambda|\psi\rangle\right) + \text{c.c.} = 0, \tag{5.13}$$

it has an extremum at the state $|\psi\rangle$. More explicitly, this condition reads

$$\langle e|\left(\frac{\partial f}{\partial x}\frac{\delta x}{\delta\langle\psi|} + \frac{\partial f}{\partial y}\frac{\delta y}{\delta\langle\psi|} + \frac{\partial f}{\partial w}\frac{\delta w}{\delta\langle\psi|} - \lambda|\psi\rangle\right) + \text{c.c.} = 0\,, \tag{5.14}$$

which should hold for arbitrary variations. Since the vector $|e\rangle$ and its dual, $\langle e|$, are linearly independent (just consider their position representations $e^*(x)$ and $e(x)$), the expression in round brackets must vanish identically which implies that the complex conjugate term will also vanish. Using

$$\frac{\delta x}{\delta\langle\psi|} \equiv \frac{\delta\Delta^2 p}{\delta\langle\psi|} = \frac{\delta\langle\psi|\hat{p}^2|\psi\rangle}{\delta\langle\psi|} = \hat{p}^2|\psi\rangle\,, \tag{5.15}$$

a similar relation for $\delta y/\delta\langle\psi|$, and the identity

$$\frac{\delta w}{\delta\langle\psi|} \equiv \frac{1}{2}\left(\hat{q}\hat{p} + \hat{p}\hat{q}\right)|\psi\rangle\,, \tag{5.16}$$

we arrive at an *Euler-Lagrange*-type equation,

$$\left(\frac{\partial f}{\partial x}\hat{p}^2 + \frac{\partial f}{\partial y}\hat{q}^2 + \frac{1}{2}\frac{\partial f}{\partial w}\left(\hat{q}\hat{p} + \hat{p}\hat{q}\right) - \lambda\right)|\psi\rangle = 0. \tag{5.17}$$

The parameter λ can be eliminated by multiplying this equation with the bra $\langle\psi|$ from the left and solving for λ; substituting the value obtained back into Eq. (5.17), one finds a nonlinear eigenvector-eigenvalue equation,

$$\left(f_x\hat{p}^2 + f_y\hat{q}^2 + \frac{f_w}{2}\left(\hat{q}\hat{p} + \hat{p}\hat{q}\right)\right)|\psi\rangle = \left(f_x x + f_y y + f_w w\right)|\psi\rangle\,, \tag{5.18}$$

using the standard shorthand for partial derivatives.

Eq. (5.18) is our first result: the extrema of a smooth function of the second moments are encoded in an eigenvalue equation for a Hermitean operator quadratic in position and momentum. However, the equation is not linear in the state $|\psi\rangle$ because the quantities $x, y, \ldots, f_w$ are functions of expectation values taken in the yet unknown state.

Let us briefly illustrate the crucial features of Eq. (5.18) in a simple case before systematically investigating its solutions. For a function linear in x, y, and w, the derivatives f_x, f_y, and f_w will be constant numbers. In this case, the operator on the left-hand-side of (5.18) turns into an explicitly given quadratic form in the position and momentum operators, falling into one of three possible categories [79]: up to a multiplicative constant, the operator will be unitarily equivalent to the Hamiltonian of (i) a harmonic oscillator with unit mass and frequency, $\hat{p}^2 + \hat{q}^2$, (ii) a free particle, $\hat{p}^2$, or (iii) an inverted harmonic oscillator, $\hat{p}^2 - \hat{q}^2$. In the first case, the spectrum of the operator will be discrete and bounded from below (or above); the spectra of the operators in the other two cases are continuous which is tantamount to the absence of normalizable eigenstates. Thus, a linear function $f(x, y, w)$ possesses a non-trivial bound only if it gives rise to an operator in (5.18) which is unitarily equivalent to the Hamiltonian of a harmonic oscillator. Our method will signal the absence of lower bounds corresponding to the cases (ii) and (iii).

5.2.2 Universality

To find a lower bound of the functional $J[\psi]$, we will determine all its extrema and then pick those where $J[\psi]$ assumes its smallest value. However, Eq. (5.18) is not a standard eigenvalue equation: even for a linear function f, the right-hand-side of (5.18) depends non-linearly on the as yet unknown state $|\psi\rangle$, and if the function f is non-linear, the operators on the left-hand-side of the equation acquire state-dependent coefficients given by its partial derivatives.

Nevertheless, the eigenvalue problem can be solved systematically, in a *self-consistent* way. Initially, we treat the expectations x, y, w, and f_x, f_y, f_w in (5.18) as parameters with given values, i.e. independent of $|\psi\rangle$. The solutions $|\psi(x, y, w)\rangle$ will depend on these parameters which means that the solutions must be checked for consistency since the relations (5.7) now require that $x = \langle\psi(x, y, w)|\hat{p}^2|\psi(x, y, w)\rangle$, for example. Typically, additional restrictions will arise from these constraints; entirely inconsistent cases exist, too.

Let us begin by writing the operator on the left-hand side of (5.18) in matrix form,

$$(\hat{p}\,,\hat{q})\begin{pmatrix} f_x & f_w/2 \\ f_w/2 & f_y \end{pmatrix}\begin{pmatrix} \hat{p} \\ \hat{q} \end{pmatrix} \equiv \hat{\mathbf{z}}^\top \cdot \mathbf{F}\,\cdot\,\hat{\mathbf{z}}\,. \tag{5.19}$$

Williamson's theorem [84] ensures that any positive or negative definite matrix can be mapped to a diagonal matrix by conjugation with a symplectic matrix $\mathbf{\Sigma}$. For simplicity, we will assume from now on that $\mathbf{F}$ is *positive* definite. The negative definite case is easily dealt with by "flipping over" the functional, i.e. by considering $-f(x, y, w)$ instead of $f(x, y, w)$. Applied to the 2×2 matrix $\mathbf{F}$, Williamson's result takes the simple form

$$\mathbf{\Sigma}^\top \cdot \mathbf{F} \cdot \mathbf{\Sigma} = \mathrm{c}\,\mathbf{I}\,, \qquad \mathbf{\Sigma} \in \mathrm{Sp}(2, \mathbb{R})\,, \quad c > 0\,, \tag{5.20}$$

where $\mathbf{I}$ is the identity matrix, whenever $\mathbf{F} > 0$ holds. This requires both

$$\det \mathbf{F} \equiv f_x f_y - f_w^2/4 > 0 \quad \text{and} \quad f_x > 0\,, \tag{5.21}$$

implying that $f_y > 0$ will hold, too. These requirements clearly agree with the observations made for linear uncertainty functionals $f(x, y, w)$: since the operators $\hat{p}^2$ and

$\hat{p}^2 - \hat{q}^2$ result in matrices $\mathbf{F}$ with zero or negative determinant, the left-hand-side of (5.18) cannot be mapped to an oscillator Hamiltonian by means of a symplectic transformation.

A direct calculation shows that the matrix $\mathbf{F}$ is diagonalized by the symplectic matrix $\Sigma = (\mathbf{S}_\gamma \mathbf{G}_b)^{-1}$, where

$$\mathbf{G}_b = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}, \quad \text{and} \quad \mathbf{S}_\gamma = \begin{pmatrix} e^{-\gamma} & 0 \\ 0 & e^{\gamma} \end{pmatrix}, \tag{5.22}$$

with real parameters

$$b = \frac{f_w}{2f_y} \in \mathbb{R} \quad \text{and} \quad \gamma = \frac{1}{2} \ln\left(\frac{f_y}{\sqrt{\det \mathbf{F}}}\right) \in \mathbb{R}, \tag{5.23}$$

leading to $c = \sqrt{\det \mathbf{F}}$ in Eq. (5.20). The symplectic matrices $\mathbf{S}_\gamma$ and $\mathbf{G}_b$ give rise to the Iwasawa (or $\mathscr{K}\mathscr{A}\mathscr{N}$) decomposition of the matrix $\mathbf{\Sigma}^{-1} \in \mathrm{Sp}(2, \mathbb{R})$ (cf. [8], for example) if they are written in opposite order and the parameter b is replaced by be^{γ}; the third factor happens to be the identity.

Next, we observe that the linear action of the matrices $\mathbf{G}$ and $\mathbf{S}$ on the canonical pair of operators $(\hat{p}, \hat{q})^\top$ can be implemented by conjugation with suitable unitary operators, known as metaplectic operators [8]. We have, for example,

$$\begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix} \begin{pmatrix} \hat{p} \\ \hat{q} \end{pmatrix} = e^{ib\hat{p}^2/2\hbar} \begin{pmatrix} \hat{p} \\ \hat{q} \end{pmatrix} e^{-ib\hat{p}^2/2\hbar}, \tag{5.24}$$

or, in matrix notation,

$$\mathbf{G}_b \cdot \hat{\mathbf{z}} = \hat{G}_b \, \hat{\mathbf{z}} \, \hat{G}_b^\dagger \tag{5.25}$$

where the unitary operator

$$\hat{G}_b = e^{ib\hat{p}^2/2\hbar} \tag{5.26}$$

describes a *momentum gauge transformation*. Similarly, the squeeze operator

$$\hat{S}_\gamma = e^{i\gamma(\hat{q}\hat{p}+\hat{p}\hat{q})/2\hbar}, \tag{5.27}$$

symplectically *scales* position and momentum according to

$$\mathbf{S}_\gamma \cdot \hat{\mathbf{z}} = \hat{S}_\gamma \, \hat{\mathbf{z}} \, \hat{S}_\gamma^\dagger \,. \tag{5.28}$$

With $\mathbf{\Sigma}^{-1} = \mathbf{S}_\gamma \mathbf{G}_b$ in (5.20), we rewrite (5.19) as

$$\hat{\mathbf{z}}^\top \cdot \mathbf{F} \, \cdot \, \hat{\mathbf{z}} = \sqrt{\det \mathbf{F}} \, (\mathbf{S}_\gamma \cdot \mathbf{G}_b \cdot \hat{\mathbf{z}})^\top \cdot (\mathbf{S}_\gamma \cdot \mathbf{G}_b \cdot \hat{\mathbf{z}}) \,. \tag{5.29}$$

Finally, using the identities (5.25) and (5.28) and multiplying Eq. (5.18) with the unitary $\hat{S}_\gamma^\dagger \hat{G}_b^\dagger$ from the left, the condition for the existence of extrema of the functional $J[\psi]$ takes a simple form,

$$\frac{1}{2} \left(\hat{p}^2 + \hat{q}^2 \right) |\psi(b,\gamma)\rangle = \left(\frac{x f_x + y f_y + w f_w}{2\sqrt{\det \mathbf{F}}} \right) |\psi(b,\gamma)\rangle \, ; \tag{5.30}$$

thus, the solutions

$$|\psi(b,\gamma)\rangle \equiv \hat{S}_\gamma^\dagger \hat{G}_b^\dagger |\psi\rangle \tag{5.31}$$

are proportional to the eigenstates $|n\rangle, n \in \mathbb{N}_0$, of a *unit oscillator*, i.e. a quantum mechanical oscillator with unit mass and unit frequency. Equivalently, the candidates for states extremising the functional $J[\psi]$ are given by the family of states,

$$|n(b,\gamma)\rangle = \hat{G}_b \hat{S}_\gamma |n\rangle \,, \qquad b,\gamma \in \mathbb{R} \,, \quad n \in \mathbb{N}_0 \,. \tag{5.32}$$

Upon rewriting the operator on the right-hand-side these states are seen to coincide with the *squeezed number states* [28]. As shown in in Appendix A.1, the product of a squeeze transformation $\hat{S}_\gamma$ (with real parameter γ) and a momentum gauge transformation $\hat{G}_b$ equals

$$\hat{G}_b \hat{S}_\gamma = \hat{S}(\xi) \hat{R}(\chi) \,, \tag{5.33}$$

i.e. the product of a rotation in phase space,

$$\hat{R}(\chi) = e^{i\chi \hat{a}^\dagger \hat{a}} \qquad \chi \in [0, 2\pi) \,, \tag{5.34}$$

and a squeeze transformation (with complex ξ) along a line with inclination θ,

$$\hat{S}(\xi) = e^{\frac{1}{2}\left(\xi\hat{a}^{\dagger 2} - \bar{\xi}\hat{a}^2\right)}, \qquad \xi = re^{i\theta} \in \mathbb{C}\,. \tag{5.35}$$

Summarizing our findings, we draw two conclusions:

1. The complete set of solutions of Eq. (5.30) coincides with the *squeezed number states*,

$$\mathscr{E} = \bigcup_{n=0}^{\infty} \mathscr{E}_n \equiv \bigcup_{n=0}^{\infty} \left\{|n(\alpha,\xi)\rangle = \hat{T}_\alpha \hat{S}(\xi)|n\rangle\,, \alpha, \xi \in \mathbb{C}\right\}\,, \tag{5.36}$$

   where non-zero expectation values of position and momentum have been reintroduced via the translation operator $\hat{T}_\alpha$ (see Eq. (5.10)) and an irrelevant constant phase has been suppressed.

2. The value of the right-hand-side of Eq. (5.30) can take only specific values,

$$\frac{xf_x + yf_y + wf_w}{2\sqrt{\det \mathbf{F}}} = \left(n + \frac{1}{2}\right)\hbar\,, \qquad n \in \mathbb{N}_0\,, \tag{5.37}$$

   given by the eigenvalues of the unit oscillator. This relation constrains the state-dependent quantities of the left-hand-side which needs to be checked for consistency, just as Eq. (5.21) does.

We have thus obtained our second main result. The extrema $\mathscr{E}$ of an arbitrary functional $J[\psi]$ characterized by a function $f(x,y,w)$ are necessarily *squeezed number states*, a set which is *independent* of the function at hand. In other words, the set $\mathscr{E}$ containing all the states which may arise as minima of an uncertainty functional $J[\psi]$ is *universal*. The minima of any functional must be a subset $\mathscr{E}(f) \subseteq \mathscr{E}$ which will depend explicitly on the function $f(x,y,w)$, determined by the consistency conditions to be studied next.

5.2.3 Consistency conditions

We now spell out the conditions which must be satisfied by the states $|n(b,\gamma)\rangle$ in (5.32) – or, equivalently, the states $|n(\alpha,\xi)\rangle$ in (5.36) – to qualify as extrema for a specific functional $J[\psi]$:

1. Recalling that $x \equiv \Delta^2 p$, etc., the relations

$$x = \langle n(b,\gamma)|\hat{p}^2|n(b,\gamma)\rangle\,, \qquad y = \langle n(b,\gamma)|\hat{q}^2|n(b,\gamma)\rangle\,, \tag{5.38}$$

   and

$$w = \frac{1}{2}\langle n(b,\gamma)|\,(\hat{p}\hat{q}+\hat{q}\hat{p})\,|n(b,\gamma)\rangle\,, \tag{5.39}$$

   represent three, generally nonlinear *consistency equations* between the second moments since the parameters b and γ are functions of x, y and w (cf. Eq. (5.23)).

2. The values of the moments x, y and w must satisfy Eq. (5.37).

3. The matrix $\mathbf{F}$ of the first derivatives must be positive definite.

Using (5.32), (5.25) and (5.28), the first consistency condition in (5.38) leads to

$$x = \langle n(b,\gamma)|\hat{p}^2|n(b,\gamma)\rangle = e^{2\gamma}\langle n|\hat{p}^2|n\rangle = e^{2\gamma}\left(n+\frac{1}{2}\right)\hbar\,, \quad n \in \mathbb{N}_0\,, \tag{5.40}$$

or, recalling the definition of γ in (5.23),

$$x\sqrt{\det\mathbf{F}} = \left(n+\frac{1}{2}\right)\hbar f_y\,, \qquad n \in \mathbb{N}_0\,. \tag{5.41}$$

Similar calculations result in

$$y\sqrt{\det\mathbf{F}} = \left(n+\frac{1}{2}\right)\hbar f_x\,, \qquad n \in \mathbb{N}_0\,, \tag{5.42}$$

and

$$-2w\sqrt{\det\mathbf{F}} = \left(n+\frac{1}{2}\right)\hbar f_w\,, \qquad n \in \mathbb{N}_0\,, \tag{5.43}$$

respectively. These conditions may be expressed in matrix form,

$$\frac{\mathbf{F}\cdot\mathbf{C}}{\sqrt{\det\mathbf{F}}} = \left(n+\frac{1}{2}\right)\hbar\,\mathbf{I}, \qquad n \in \mathbb{N}_0\,, \tag{5.44}$$

involving both the covariance matrix $\mathbf{C}$ and $\mathbf{F}$.

Taking the trace of the last relation shows that Eq. (5.37) is satisfied automatically. Without specifying a function $f(x,y,w)$, no conclusions can be drawn about the validity of Eqs. (5.41-5.43) or the positive definiteness of the matrix $\mathbf{F}$.

5.2.4 Geometry of extremal states

In order to visualize the interplay of the consistency conditions we express them as

$$x f_x = y f_y \,, \qquad x f_w = -2w f_y \,, \tag{5.45}$$

and

$$xy - w^2 = \left(n + \frac{1}{2}\right)^2 \hbar^2 \,, \qquad n \in \mathbb{N}_0 \,, \tag{5.46}$$

following easily from either (5.41-5.43) or (5.44). The third constraint is *universal* since it does not depend on the function $f(x,y,w)$. Using the variables

$$u = \frac{1}{2}(x+y) > 0 \,, \quad v = \frac{1}{2}(x-y) \in \mathbb{R} \,,$$

we define the three-dimensional *space of* (*second*) *moments*, with coordinates (u,v,w). The third condition

$$u^2 - v^2 - w^2 = e_n^2 \,, \qquad e_n = \left(n + \frac{1}{2}\right)\hbar \,, \quad n \in \mathbb{N}_0 \,, \tag{5.47}$$

determines, for each non-negative integer, one sheet of a two-sheeted hyperboloid, located in the "upper" half of the space of moments, i.e. $u > 0$ and $v, w \in \mathbb{R}$ (cf. Fig. 5.1). The points on the n-th sheet, which intersects the u-axis at $u = +e_n$, are in one-to-one correspondence with the squeezed states originating from the number state $|n\rangle$, forming the set $\mathscr{E}_n$ in (5.36).

The consistency conditions (5.45) introduces constraints which depend on the function $f(x,y,w)$ at hand. These equations will only be satisfied for specific subsets $\mathscr{E}_n(f)$ of points on the hyperboloids $\mathscr{E}_n$, resulting in a set of f-dependent states

$$\mathscr{E}(f) = \bigcup_{n=0}^{\infty} \mathscr{E}_n(f)$$

which contains candidates to possibly minimize the functional $J[\psi]$. The candidate sets $\mathscr{E}(f)$ may depend on one or two parameters, or contain isolated points only. If the consistency conditions cannot be satisfied, then the functional $J[\psi]$ has no lower bound. Furthermore, if the matrix $\mathbf{F}$ is not positive definite for any of these states,

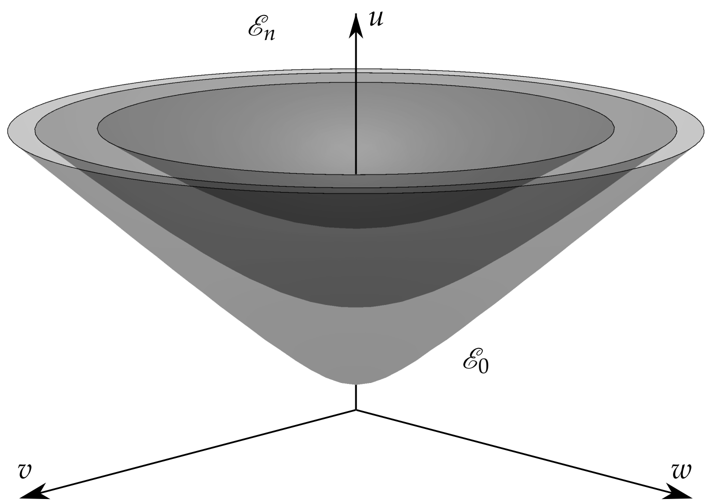

Figure 5.1: Space of (second) moments, with points (u, v, w): the extremal states of smooth functionals $J[\psi]$ are located on a discrete set of nested hyperboloids $\mathscr{E} = \bigcup_{n=0}^{\infty} \mathscr{E}_n$ the first three of which are shown, using light ($n = 0$), medium ($n = 1$) and dark shading ($n = 2$), respectively. The accessible *uncertainty region* for a quantum particle is given by the points on and inside of the convex surface $\mathscr{E}_0 : u^2 - v^2 - w^2 = \hbar^2/4$ which coincides with the minima $\mathscr{M}(f^{RS})$ of the RS inequality, i.e. squeezed states with minimal uncertainty.

the method makes no predictions about the minima of the functional $J[\psi]$. We have, however, not found any non-trivial cases of this behaviour.

The final step in determining the minima of the functional $J[\psi]$ will be to evaluate it for the candidate states $\mathscr{E}(f)$ and pick the lowest possible value. The states achieving this minimum value constitute the solutions $\mathscr{M}(f) \subseteq \mathscr{E}(f)$ of the minimization problem. In their entirety, the minima $\mathscr{M}(f)$ may consist of isolated states or of sets depending on one or two parameters. Usually, states on the sheet $\mathscr{E}_0$ are found to saturate the bound but is possible to manufacture cases where the minima are located inside the uncertainty region (see the discussion in Sec. 5.4).

Before turning to the discussion of known and new uncertainty relations, we briefly turn our attention to the space of moments. For $n = 0$, Eq. (5.47) is equivalent to (5.3) which implies that not all points $(u, v, w) \in \mathbb{R}^3$ can arise as moment triples. The accessible part of the space is called the *uncertainty region* shown in Fig. 5.1; it is bounded by $\mathscr{E}_0$, the hyperboloid defined in Eq. (5.47) for $n = 0$. Since the points on this surface coincide with the squeezed states of minimum uncertainty, the hyperboloid can also

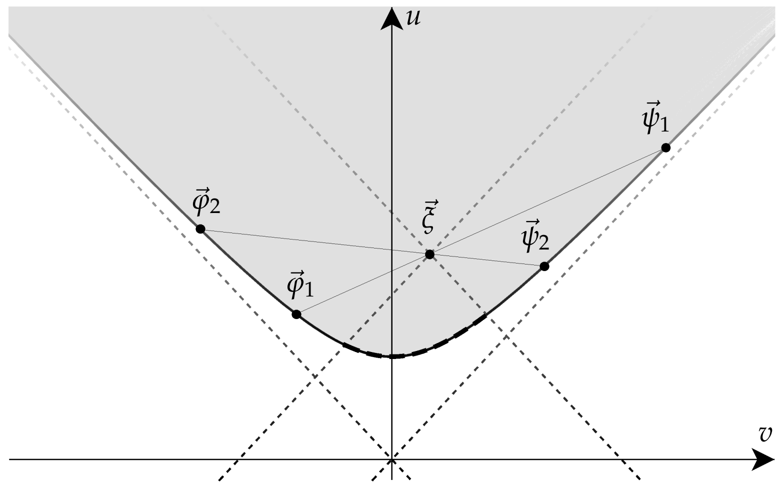

Figure 5.2: Cross-section ($w = 0$) of the uncertainty region (shaded) illustrating the convexity of its boundary $u^2 - v^2 - w^2 = \hbar^2/4$; convex combinations of moment triples located on the hyperboloid (associated with pure Gaussian states with minimal uncertainty) reproduce any given moment vector $\vec{\xi}$ inside the uncertainty region (the points must be outside of the "backward light-cone" of the point $\vec{\xi}$, indicated by the dashed segment of the hyperbola).

be described by the known condition for pure states to be Gaussian [67],

$$\det\left(\mathbf{C} + i\frac{\hbar}{2}\mathbf{\Omega}\right) = 0\,, \tag{5.48}$$

where $\mathbf{\Omega}$ and $\mathbf{C}$ are the standard symplectic and the covariance matrix, respectively.

The analogously defined uncertainty region of a quantum spin s [25] does not have a convex boundary from the outset. It is instructive to briefly discuss some consequences which follow from the fact that the uncertainty region of a quantum particle is bounded by a convex surface in the space of moments.

Consider a state $|\xi\rangle$ giving rise to moment vector $\vec{\xi} = (u_\xi, v_\xi, w_\xi)$ *inside* the uncertainty region. It is possible to *identify infinitely many pairs of Gaussian states on the boundary such that their mixture reproduces the given triple* $\vec{\xi}$. On the level of moments, it is geometrically obvious that any moment triple $\vec{\xi}$ can be reached as a convex combination of two points located on the boundary (cf. Fig. 5.2). It is sufficient to consider states with vanishing covariance, $w = 0$. Picking any point $\vec{\varphi}$ "space-like" relative to $\vec{\xi}$ and located on the hyperboloid, the pair determines a line intersecting the boundary in a unique point $\vec{\psi}$. Then, the desired point $\vec{\xi}$ must lie on the line segment

$\vec{\xi}(t) = \vec{\varphi} + t(\vec{\psi} - \vec{\varphi})$, $t \in [0,1]$, connecting the points $\vec{\varphi}$ and $\vec{\psi}$; it will pass through the point $\vec{\xi}$ if

$$t_0 = \frac{u_\xi - u_\varphi}{u_\psi - u_\varphi} \equiv \frac{v_\xi - v_\varphi}{v_\psi - v_\varphi} \in [0,1] \,. \tag{5.49}$$

When writing the line segment in the form $\vec{\xi}(t) = t\vec{\psi} + (1-t)\vec{\varphi}$, it becomes obvious that the reasoning valid in the space of moments extends to quantum states, i.e. the mixture

$$\hat{\rho}_{t_0} = t_0 \hat{P}_\psi + (1 - t_0)\hat{P}_\varphi \tag{5.50}$$

of the rank-1 projectors $\hat{P}_\psi = |\psi\rangle\langle\psi|$ and $\hat{P}_\varphi$ onto the associated Gaussian states defines a mixed quantum state with the desired moment triple $\vec{\xi}$. Clearly, any such convex combination can be used to create other mixtures with the same moment triple by rotating it rigidly about the u-axis.

Next, the moment vector (u_t, v_t, w_t) calculated from the mixed states $\hat{\rho}_t$ in (5.50) *can only lie inside the uncertainty region, not on its boundary*, as is shown in Appendix A.2, without reference to the geometric picture developed here. This observation retrospectively justifies our initial limitation to pure states when searching for minima of the functionals $J[\psi]$.

The relation between quantum states and points inside of the uncertainty region is, of course, many-to-one. For example, the first excited state of a unit oscillator $|1\rangle$ with moment vector $\vec{\xi}_1 = (9\hbar^2/4, 0, 0)$, being a pure state, cannot be written as a mixture of two Gaussians. Nevertheless, suitable mixtures of Gaussian states will produce the moment vector $\vec{\xi}_1$. The only moment vectors $\vec{\xi}$ which cannot be obtained from mixtures are those on the boundary. Here, the relation between states and moment vectors is one-to-one, in agreement with the fact that Gaussian states are determined uniquely by the covariance matrix $\mathbf{C}$.

Let us restrict our attention to Gaussian states of minimal uncertainty and their convex combinations only. Then, the uncertainty region of a quantum particle has a surprising number of features in common with the Bloch ball used to visualize the states of a qubit. Each state is characterized uniquely by a triple of numbers, the states on the convex boundary are the only pure states, and the decomposition of mixed states into pairs of pure states is clearly not unique. The group $\mathrm{SO}(2,1)$ of transformations which leave the uncertainty region invariant plays the role of the SU(2) transformations mapping the Bloch ball onto itself.

5.2.5 Mixed states

So far, we only considered pure states of a quantum particle when searching for the extrema of functions of its second moments. We now show that our findings do not change if we also allow for convex combinations of pure states. Each mixed state $\hat{\rho}$ generates a moment triple $\vec{\mu}$ with components $x = \text{Tr}\left(\hat{\rho}\hat{x}^2\right)$, etc., satisfying the Robertson-Schrödinger inequality [27]. Thus, the uncertainty region necessarily contains all potential mixed-state minima $\vec{\mu}$ of a given functional. Two cases occur.

If the triple $\vec{\mu}$ is located on one of the hyperboloids $\mathscr{E}$, then there exists a squeezed number state—i.e. a pure state— which gives rise to the same three expectations. Hence, the point $\vec{\mu}$ has already been included in the search for extrema.

Alternatively, the point $\vec{\mu}$ is located between two hyperboloids, $\mathscr{E}_n$ and $\mathscr{E}_{n+1}$, say, with $n \in \mathbb{N}_0$. Again, there is a pure state with moments given by $\vec{\mu}$. To see this, we first consider the points on the line segment connecting the points $(u_n, 0, 0)$ and $(u_{n+1}, 0, 0)$ associated with the number states $|n\rangle$ and $|n+1\rangle$, respectively. The moments of the superposition

$$|n\rangle_t = \sqrt{t}|n\rangle + \sqrt{1-t}|n+1\rangle, \quad t \in [0,1], \tag{5.51}$$

indeed lead to the moment triple

$$\vec{n}_t = \left(u_{n+1} + t\left(u_n - u_{n+1}\right), 0, 0\right), \quad t \in [0,1], \tag{5.52}$$

since all matrix elements of the the second moments between states of different parity vanish. Finally, any moment triple $\vec{\mu}$ off the u-axis will lie on a hyperboloid with a specific value of $t = t_0$, say. This moment triple can be obtained, however, from the state $\hat{S}(\xi)|n\rangle_{t_0}$, with a suitable value ξ. Using relativistic terminology, the operator $\hat{S}(\xi)$ must induce a Lorentz transformation which maps the given point on the u-axis to the desired point $\vec{\mu}$ on the same hyperboloid. In conclusion, each triple $\vec{\mu}$ of the uncertainty region can be achieved by the moments of a suitable pure state. Thus, mixed states do not give rise to candidates for minima different from those associated with pure states.

5.2.6 Known uncertainty relations

To illustrate our approach we re-derive three of the four bounds mentioned in the introduction: the uncertainty relations by *Robertson-Schrödinger*, by *Heisenberg-Kennard*, and the *triple-product* inequality.

Robertson-Schrödinger uncertainty relation. Defining

$$f^{RS}(x,y,w) = xy - w^2 \,, \tag{5.53}$$

the matrix of first-order derivatives associated with the quadratic form (5.19) is given by

$$\mathbf{F} = \begin{pmatrix} y & -w \\ -w & x \end{pmatrix} , \tag{5.54}$$

and, interestingly, its determinant

$$\det \mathbf{F} = xy - w^2 \equiv f^{RS}(x,y,w) \tag{5.55}$$

coincides with the original functional. At this point of the derivation, it is not yet known whether the matrix $\mathbf{F}$ is strictly positive.

The relations Eq. (5.45) do not produce any constraints on the parameters x, y, and w, since they are satisfied automatically leaving Eq. (5.46) as the only restriction. Since the left-hand-side of (5.46) coincides with the function $f^{RS}(x,y,w)$, all squeezed states are candidates to minimize the RS functional,

$$\mathscr{E}(f^{RS}) = \mathscr{E} \,.$$

Therefore, the function f^{RS} comes with the largest possible set of candidates to minimize it, given by the union of the sets $\mathscr{E}_n$ in Fig. 5.1. The lower bound on f^{RS} now follows directly from combining (5.46) with (5.53),

$$f^{RS}(x,y,w) = \left(n + \frac{1}{2}\right)^2 \hbar^2 \geq \frac{\hbar^2}{4}\,, \tag{5.56}$$

reproducing the RS inequality. The identity (5.55) implies that the determinant of **F** is positive everywhere in the uncertainty region.

The hyperboloid closest to the origin of the (u,v,w)-space provides the states minimizing the function f^{RS},

$$\mathscr{M}(f^{RS}) = \mathscr{E}_0\,,$$

i.e. the set of squeezed states based on the ground state $|0\rangle$ of a unit oscillator. This is, of course, a two-parameter set: the relations (5.23) take the form

$$b = -\frac{z}{x}\,, \quad \text{and} \quad \gamma = \frac{1}{2}\ln\left(2x\right)\,, \tag{5.57}$$

which means that, with $x > 0$ and $z \in \mathbb{R}$, both b and γ take indeed arbitrary real values. Thus, each squeezed state can be reached and, when adding phase-space translations, we obtain the four-parameter family of *all* squeezed states as minima of f^{RS}:

$$\mathscr{M}_\alpha(f^{RS}) = \hat{T}_\alpha \mathscr{M}(f^{RS})\,.$$

This property singles out the RS functional among all uncertainty functionals.

If an uncertainty functional, given by some function f, is different from the RS functional, the first two consistency relations will, in general, *not* be satisfied automatically but impose non-trivial constraints on the second moments. Therefore, the extrema of the functional must be a proper subset of those of the RS functional, i.e. $\mathscr{E}(f) \subset \mathscr{E}$, as the following example shows.

Heisenberg's uncertainty relation. Let us determine the minimum of the product of the standard deviations Δp and Δq by considering the function

$$f^H(x,y,w) = \sqrt{xy}\,. \tag{5.58}$$

Its partial derivatives satisfy

$$2f_x^H = \sqrt{\frac{y}{x}}\,, \qquad 2f_y^H = \sqrt{\frac{x}{y}}\,, \qquad f_w^H = 0\,, \tag{5.59}$$

resulting in a diagonal matrix $\mathbf{F}$ which is positive definite,

$$\mathbf{F} = \frac{1}{2}\begin{pmatrix} \sqrt{y/x} & 0 \\ 0 & \sqrt{x/y} \end{pmatrix}\,, \quad \det \mathbf{F} = \frac{1}{4} > 0\,. \tag{5.60}$$

Eqs. (5.41) and (5.42) collapse into a single set of conditions, namely

$$\sqrt{xy} = \left(n + \frac{1}{2}\right)\hbar \equiv e_n\,, \quad n \in \mathbb{N}_0\,, \tag{5.61}$$

which determines the value of the product of the standard deviations at the extrema of $f^H(x, y, w)$, labelled by the positive integers. In the (u, v, w)-space, the intersections of the surfaces defined by (5.61) and the hyperboloids (5.47) consist of hyperbolas in the (u, v)-plane containing the points $(e_n, 0)$, $n \in \mathbb{N}_0$. The union of these hyperbolas define the set $\mathscr{E}(f^H)$, which correspond to the potential minima of the function $f^H(x, y)$, illustrated in Fig. 5.3. The third condition, Eq. (5.43), implies that $w = 0$. Combining (5.61) with (5.58), we obtain the bound,

$$f^H(x, y, w) = \left(n + \frac{1}{2}\right)\hbar \geq \frac{\hbar}{2}\,, \tag{5.62}$$

reproducing Heisenberg's uncertainty relation (5.1).

The family of states minimizing Heisenberg's uncertainty relation is found by using the identity (5.61) for $n = 0$ in the definition of the parameter γ in (5.23), leading to $f_y = x/\hbar$. Since the consistency conditions do not impose any other condition on the variance x, it may take any positive value implying that $\gamma \in \mathbb{R}$. Since $\mathrm{f}_w = 0$ leads to $b = 0$, the set of states minimizing the Heisenberg's uncertainty relation is given by squeezed states with real squeezing parameter,

$$\mathscr{M}_\alpha(f^H) = \hat{T}_\alpha \mathscr{E}(f^H) \equiv \left\{ \hat{T}_\alpha \hat{S}_\gamma |0\rangle\,,\, \alpha \in \mathbb{C}, \gamma \in \mathbb{R} \right\}\,, \tag{5.63}$$

where we have re-introduced arbitrary phase-space displacements.

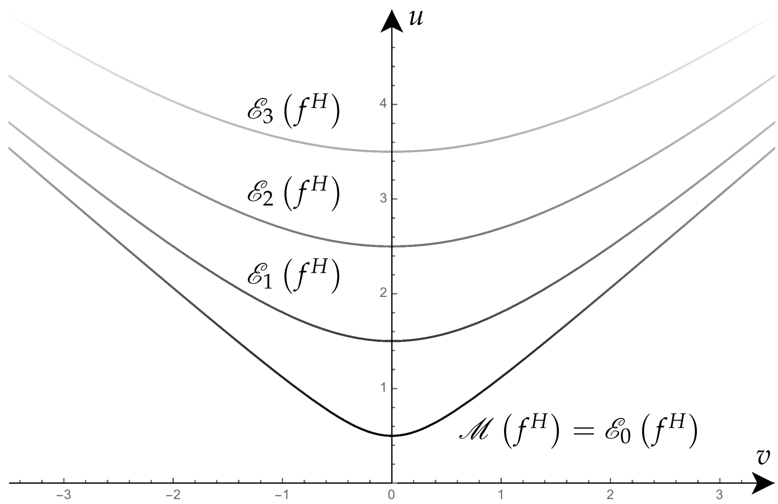

Figure 5.3: Hyperbolas in the (u, v)-plane through the points $(e_n, 0)$, $n \in \mathbb{N}_0$, stemming from intersections of the hyperboloids (5.47) and the surfaces defined by (5.61). The union of the hyperbolas defines the set $\mathscr{E}(f^H)$ which represents the location of all possible minima of the function $f^H(x, y)$; the points on the "lowest" (darkest) hyperbola $\mathscr{E}_0(f^H)$ correspond to the set of states $\mathscr{M}(f^H)$ which saturate Heisenberg's uncertainty relation.

Triple product inequality. Using Eq. (5.5), we see that we need to find the minimum of the expression

$$f^T(x, y, w) = xy\,(x + y + 2w) \tag{5.64}$$

in order to reproduce the triple product uncertainty relation (5.4). The first consistency condition in Eq. (5.45) implies that $x = y$; using this identity in the second condition, one finds

$$x(w + x)(w + \frac{x}{2}) = 0\,. \tag{5.65}$$

Recalling that variances must be positive, $x > 0$, the correlation w must equal either $-x$ or $-x/2$. According to (5.5), the first case would imply $\Delta^2 r \equiv 0$, which is impossible since the operator $\hat{r}$ has no normalizable eigenstates. Therefore, using the only solution $w = -x/2$ of (5.65) and $x = y$ in the third consistency condition, one finds that

$$x^2 = \frac{4}{3}\left(n + \frac{1}{2}\right)^2 \hbar^2\,, \quad n \in \mathbb{N}_0\,, \tag{5.66}$$

must hold. It is now straightforward to evaluate $f^T(x,y,w)$ at its extrema to find its global minimum,

$$f^T(x,y,w) = x^3 = \left(\sqrt{\frac{4}{3}}\left(n+\frac{1}{2}\right)\hbar\right)^3 \geq \left(\tau\frac{\hbar}{2}\right)^3 ,$$

which reproduces (5.4). It is easy to confirm that the matrix $\mathbf{F}$ is positive definite with determinant $\det \mathbf{F} = \hbar^4/3$. Since the minimum occurs for $n = 0$, the values of the second moments are given by

$$x = y = -2w = \frac{\hbar}{\sqrt{3}} \equiv \tau\frac{\hbar}{2} . \tag{5.67}$$

These relations fix the values of the parameters in (5.23),

$$b = \frac{1}{2} \quad \text{and} \quad \gamma = \frac{1}{4}\ln\tau . \tag{5.68}$$

Using (5.32) or (5.36) we obtain one single state which saturates the triple uncertainty, namely

$$|\Xi_0\rangle \equiv \hat{G}_{\frac{1}{2}}\hat{S}_{\frac{1}{2}\ln\tau}|0\rangle = \hat{S}_{\frac{i}{4}\ln 3}|0\rangle , \tag{5.69}$$

in agreement with the findings of Chapter 3. If one includes rigid phase-space translations, the set of states minimizing the triple uncertainty is finally given by the two-parameter family

$$\mathscr{M}_\alpha(f^T) = \left\{\hat{T}_\alpha|\Xi_0\rangle , \alpha \in \mathbb{C}\right\} , \tag{5.70}$$

in agreement with [39].

Geometrically, the state $|\Xi_0\rangle$ arises from the intersection of the sequence of hyperboloids with the surfaces defined by

$$u = \tau e_n , \quad v = 0 , \quad w = -\frac{1}{2}\tau e_n , \quad n \in \mathbb{N}_0 . \tag{5.71}$$

The planes defined by constant values of u have concentric circles in common with the hyperboloids, and the vertical uw-plane (given by $v = 0$) intersects with each of the circles in two points only. Finally, the condition on the variable w selects a single one of the points with the same value of u. According to (5.71), the candidate states in

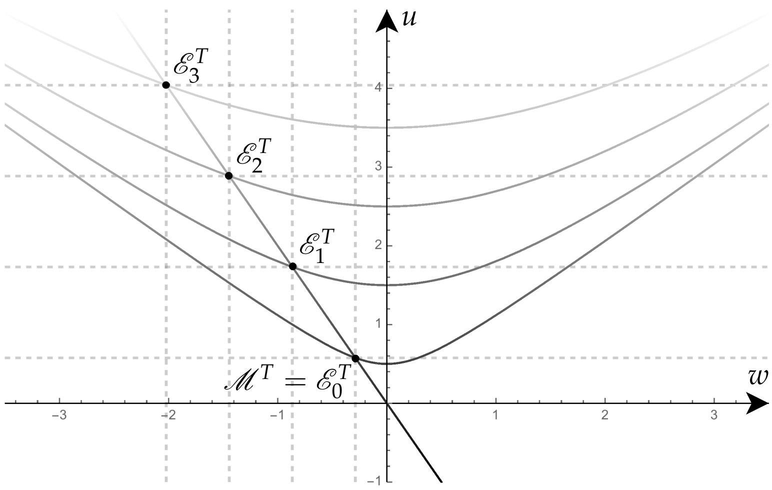

Figure 5.4: Candidate states $\mathscr{E}(f^T) \equiv \mathscr{E}^T$ possibly minimizing the product of three variances $f^T(u,v,w)$, represented by dots located on the intersections of the hyperboloids (5.47) and the planes defined by the consistency conditions (5.71); the point closest to the origin, $\mathscr{M}(f^T) \equiv \mathscr{M}^T$, represents the state $|\Xi_0\rangle$ achieving the minimum of the triple product uncertainty relation (5.4) (and of any other S_3-invariant inequality associated with a functional $f_N^{(3)}$ in (5.91)).

(u,v,w)-space are located on a straight line,

$$\mathscr{E}(f^T) = \left\{ \tau e_n \begin{pmatrix} 1 \\ 0 \\ -1/2 \end{pmatrix}, n \in \mathbb{N}_0 \right\}, \tag{5.72}$$

and the state $|\Xi_0\rangle$ corresponds to the point closest to the origin.

5.3 New uncertainty relations

5.3.1 Generalizing known relations

The *linear combination* of second moments

$$f^L(x,y,w) = \mu x + \nu y + 2\lambda w\,, \qquad \mu,\nu,\lambda, \in \mathbb{R}\,, \tag{5.73}$$

is now shown to imply the uncertainty relation

$$\mu\Delta^2 p + \nu\Delta^2 q + 2\lambda C_{pq} \geq \hbar\sqrt{\mu\nu - \lambda^2}, \qquad \mu,\nu > 0\,, \quad \mu\nu > \lambda^2\,. \tag{5.74}$$

The constraints on the parameters follow from the matrix $\mathbf{F}$ in (5.19) being strictly positive definite. The consistency conditions (5.45) associated with f^L relate both y and w to x according to

$$y = \frac{\mu}{\nu}x\,, \qquad w = -\frac{\lambda}{\nu}x\,. \tag{5.75}$$

Then, Eq. (5.46) simplifies to

$$\frac{\left(\mu\nu - \lambda^2\right)}{\nu^2}x^2 = \left(n + \frac{1}{2}\right)^2 \hbar^2 = e_n^2\hbar^2\,, \tag{5.76}$$

which is consistent due to $\det \mathbf{F} = \mu\nu - \lambda^2 > 0$. Expressing the functional $f^L(x,y,w)$ in terms of x only, we obtain the bound given in (5.74),

$$f^L(x,y,w) = \frac{2\left(\mu\nu - \lambda^2\right)}{\nu}x = 2e_n\hbar\sqrt{\mu\nu - \lambda^2} \geq \hbar\sqrt{\mu\nu - \lambda^2}\,. \tag{5.77}$$

Up to phase-space translations $\hat{T}_\alpha$, a single squeezed state saturates the bound, namely

$$\mathscr{M}(f^L) = \left\{ |\mu,\nu,\lambda\rangle = \hat{G}_{\frac{\lambda}{\nu}}\hat{S}_{\frac{1}{2}\ln\left(\frac{\nu}{\sqrt{\mu\nu-\lambda^2}}\right)}|0\rangle \right\}\,. \tag{5.78}$$

When expressing the correlation term C_{pq} in terms of the variance $\Delta^2 r$ according to Eq. (5.5), we obtain, for $\mu = \nu = 2\lambda = 1$ in (5.74), the *triple sum* uncertainty relation

$$\Delta^2 p + \Delta^2 q + \Delta^2 r \geq \sqrt{3}\hbar\,, \tag{5.79}$$

derived in [39], and the minimum is achieved for the state $|1,1,1/2\rangle \equiv |\Xi_0\rangle$ which also minimizes the triple product uncertainty (cf. Eq. (5.69) and Fig. 5.4).

Sums of powers of position and momentum variances are bounded from below according to the inequality

$$\mu\left(\Delta^2 p\right)^m + \nu\left(\Delta^2 q\right)^{m'} \geq \left(\frac{\hbar}{2}\right)^{\frac{2mm'}{m+m'}} \left(\mu\left(\frac{\nu}{\mu}\frac{m'}{m}\right)^{\frac{m}{m+m'}} + \nu\left(\frac{\mu}{\nu}\frac{m}{m'}\right)^{\frac{m'}{m+m'}}\right), \quad \mu,\nu > 0\,,\ m,m' \in \mathbb{N}\,, \tag{5.80}$$

reducing to the *pair sum* uncertainty relation (5.2) in the simplest case ($\mu = \nu = m = m' = 1$).

Next, we study a *generalized RS-uncertainty functional* (5.53),

$$f^{gRS}_{m,m'}(x,y,w) = (xy)^m - \mu w^{m'}\,, \qquad \mu > 0\,, \quad m,m' \in \mathbb{N}\,, \tag{5.81}$$

which provides an example for which the consistency conditions cannot be solved in closed form for arbitrary integers m and n. Setting $m' = 2m$ and assuming that both $m > 1$ and $\mu > 1$ hold, we obtain the explicit bound

$$\left(\Delta^2 p \cdot \Delta^2 q\right)^m - \mu\left(C_{pq}\right)^{2m} \geq \left(\frac{\hbar}{2}\right)^{2m} \frac{\mu}{\left(\mu^{\frac{1}{m-1}} - 1\right)^m}\,. \tag{5.82}$$

An interesting special case of f^{gRS} occurs for $m = 1/2$ and $0 < \mu < 1$,

$$\Delta p \Delta q - \mu\left|C_{pq}\right| \geq \frac{\hbar}{2}\sqrt{1-\mu^2}\,, \tag{5.83}$$

which can be treated in spite of the presence of a non-differentiable term. The extremal states depend on one free parameter,

$$\mathscr{E}_\alpha(f^{gRS}_{1,1/2}) = \left\{|\alpha,n\rangle = \hat{T}_\alpha \hat{G}_{\pm\frac{\mu e_n \hbar}{2x\sqrt{1-\mu^2}}} \hat{S}_{\frac{1}{2}\ln\left(\frac{x}{e_n\hbar}\right)}|n\rangle\right\}\,, \qquad x > 0\,, \tag{5.84}$$

reducing for $\mu = 0$, i.e. in the absence of the covariance term, to the squeezed number states with a real parameter extremising Heisenberg's inequality.

Next, we present an example of an uncertainty relation which seems to be entirely out of reach of traditional derivations. Defining the functional

$$f^e(x,y) = x + \mu e^{y/\nu}\,, \quad \mu,\nu > 0\,, \tag{5.85}$$

we obtain the inequality

$$\Delta^2 p + \mu e^{\Delta^2 q/\nu} \geq \left(1 + 2W(\hbar/4\sqrt{\mu\nu})\right) e^{2W(\hbar/4\sqrt{\mu\nu})}\,, \tag{5.86}$$

using the fact that Lambert's W-function $W(s)$, defined as the inverse of $s(W) = W \exp W$, is a strictly increasing function. In the limit of $\mu \to \infty$ and after setting $\mu = \nu$, the left-hand-side of (5.86) turns into $(\mu + \Delta^2 p + \Delta^2 q + \mathscr{O}(1/\mu))$ while the expansion of its right-hand-side produces the correct bound $(\mu + \hbar + \mathscr{O}(1/\mu))$, since $W(s) = s + \mathscr{O}(s^2)$.

The position and momentum variances at the extremum with label $n \in \mathbb{N}_0$ are given by

$$x = 2\mu W\left(\frac{e_n\hbar}{2\sqrt{\mu\nu}}\right) e^{W\left(\frac{e_n\hbar}{2\sqrt{\mu\nu}}\right)} \tag{5.87}$$

and

$$y = 2\nu W\left(\frac{e_n\hbar}{2\sqrt{\mu\nu}}\right)\,, \tag{5.88}$$

respectively. Using Eqs. (5.23) with $\det \mathbf{F} = (\mu/\nu)e^{y/\nu}$, one finds

$$b = 0\,, \qquad \gamma = \frac{1}{4}\ln\left(\frac{\mu}{\nu}\right) + \frac{1}{2}W\left(\frac{e_n\hbar}{2\sqrt{\mu\nu}}\right)\,, \tag{5.89}$$

which means that only a single state (and its rigid displacements) will saturate the inequality (5.86). If $\mu = \nu$, we recover $x = y = \hbar^2/4$ as well as $b = \gamma = 0$, i.e. the ground state of a unit oscillator since $W(0) = 0$.

Finally, we point out that some general statements can be made for functionals of the form $f = f(xy, w)$ and $f = f(\mu x^m + \nu y^{m'}, w)$, i.e. generalizations of the expressions in (5.80) and (5.81), respectively. By examining the consistency conditions one can show any existing extrema of the first expression come as a one-parameter set, while they are isolated or a one-parameter family in the second case. However, without knowing the explicit form of the functions no further conclusions can be drawn (see Appendix A.3).

5.3.2 Uncertainty functionals with permutation symmetries

The triple product uncertainty relation and the one derived by Kennard possess discrete symmetries. Here we investigate more general uncertainty functionals which are invariant under the exchange of two and three variances.

S_3-invariant functionals

Consider a function of three variables which is invariant under the exchange of any pair,

$$f^{(3)}(x,y,z) = f^{(3)}(y,x,z) = f^{(3)}(x,z,y)\,. \tag{5.90}$$

We now derive the lower bound of a large class of S_3-invariant uncertainty functionals $J[\psi]$ and show that their minima coincide with the state $|\Xi_0\rangle$ minimizing the triple product inequality. The variables x, y and z will be taken as the variances of the operators $\hat{p}, \hat{q}$ and $\hat{r}$, respectively.

More specifically, we study the minima of sums of completely homogeneous polynomials of degree n, with arbitrary *non-negative* coefficients,

$$f_N^{(3)}(x,y,z) = \sum_{n=1}^{N} \sum_{j+k+\ell=n} a_{jk\ell} x^j y^k z^\ell\,, \quad a_{jk\ell} \geq 0\,, \tag{5.91}$$

dropping the unimportant constant term a_{000}. The associated $\mathbf{F}$-matrix

$$\det \mathbf{F} \equiv f_x f_y + f_y f_z + f_z f_x \tag{5.92}$$

is positive definite since $x, y, z > 0$ and the partial derivatives are just positive polynomials.

The symmetry under S_3-permutations (5.90) implies that the coefficients must satisfy the conditions

$$a_{jk\ell} = a_{kj\ell} = a_{j\ell k}\,, \quad 0 \leq j, k, \ell \leq n\,. \tag{5.93}$$

The first terms of the polynomials we consider are given by

$$
\begin{aligned}
f_N^{(3)}(x,y,z) \quad = \quad & a_{100}(x+y+z) \\
& + a_{200}\left(x^2+y^2+z^2\right) + a_{110}\left(xy+yz+zx\right) \\
& + a_{300}\left(x^3+y^3+z^3\right) + a_{210}\left(x^2(y+z)+y^2(z+x)+z^2(x+y)\right) \\
& + a_{111}xyz + \dots
\end{aligned}
\tag{5.94}
$$

If the only nonzero coefficients are $a_{100} = 1$ or $a_{111} = 1$, we recover the functionals associated with the triple sum (5.79) or the triple product inequality (3.5), respectively. In general, a completely homogeneous S_3-symmetric polynomial in three variables of degree $n \geq 1$ consists of $\kappa_n = \left\lfloor \frac{(n+3)^2+6}{12} \right\rfloor$ terms where the floor function $\lfloor s \rfloor$ denotes the integer part of the number s: each term arises from one way to partition $j+k+\ell = n$ objects into three sets with j, k and ℓ elements, respectively [68]. Thus, a symmetric polynomial of degree up to N depends on $\left(\sum_{n=1}^{N} \kappa_n\right)$ independent coefficients if one ignores the constant term.

The main result of this section follows from rewriting the consistency conditions of Chapter 5, (5.45) and (5.46) in terms of the variables x, y and z,

$$
x f_x - y f_y + (x-y) f_z = 0\,, \tag{5.95}
$$

$$
z f_z - x f_x + (z-x) f_y = 0\,, \tag{5.96}
$$

$$
2\left(xy+yz+zx\right) - x^2 - y^2 - z^2 = (2n+1)^2 \hbar^2\,, \tag{5.97}
$$

where we have used the identity $z = x + y + 2w$ given in (5.5). The first two consistency conditions imply that the extrema of any symmetric polynomial $f_N^{(3)}(x,y,z)$ occur whenever the three variances take the same value,

$$
x = y = z\,. \tag{5.98}
$$

To show that $x = y$ holds we pick any nonzero term $a_{jk\ell}x^j y^k z^\ell$ in the expansion (5.91) and assume that the powers of x and y are different, i.e. $j \neq k$; the case $j = k$ will be considered later. Due to the symmetry under the exchange $x \leftrightarrow y$, the sum also must contain the term $a_{kj\ell}x^k y^j z^\ell$, with $a_{kj\ell} \equiv a_{jk\ell}$. Defining $t(x,y,z) =$

$a_{jk\ell}\left(x^j y^k + x^k y^j\right) z^\ell$, the first two terms of (5.95) take the form

$$xt_x - yt_y = (j-k)a_{jkl}\left(x^j y^k - x^k y^j\right) z^\ell \,. \tag{5.99}$$

Assuming that $j = k + \delta$, with $\delta > 0$, we find

$$\begin{aligned} xt_x - yt_y &= a_{k+\delta\, kl}\delta\left(x^\delta - y^\delta\right) x^k y^k z^\ell & (5.100)\\ &= (x-y)\delta a_{k+\delta\, k\ell}\left(x^{\delta-1} + x^{\delta-2}y + \ldots + xy^{\delta-2} + y^{\delta-1}\right) x^k y^k z^\ell & (5.101)\\ &\equiv (x-y)g_+(x,y,z)\,, & (5.102)\end{aligned}$$

where $g_+(x,y,z) > 0$. Using this expression in (5.95), the consistency condition takes the form

$$(x-y)\left(g_+(x,y,z) + t_z\right) = 0\,, \tag{5.103}$$

with another positive function $t_z(x,y,z)$. If $\delta < 0$ we write $k = j - \delta \equiv j + |\delta|$ and eliminate k instead of j from (5.99), only to find that its left-hand-side again turns into $(x-y)$ multiplied with a positive function. If the powers of x and y of the term $a_{kj\ell}x^k y^j z^\ell$ are equal, $j = k$, one immediately finds that $(x\partial_x - y\partial_y)a_{jkl}x^j y^j z^\ell = 0$, also reducing Eq. (5.95) to $(x-y)\partial_z a_{jkl}x^j y^j z^\ell = 0$.

The argument just given covers *all* terms in the sum (5.91), and the positivity of the coefficients $a_{jk\ell}$ implies that the first consistency condition can only be satisfied for $x = y$. Using the symmetry of $f_N^{(3)}(x,y,z)$ under the exchange $y \leftrightarrow z$, an identical argument leads to the identity $y = z$.

Using (5.98) to evaluate the left-hand-side of the third consistency condition (5.97) results in

$$x = y = z = \tau e_n \hbar\,, \qquad n \in \mathbb{N}_0\,, \tag{5.104}$$

where $\tau = \sqrt{4/3}$, so that we obtain an uncertainty relation for any S_3-invariant function

$$f_N^{(3)}(x,y,z) \geq f_N^{(3)}(x,y,z)\Big|_{x=y=z=\hbar/\sqrt{3}}\,. \tag{5.105}$$

This result correctly reproduces the special cases of Eqs. (3.5) and (5.79), and there is only one state which saturates the inequality, namely $|\Xi_0\rangle$ given in Eq. (5.69). Letting

$N \to \infty$ in Eq. (5.91), we conclude that the main result of this section, Eq. (5.105), also applies to any S_3-symmetric function $f_\infty^{(3)}(x,y,z)$ with a Taylor expansion with positive coefficients and infinite radius of convergence, as long as its first partial derivatives exist.

S_2-invariant functionals

Assume now that, in analogy to Eq. (5.90), we have a functional depending on just two variances in a symmetric way,

$$f^{(2)}(x,y) = f^{(2)}(y,x)\,. \tag{5.106}$$

An argument similar to the one given for the function $f^{(3)}(x,y,z)$ results in the uncertainty relation

$$f_N^{(2)}(x,y) \geq f_N^{(2)}(x,y)\Big|_{x=y=\hbar/2}\,, \tag{5.107}$$

which covers the cases of Heisenberg's relation (5.1) and the pair sum inequality (5.2). Thus, the actual form of the function at hand determines whether the set of minima $\mathscr{M}(f_N^{(2)})$ will depend on a continuous parameter or not. If the functional is invariant under scaling transformation $x \to \lambda x$, $y \to y/\lambda$, in addition to the permutation symmetry, there is a one-parameter family of solutions and the right-hand-side of Eq. (5.107) achieves its minimum on the set of points with $xy = (\hbar/2)^2$, not just those with $x = y = \hbar/2$.

To derive (5.107), we suppose that the function $f_N^{(2)}(x,y)$ has an expansion in analogy to $f_N^{(3)}(x,y,z)$ in Eq. (5.91) but without the variable z. Adapting the reasoning valid for $f_N^{(3)}(x,y,z)$, the consistency equations (5.45) are found to imply $x = y$ and $w = 0$. Using this result in (5.46), the bound $x \geq \hbar^2/4$ follows immediately, so that the inequality (5.107) must hold for S_2-invariant functionals.

5.4 Summary and discussion

In this chapter we studied the extrema of smooth functions of the second moments, $f(\Delta^2 p, \Delta^2 q, C_{pq})$, and the resulting framework explains the universal role of squeezed states for preparational uncertainty relations of the second moments, while it completely charts the landscape of inequalities of this type.

The chain of inclusions

$$\mathscr{E} \supseteq \mathscr{E}(f) \supseteq M(f) \tag{5.108}$$

concisely summarizes the general structure of our findings. First, we have shown that squeezed number states of a unit oscillator, i.e. a quantum mechanical harmonic oscillator with unit frequency and mass, occur naturally as extrema of an uncertainty functional $J[\psi]$ depending on second moments. We denote this *universal* set of states by $\mathscr{E}$. Second, the extrema of a *specific* functional $J_f[\psi]$, associated with a specific function $f(\Delta^2 p, \Delta^2 q, C_{pq})$, form the subset $\mathscr{E}(f)$ of the universal set $\mathscr{E}$. Third, the functional will assume its minimum for one or more of the extrema $\mathscr{E}(f)$, a subset which we denote by $\mathscr{M}(f)$. The set of minima may be empty, $\mathscr{M}(f) = \varnothing$. If it is not empty, a lower bound on the functional $J_f[\psi]$ has been found, and it represents a preparational uncertainty relation in terms of the second moments.

Strictly speaking, we obtained the relations $\mathscr{E}_\alpha \supseteq \mathscr{E}_\alpha(f) \supseteq \mathscr{M}_\alpha(f)$ instead of Eq. (5.108) as it is possible to move quantum states in phase space without affecting the values of the second moments. The four-parameter set $\mathscr{E}_\alpha \equiv \hat{T}_\alpha \mathscr{E}$, for example, consists of the squeezed states $\mathscr{E}$ plus those obtained from them by means of the translation operator $\hat{T}_\alpha$ which rigidly displaces the origin of phase space to the point α. Thus, each state saturating a specific inequality with vanishing expectation values gives rise to a two-parameter family of minima.

Our results have a useful geometric representation in the real three-dimensional space of moments. In this space, the uncertainty region consists of all triples of moments which can arise from (pure or mixed) states of a quantum particle. The region turns out to be a convex set bounded by a one-sheeted hyperboloid. Each point on this hyperboloid is associated with a unique Gaussian state saturating the Robertson-Schrödinger inequality. The RS-functional is invariant under the group $\mathrm{Sp}(2, \mathbb{R})$ applied to the canonical pair $(\hat{p}, \hat{q})$, i.e. rotations, scalings and linear gauge transformations. They induce $\mathrm{SO}(1,2)$ transformations in the space of moments, leaving invariant non-overlapping hyperboloids which foliate the uncertainty region. The boundary of the uncertainty region is, in particular, invariant under the elliptic rotations,

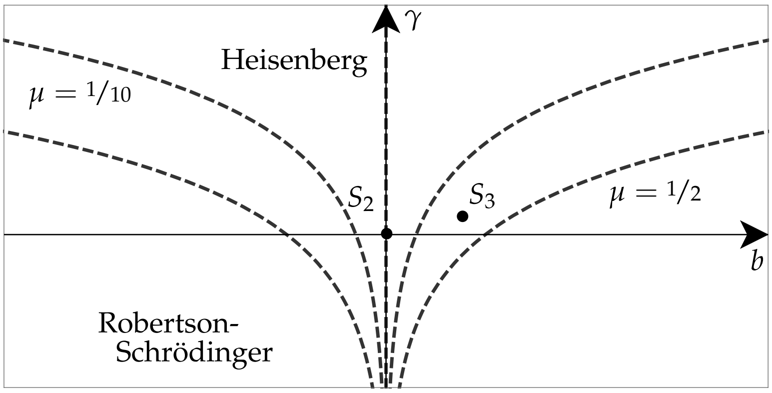

Figure 5.5: States on the boundary of the uncertainty region minimizing known and new uncertainty relations parametrized by the real numbers (b, γ), with $\hbar = 1$; each point of the plane corresponds to a squeezed state saturating the RS-inequality (5.3); points on the vertical dashed line represent minima of Heisenberg's uncertainty relation (5.1); the two curved dashed lines indicate the minima of the modified RS-inequality (5.83) with $m = 1/2$ and values $\mu = 1/2$ (bottom) and $\mu = 1/10$ (top); the full dots correspond to minima of S_2-invariant functionals (5.107) such as the pair sum (5.2) and S_3-invariant functionals (5.105) such as the triple product (5.4).

hyperbolic boosts and parabolic transformations which generate the group $\mathrm{SO}(1,2)$. This observation agrees with the importance of the group $\mathrm{SO}(1,2)$ in quantum optics, where coherent and squeezed states are ubiquitous.

We also derived new and explicit uncertainty relations; Fig. 5.5 illustrates the sets of states which minimize (some of) the uncertainty relations discussed in this chapter. The figure shows the (b, γ)-plane, which provides a visualisation of the boundary of the uncertainty region. The sets of minima $\mathscr{M}$ may depend on *two* parameters (all squeezed states minimizing the RS-inequality), on *one* parameter (such as the real squeezed states saturating Heisenberg's uncertainty relation) or consist of a *single point* only (associated with S_3-invariant inequalities such as the triple product inequality, for example). We have not been able to backward-engineer functionals which would be minimized by prescribed subsets of the plane such as a circle or a disk.

The minimizing states found so far were all pure, located on the *boundary* of the uncertainty region. However, functionals may, of course, also take their minima *inside* this region which would allow for extremal *mixed* states. The trivially bounded functional

$$f_0(x, y, w) = (x - x_0)^2 + (y - y_0)^2 + (w - w_0)^2 \,, \tag{5.109}$$

where the triple (x_0, y_0, w_0) denotes any point inside the uncertainty region, is a simple example. By construction, the minimum $f_0 = 0$ is achieved by any pure or mixed state with moment triple (x_0, y_0, w_0). In spite of its simplicity, this result does not follow from our approach since the sign of the determinant of $\mathbf{F}_0$,

$$\det \mathbf{F}_0 = 4(x - x_0)(y - y_0) - (w - w_0)^2 , \tag{5.110}$$

is not definite: if $w = w_0$, for example, the remaining term divides the xy-plane into four quadrants where $\det \mathbf{F}_0$ takes alternating signs. However, such trivial examples are excluded by the extra conditions imposed from the requirement that the function f is a measure of the overall uncertainty, as defined in the introduction of this chapter.

Three conceptually interesting generalizations of our approach are worth investigating. First, there is no fundamental reason to restrict oneself to uncertainty functionals depending only on the second moments in position and momentum [17]. On the contrary, higher order expectation values would enable us to move away from Gaussian quantum mechanics which is largely reproducible in terms of a classical model "with an epistemic restriction" of the allowed probability distributions [10]. Including fourth-order terms $\langle\psi|\hat{q}^4|\psi\rangle$, for example, will result in a eigenvalue equation (5.18) which is not related to a unit oscillator in a simple way. It is known that fourth-order moments for single-particle expectations can give rise to inequalities which cannot be reproduced by models based on classical probabilities [12, 42]. Thus, it might become possible to study truly non-classical behaviour in a systemic manner using suitable uncertainty functionals.

Secondly, our approach can be generalized to the case of two or more continuous variables. We expect that a systematic study of uncertainty functionals becomes possible, leading to criteria which would detect pure entangled states. More specifically, considering monotonically increasing functions of the variances of commuting operators and deriving the lower bounds in separable and entangled states would allows us to construct entanglement detection criteria. Known results appear to mainly rely on intuitive choices of suitable bi-linear observables [30, 63].

Finally, the comprehensive study [25] of uncertainty relations for a single spin s has been limited to observables transforming covariantly under the group SU(2). The method proposed here is easily adapted to investigate functionals depending on arbitrary functions of moments.

Chapter 6

Uncertainty relations for multiple degrees of freedom and states

6.1 Introduction

The aim of this chapter is to extend the framework developed in Chapter 5 to the case of more than one spatial degree of freedom. Our starting point is again an uncertainty functional, which in this case, depends on the $N(2N+1)$ second moments in an arbitrary quantum state. In analogy to the one dimensional case, under the assumption that it is suitably well behaved, looking for its extrema leads to an eigenvalue equation, quadratic in the $2N$, in total, position and momentum operators. Whenever the coefficients of the quadratic operator form a matrix that is positive (or negative) definite, the theorem of Williamson ensures that it can be diagonalised using a symplectic transformation. Although the eigenvalue equation is highly non-linear in the state, since every coefficient is in general a function of it, we can still produce a solution by initially assuming that they are constants, and only later re-introducing their dependence on the state through the *consistency conditions*. Similarly to the one dimensional case, we find a set of states that are universal in that every uncertainty functional attains its minimum, when there is one, for some states belonging to that set.

6.2 Extremal Uncertainty

6.2.1 Solving the eigenvalue equation

Any potential quantity of the second moments in N spatial degrees of freedom can be expressed as a function f of $N(2N+1)$ variables, and thus the functional can be explicitly written as

$$J[\psi] = f\left(\Delta^2 p_1, \Delta^2 q_1, C_{p_1 q_1}, \dots C_{p_1 p_2}, C_{p_1 q_2}, \dots ; \psi\right) - \lambda(\langle\psi|\psi\rangle - 1)\,, \tag{6.1}$$

where the Lagrange multiplier ensures that the potential solutions are normalised, and in this notation, we write first the "local " second moments for each degree of freedom and then the ones mixing them. Varying the functional as in Sec. 5.2.1, we obtain an eigenvalue equation, quadratic in all position and momentum operators.

For example, the eigenvalue equation for two degrees of freedom is given by

$$\begin{aligned}\left(\sum_{i=1}^{2}\left(f_{x_i}\hat{p}_i^2 + f_{y_i}\hat{q}_i^2 + \frac{f_{w_i}}{2}\left(\hat{q}_i\hat{p}_i + \hat{p}_i\hat{q}_i\right)\right) + f_{w_3}\hat{p}_1\hat{p}_2 + f_{w_4}\hat{p}_1\hat{q}_2 + f_{w_5}\hat{q}_1\hat{p}_2 + f_{w_6}\hat{q}_1\hat{q}_2\right)|\psi\rangle = \\ = \left(\sum_{i=1}^{2}\left(x_i f_{x_i} + y_i f_{y_i} + w_i f_{z_i}\right) + \sum_{j=3}^{6} w_j f_{z_j}\right)|\psi\rangle\,,\end{aligned} \tag{6.2}$$

where we have exploited the invariance of the second moments under phase space translations. In matrix notation the last equation becomes

$$\left(\hat{\mathbf{z}}^\top \cdot \mathbf{F} \cdot \hat{\mathbf{z}}\right)|\psi\rangle = L|\psi\rangle\,, \tag{6.3}$$

where $\hat{\mathbf{z}}^\top = (\hat{p}_1, \hat{q}_1, \hat{p}_2, \hat{q}_2)$, and $w_3, \dots, w_6$ the moments mixing the degrees of freedom. The form of (6.3) is the same for N degrees of freedom with the obvious extensions of the definitions, which we will assume from now on.

As in the one dimensional case, in order to provide the general solution of Eq. (6.3), we initially make the assumption that the matrix of partial derivatives of f, $\mathbf{F}$, does not depend on the state and is constant. In addition, if we assume that $\mathbf{F}$ is positive (or negative) definite, it follows from Williamson's theorem that there is a matrix $\mathbf{S}$ such that $\mathbf{F} = \mathbf{S}^\top \mathbf{D} \mathbf{S}$, where $\mathbf{D} = \mathrm{diag}(\lambda_1, \lambda_1, \ldots, \lambda_N, \lambda_N)$ and the positive real numbers λ_i are the symplectic eigenvalues of $\mathbf{F}$. Multiplying both sides of Eq. (6.3) by the metaplectic unitary operator from the left $\hat{S}^\dagger$, defined through

$$\mathbf{S} \cdot \hat{\mathbf{z}} = \hat{S}\, \hat{\mathbf{z}}\, \hat{S}^\dagger , \tag{6.4}$$

we find that the left hand side is equivalent to

$$\begin{aligned} \hat{S}^\dagger \left(\hat{\mathbf{z}}^\top \cdot \mathbf{F} \cdot \hat{\mathbf{z}} \right) \hat{S} \left(\hat{S}^\dagger |\psi\rangle \right) &= \left(\hat{S}^\dagger \hat{\mathbf{z}}^\top \hat{S} \right) \mathbf{F} \left(\hat{S}^\dagger \hat{\mathbf{z}} \hat{S} \right) \left(\hat{S}^\dagger |\psi\rangle \right) \\ &= \left(\mathbf{S}^{-1} \hat{\mathbf{z}} \right)^\top \left(\mathbf{S}^\top \mathbf{D} \mathbf{S} \right) \left(\mathbf{S}^{-1} \hat{\mathbf{z}} \right) \left(\hat{S}^\dagger |\psi\rangle \right) , \end{aligned} \tag{6.5}$$

and Eq. (6.3) becomes

$$\left(\hat{\mathbf{z}}^\top \cdot \mathbf{D} \cdot \hat{\mathbf{z}} \right) \left(\hat{S}^\dagger |\psi\rangle \right) = L \left(\hat{S}^\dagger |\psi\rangle \right) , \tag{6.6}$$

which can also be written as

$$\sum_{i=1}^{N} \lambda_i \left(\frac{\hat{p}_i^2 + \hat{q}_i^2}{2} \right) \left(\hat{S}^\dagger |\psi\rangle \right) = \frac{L}{2} \left(\hat{S}^\dagger |\psi\rangle \right) . \tag{6.7}$$

The solutions of the last equation are given by tensor products of number states, one for each degree of freedom,

$$|\psi\rangle = \hat{S} \left(|n_1\rangle \otimes \ldots \otimes |n_N\rangle \right) \equiv \hat{S} \left(\bigotimes_{m=1}^{N} |n_m\rangle \right) . \tag{6.8}$$

The condition

$$\frac{L}{2} = \sum_{i=1}^{N} \lambda_i \left(n_i + \frac{1}{2} \right) \hbar \tag{6.9}$$

is a condition that must be satisfied by all potential extremal states.

6.2.2 Consistency conditions

From the definition of the covariance matrix in the state $|\psi\rangle$, we obtain

$$\begin{aligned}\mathbf{C} = \langle\psi|\hat{\mathbf{C}}|\psi\rangle &\equiv \langle\psi|\left(\frac{\hat{\mathbf{z}}\otimes\hat{\mathbf{z}}+\hat{\mathbf{z}}^\top\otimes\hat{\mathbf{z}}^\top}{2}\right)|\psi\rangle \\ &= \left(\otimes_m\langle n_m|\right)\hat{S}^\dagger\left(\frac{\hat{\mathbf{z}}\otimes\hat{\mathbf{z}}+\hat{\mathbf{z}}^\top\otimes\hat{\mathbf{z}}^\top}{2}\right)\hat{S}\left(\otimes_m|n_m\rangle\right)\,,\end{aligned} \tag{6.10}$$

the C_{ij} element of which is given by

$$\begin{aligned}C_{ij} &= \left(\otimes_m\langle n_m|\right)\hat{S}^\dagger\left(\frac{\hat{z}_i\hat{z}_j+\hat{z}_j\hat{z}_i}{2}\right)\hat{S}\left(\otimes_m|n_m\rangle\right) \\ &= \frac{1}{2}\left(\otimes_m\langle n_m|\right)\left(\hat{S}^\dagger\hat{z}_i\hat{S}\hat{S}^\dagger\hat{z}_j\hat{S}+\hat{S}^\dagger\hat{z}_j\hat{S}\hat{S}^\dagger\hat{z}_i\hat{S}\right)\left(\otimes_m|n_m\rangle\right) \\ &= \frac{1}{2}\left(\otimes_m\langle n_m|\right)\sum_{k,l}S^{-1}_{ik}S^{-1}_{jl}\left(\hat{z}_i\hat{z}_j+\hat{z}_j\hat{z}_i\right)\left(\otimes_m|n_m\rangle\right) \\ &= \sum_{k,l}S^{-1}_{ik}S^{-1}_{jl}\left(\otimes_m\langle n_m|\right)\left(\frac{\hat{z}_i\hat{z}_j+\hat{z}_j\hat{z}_i}{2}\right)\left(\otimes_m|n_m\rangle\right) \\ &= \sum_{k,l}S^{-1}_{ik}S^{-1}_{jl}N_{kl}\,,\end{aligned} \tag{6.11}$$

where $\mathbf{N} = \operatorname{diag}\left(c_{n_1}, c_{n_1}, \ldots, c_{n_N}, c_{n_N}\right)$ and $c_{n_m} = \left(n_m + 1/2\right)\hbar$. Since the matrix $\mathbf{N}$ is diagonal, we find

$$C_{ij} = \sum_{k,l}S^{-1}_{ik}S^{-1}_{jl}N_{kl} = \sum_{k}S^{-1}_{ik}S^{-1}_{jk}N_{kk}\,, \tag{6.12}$$

or finally,

$$\mathbf{C} = \mathbf{S}^{-1}\mathbf{N}(\mathbf{S}^{-1})^\top\,, \tag{6.13}$$

which are the *consistency conditions* for N degrees of freedom in matrix form. In analogy with the one dimensional case, from the set of all potential extrema, they select the ones that are consistent with the specific function of the second moments considered.

The take-away message from the conditions (6.13) can be summarised in the following statement: a function f of the second moments of N positions and momenta, has an extremum in a pure state $|\psi\rangle$, if there exists a symplectic matrix S that diagonalises the corresponding covariance matrix, and at the same time, the transpose of its inverse, $\left(S^{-1}\right)^\top$, diagonalises the matrix of the partial derivatives of the function f.

Observe that the consistency conditions imply (6.9), since it holds that $L = \operatorname{Tr}(\mathbf{CF})$

and $\sum_i \lambda_i c_{n_i} = \frac{1}{2}\mathrm{Tr}(\mathbf{DN})$. Thus

$$\frac{L}{2} = \sum_{i=1}^{N} \lambda_i \left(n_i + \frac{1}{2} \right) \hbar \;\Rightarrow\; \mathrm{Tr}(\mathbf{CF}) = \mathrm{Tr}(\mathbf{DN})\,, \tag{6.14}$$

which is trivially satisfied if the consistency conditions, Eq. (6.13), are satisfied.

According to Eq. (6.13) the determinant of the covariance matrix deriving from states which extremise the uncertainty functional takes the value

$$\det \mathbf{C} = \prod_{i=1}^{N} \left(n_i + \frac{1}{2} \right)^2 \hbar^2\,. \tag{6.15}$$

Clearly, the minimum is achieved when each oscillator resides in its ground state,

$$\det \mathbf{C} \geq \left(\frac{\hbar}{2} \right)^{2N}\,, \tag{6.16}$$

corresponding to $n_1 = \ldots = n_N = 0$ in Eq. (6.15). No pure N-particle state can give rise to a covariance matrix $\mathbf{C}$ violating the inequality (6.16). This universally valid constraint generalizes the single-particle inequality derived by Robertson and Schrödinger to N particles, expressing it elegantly as a condition on the determinant of the covariance matrix of a state. Supplying (6.15) with the one dimensional Robertson-Schrödinger inequalities that need to be obeyed in addition by each subsystem, we get the general uncertainty statement for more than one degree of freedom. This is usually expressed in the form [67],

$$\mathbf{C} + \frac{i}{2}\mathbf{\Omega} \geq 0\,, \tag{6.17}$$

where $\mathbf{\Omega}$ is a fixed non-singular, skew-symmetric matrix that any symplectic matrix must preserve. Its explicit form depends on the choice of ordering of the elements in the vector $\mathbf{z}$, defined through $[\hat{z}_k, \hat{z}_l] = i\hbar\omega_{kl}$. The explicit form was given in Eq. (2.24) of Chapter 2.

6.2.3 Convexity

Let us now show that the region defined by (6.16) is a convex set in the $N(2N+1)$-dimensional space of second moments. We first show that any convex combination

$$\mathbf{C}(t) = t\mathbf{C}_1 + (1-t)\mathbf{C}_2\,, \qquad t \in [0,1]\,, \tag{6.18}$$

of two covariance matrices $\mathbf{C}_1$ and $\mathbf{C}_2$ located on the boundary,

$$\det \mathbf{C}_1 = \det \mathbf{C}_2 = \left(\frac{\hbar}{2}\right)^{2N}\,, \tag{6.19}$$

satisfies the inequality (6.16). This is consequence of the fact that

$$g(\mathbf{C}) = -\ln \det \mathbf{C} \tag{6.20}$$

is a convex function, i.e.

$$g(t\mathbf{C}_1 + (1-t)\mathbf{C}_2) \leq t g(\mathbf{C}_1) + (1-t) g(\mathbf{C}_2) \tag{6.21}$$

holds for arbitrary for strictly positive definite matrices, $\mathbf{C}_1, \mathbf{C}_2 > 0$. Rewriting (6.19) in the form

$$-\ln \det (\mathbf{C}_1/\hbar) = -\ln \det (\mathbf{C}_2/\hbar) = 2N \ln 2\,, \tag{6.22}$$

one finds that

$$-\ln \det \left[(t\mathbf{C}_1 + (1-t)\mathbf{C}_2)\,/\hbar\right] \leq -t \ln \det (\mathbf{C}_1/\hbar) - (1-t) \ln \det (\mathbf{C}_2/\hbar) = 2N \ln 2\,. \tag{6.23}$$

In other words, we have shown that

$$\det (t\mathbf{C}_1 + (1-t)\mathbf{C}_2) \geq \left(\frac{\hbar}{2}\right)^{2N}\,; \tag{6.24}$$

all convex combinations of covariance matrices lie inside the uncertainty region. Equality in (6.24) is only achieved if $t = 0$ or $t = 1$. Therefore, states of minimal uncertainty must be pure and necessary lie on the surface bounding the uncertainty region. Clearly, the argument extends to convex combinations of covariance matrices inside the uncertainty region.

6.2.4 Mixed states

In analogy with the one-dimensional case, as discussed in Sec. 5.2.5, it can be shown that the extremal states of any functional, Eq. (6.8), cover all points in the uncertainty region.

To see this, note that any admissible covariance matrix can be diagonalised, which is guaranteed by Williamson's theorem [84, 8]. Let $s_1, \ldots, s_N$ denote its symplectic eigenvalues defined in Chapter 2, assumed to be ordered from smallest to largest and $s_N \leq m + 1/2$, for some $m \geq 2$.

Consider the state

$$|\psi_k\rangle = \sqrt{t_k}|n_k = 0\rangle + \sqrt{1 - t_k}|n_k = m\rangle\,, \quad k = 1, \ldots, N\,, \tag{6.25}$$

and observe that for suitable values of t_k, the variances of position and momentum of the k-th degree of freedom can be made equal to s_k, and their covariance is zero. In addition, note that all matrix elements of operators of the form $\hat{p}_k \hat{q}_{k'}$ with $k \neq k'$ are zero in the product state

$$|\Psi\rangle = |\psi_1\rangle \otimes \ldots \otimes |\psi_N\rangle\,, \tag{6.26}$$

and thus all possible symplectic eigenvalues can be covered.

This suffices to demonstrate that mixed states do not contribute additional points in the uncertainty region, similarly to the one-dimensional case, which justifies our restriction to pure states only.

6.3 Applications and examples

6.3.1 One spatial degree of freedom

As a consistency check, we re-derive the result of Chapter 5 that was obtained for one spatial degree of freedom. In that case, $\mathbf{N} = c_n \mathbf{I}$, and with $\mathbf{S} = \mathbf{S}_\gamma \mathbf{G}_b$, one finds that the consistency conditions take the simple form

$$\mathbf{C} = \mathbf{S}^{-1}\mathbf{N}(\mathbf{S}^{-1})^{\top} = \mathbf{S}^{-1}(\mathbf{S}^{-1})^{\top}c_n = c_n\mathbf{G}_b^{-1}\mathbf{S}_\gamma^{-1}(\mathbf{S}_\gamma^{-1})^{\top}(\mathbf{G}_b^{-1})^{\top} = |\mathbf{F}|^{1/2}\mathbf{F}^{-1}c_n\,, \tag{6.27}$$

or

$$\frac{\mathbf{F}\cdot\mathbf{C}}{\sqrt{\det\mathbf{F}}} = c_n\,\mathbf{I}\,, \qquad n\in\mathbb{N}_0\,, \tag{6.28}$$

where

$$\mathbf{G}_b = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}, \quad \text{and} \quad \mathbf{S}_\gamma = \begin{pmatrix} e^{-\gamma} & 0 \\ 0 & e^{\gamma} \end{pmatrix}, \tag{6.29}$$

with real parameters

$$b = \frac{f_w}{2f_y} \in \mathbb{R} \quad \text{and} \quad \gamma = \frac{1}{2}\ln\left(\frac{f_y}{\sqrt{\det\mathbf{F}}}\right) \in \mathbb{R}\,. \tag{6.30}$$

Thus, the general formalism correctly reproduces the findings of Chapter 5, as consistency would require.

6.3.2 N spatial degrees of freedom, product states

Let us now examine the case of more degrees of freedom but when there are no correlations between them; then, the functional depends only on the second moments for each degree of freedom, $f \equiv f(x_1, y_1, w_1, \ldots, x_N, y_N, w_N)$, i.e. the $2N(N-1)$ second moments that mix the degrees of freedom are zero: $w_{N+1} = \ldots = w_{2N^2-N} = 0$. Specialising for simplicity to $N = 2$, one can show that

$$\hat{S} = (\hat{S}_1 \otimes \mathbb{1})(\mathbb{1} \otimes \hat{S}_2), \tag{6.31}$$

of which the corresponding symplectic matrices are given in block form by:

$$\mathbf{S}_1 = \begin{pmatrix} \mathbf{S}_{\gamma_1}\mathbf{G}_{b_1} & 0 \\ 0 & \mathbf{I} \end{pmatrix} \text{ and } \mathbf{S}_2 = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{S}_{\gamma_2}\mathbf{G}_{b_2} \end{pmatrix} \tag{6.32}$$

and the consistency conditions become

$$\mathbf{C} = \mathbf{S}^{-1}\mathbf{N}(\mathbf{S}^{-1})^{\top} = \mathbf{S}^{-1}(\mathbf{S}^{-1})^{\top}\mathbf{N} = \mathbf{S}^{-1}(\mathbf{S}^{-1})^{\top}\mathbf{N} = \mathbf{F}_p^{-1}\mathbf{N}\,, \tag{6.33}$$

or

$$\mathbf{F}_p\mathbf{C} = \mathbf{N}\,, \tag{6.34}$$

where we defined the matrix

$$\mathbf{F}_p = \begin{pmatrix} |\mathbf{F}_1|^{-1/2}\mathbf{F}_1 & 0 \\ 0 & |\mathbf{F}_2|^{-1/2}\mathbf{F}_2 \end{pmatrix}, \tag{6.35}$$

which shows that this case is equivalent to two one-dimensional ones separately, as it was expected. To see this, note that (6.34) is equivalent to the six equations obtained by taking the set of Eqs. (5.41-5.43) of Chapter 5 twice, once for each degree of freedom.

6.3.2.1 Inequalities for a two dimensional quantum system

Apart from inequalities that can be obtained form the product and sum of one degree of freedom inequalities, one can consider others than are not derived from the one dimensional ones in a trivial way. An example of the former is the Robertson inequality for more than two observables [57], which we can derive by considering

$$f(x_1, y_1, z_1, x_2, y_2, z_2) = (x_1 y_1 - z_1^2)(x_2 y_2 - z_2^2)\,, \tag{6.36}$$

that leads to

$$\left(\Delta^2 p_1\, \Delta^2 q_1 - C^2_{p_1 q_2}\right)\left(\Delta^2 p_2\, \Delta^2 q_2 - C^2_{p_2 q_2}\right) \geq \left(\frac{\hbar}{2}\right)^4, \tag{6.37}$$

the boundary described by Eq. (6.16) in the absence of correlations. Note that this inequality is only invariant under $Sp(2,\mathbb{R}) \otimes Sp(2,\mathbb{R})$ transformations, instead of the full $Sp(4,\mathbb{R})$ that the Robertson-Schrödinger equivalent inequality for more degrees of freedom must be invariant under. However, the matrix inequality

$$\mathbf{C} + \frac{i}{2}\mathbf{\Omega} \geq 0\,, \tag{6.38}$$

is invariant under any symplectic transformation and serves as the required generalisation. It can also be stated as inequality conditions that need to be imposed on the "symplectic eigenvalues" of the covariance matrix [1, 2].

On the other hand, an example that is not trivially obtained from one dimensional inequalities is found by considering the functional

$$f(x_1, y_1, w_1, x_2, y_2, w_2) = x_1\, y_1 x_2\, y_2 - w_1^2\, w_2^2\,, \tag{6.39}$$

and solving Eqs. (6.34) leads to the inequality

$$\left(\Delta^2 p_1\, \Delta^2 q_1\right)\left(\Delta^2 p_2\, \Delta^2 q_2\right) \geq \left(\frac{\hbar}{2}\right)^4 + C_{p_1 q_1}^2 C_{p_2 q_2}^2\,. \tag{6.40}$$

As expected, it is stronger than the "Heisenberg"-type inequality for more than two observables mentioned in the paper by Robertson [57]:

$$\Delta p_1\, \Delta q_1\, \Delta p_2\, \Delta q_2 \geq \left(\frac{\hbar}{2}\right)^2\,, \tag{6.41}$$

but weaker than Eq. (6.37).

Other inequalities that one can derive by considering appropriate functionals are:

$$\begin{aligned}
&\alpha\left(\Delta^2 p_1\right)^n \left(\Delta^2 q_2\right)^n + \beta\left(\Delta^2 p_2\right)^n \left(\Delta^2 q_1\right)^n \geq 2\sqrt{\alpha\beta}\left(\frac{\hbar}{2}\right)^{2n}\,,\\
&\alpha\left(\Delta^2 p_1\right)^n \left(\Delta^2 p_2\right)^n + \beta\left(\Delta^2 q_1\right)^n \left(\Delta^2 q_2\right)^n \geq 2\sqrt{\alpha\beta}\left(\frac{\hbar}{2}\right)^{2n}\,,
\end{aligned} \tag{6.42}$$

with $\alpha, \beta > 0$. For $\alpha = \beta = 1$ and $n = 1$, one obtains

$$\begin{aligned}
&\Delta p_1 \Delta q_2 + \Delta p_2 \Delta q_1 \geq \hbar\,,\\
&\Delta p_1 \Delta p_2 + \Delta q_1 \Delta q_2 \geq \hbar\,,
\end{aligned} \tag{6.43}$$

which can be compared with the inequality for the sum of Heisenberg inequalities for each degree of freedom:

$$\Delta p_1 \Delta q_1 + \Delta p_2 \Delta q_2 \geq \hbar\,. \tag{6.44}$$

We conclude this section by mentioning two more inequalities which will be of some importance in the context of Sec. 6.4.1:

$$\frac{1}{2}\left(\Delta^2 p_1 \Delta^2 q_2 + \Delta^2 p_2 \Delta^2 q_1\right) - C_{p_1 q_1} C_{p_2 q_2} \geq \left(\frac{\hbar}{2}\right)^2 , \tag{6.45}$$

and

$$\frac{1}{2}\left(\Delta^2 p_1 \Delta^2 p_2 + \Delta^2 q_1 \Delta^2 q_2\right) - C_{p_1 q_1} C_{p_2 q_2} \geq \left(\frac{\hbar}{2}\right)^2 . \tag{6.46}$$

6.3.3 An inequality including correlations in two spatial degrees of freedom

Let us now study a less trivial function that involves terms mixing the degrees of freedom. We start from the functional

$$f\left(\Delta^2 p_1, \ldots, C_{q_1 q_2}\right) = a\Delta^2 p_1 + a\Delta^2 q_1 + b\Delta^2 p_2 + b\Delta^2 q_2 + cC_{p_1 p_2} - cC_{q_1 q_2} ,$$

for which the matrix $\mathbf{F}$ takes the form

$$\mathbf{F} = \begin{pmatrix} a & 0 & c/2 & 0 \\ 0 & a & 0 & -c/2 \\ c/2 & 0 & b & 0 \\ 0 & -c/2 & 0 & b \end{pmatrix} . \tag{6.47}$$

It is positive definite if the coefficients a, b, c obey the conditions $a, b > 0$ and $4ab > c^2$, which we assume. The symplectic matrix $\mathbf{S}$ that puts $\mathbf{F}$ in diagonal form is given by [76]:

$$\mathbf{S} = \begin{pmatrix} w_+ & 0 & w_- & 0 \\ 0 & w_+ & 0 & -w_- \\ w_- & 0 & w_+ & 0 \\ 0 & -w_- & 0 & w_+ \end{pmatrix} , \tag{6.48}$$

where

$$w_\pm = \sqrt{\frac{a+b\pm\sqrt{y}}{2\sqrt{y}}} , \quad \text{with} \quad y = (a+b)^2 - c^2 . \tag{6.49}$$

With these definitions, the consistency conditions, Eq. (6.13), can be solved and lead to the covariance matrix at the extrema

$$\begin{aligned}\mathbf{C} &= \mathbf{S}^{-1}\mathbf{N}(\mathbf{S}^{-1})^{\top} \\ &= \begin{pmatrix} \Delta^2 p_1^{(e)} & 0 & C_{p_1p_2}^{(e)} & 0 \\ 0 & \Delta^2 q_1^{(e)} & 0 & C_{q_1q_2}^{(e)} \\ C_{p_1p_2}^{(e)} & 0 & \Delta^2 p_2^{(e)} & 0 \\ 0 & C_{q_1q_2}^{(e)} & 0 & \Delta^2 q_2^{(e)} \end{pmatrix},\end{aligned} \tag{6.50}$$

where the second moments at the extrema are given by

$$\begin{aligned}\Delta^2 p_1^{(e)} = \Delta^2 q_1^{(e)} &= \frac{n_1 - n_2}{2} + \frac{(a+b)(n_1+n_2+1)}{2\sqrt{(a+b)^2 - c^2}} \\ \Delta^2 p_2^{(e)} = \Delta^2 q_2^{(e)} &= \frac{n_2 - n_1}{2} + \frac{(a+b)(n_1+n_2+1)}{2\sqrt{(a+b)^2 - c^2}},\end{aligned} \tag{6.51}$$

while the correlations between the two degrees of freedom are

$$C_{p_1p_2}^{(e)} = -C_{q_1q_2}^{(e)} = -\frac{c(n_1+n_2+1)}{2\sqrt{(a+b)^2-c^2}}\,. \tag{6.52}$$

One can check that $\Delta^2 p_1^{(e)}$, $\Delta^2 q_1^{(e)}$, $\Delta^2 p_2^{(e)}$, $\Delta^2 q_2^{(e)} > 0$, while $\left(C_{p_1p_2}^{(e)}\right)^2 \leq \Delta^2 p_1^{(e)}\,\Delta^2 p_2^{(e)}$ and $\left(C_{q_1q_2}^{(e)}\right)^2 \leq \Delta^2 q_1^{(e)}\,\Delta^2 q_2^{(e)}$, as required. In fact, the last two inequalities are never saturated by the extremal states but they get arbitrarily close as one of n_1, n_2 is zero, while the other tends to infinity.

Substituting the values of the moments from the solutions of the consistency conditions in the functional, we find its extremal values

$$f_{a,b,c}^{(e)}(n_1,n_2) = (a-b)(n_1-n_2) + \sqrt{(a+b)^2-c^2}(n_1+n_2+1) \geq f_{a,b,c}(0,0)\,. \tag{6.53}$$

As a result, we can deduce the inequality in an arbitrary state:

$$a\Delta^2 p_1 + a\Delta^2 q_1 + b\Delta^2 p_2 + b\Delta^2 q_2 + cC_{p_1p_2} - cC_{q_1q_2} \geq \sqrt{(a+b)^2-c^2}\,, \tag{6.54}$$

which can be contrasted with one for product states

$$a\Delta^2 p_1 + a\Delta^2 q_1 + b\Delta^2 p_2 + b\Delta^2 q_2 \geq a + b\,. \tag{6.55}$$

Note that in the limit $a = b = c/2$, which however breaks the positive definiteness of $\mathbf{F}$, the right hand side of (6.54) tends to zero and the terms on the left are just the sum of the variances of the *EPR-type* pair of operators $\hat{u}_1 = \hat{p}_1 + \hat{p}_2$ and $\hat{u}_2 = \hat{q}_1 - \hat{q}_2$, [66, 8].

6.4 Uncertainty in multiple states and degrees of freedom

6.4.1 Uncertainty functionals in more than one states

Every uncertainty functional we defined so far, in this chapter and previous ones, was assumed to depend only on one state. However, it is a straightforward generalisation to allow for more than one states in a functional and look for its extrema. In this section, we briefly study the case of more states for one or more degrees of freedom in the absence of correlations. For the one dimensional case, we re-derive an inequality by Trifonov [74], named *state-extended* uncertainty relation, along with generalisations.

The extremisation of such a functional leads to an eigenvalue equation similar to Eq. (6.2), for each state, the solution of which leads to a number of consistency conditions of the type of Eq. (6.13).

Let us first examine the one spatial degree of freedom case. The functional depends on the m states φ_i with $i = 1, .., m$ and its explicit form is:

$$\begin{aligned} J[\varphi_1, \ldots, \varphi_m] = f\left(\Delta^2 p_1, \Delta^2 q_1, C_{p_1 q_1}, \ldots, \Delta^2 p_m, \Delta^2 q_m, C_{p_m q_m}; \varphi_1, \ldots, \varphi_m\right) \\ - \sum_i \lambda_i (\langle \varphi_i | \varphi_i \rangle - 1)\,, \end{aligned} \tag{6.56}$$

where $\Delta^2 p_i, \Delta^2 q_i, C_{p_i q_i}$ denote the variances and the covariance in the i-th state φ_i, and λ_i impose the condition of normalised states. Generalising the variational technique developed in the beginning of this chapter for a multidimensional functional and keeping up to first order terms, we find

$$J[\varphi_1+\varepsilon_1,\ldots,\varphi_m+\varepsilon_m] \approx J[\varphi_1,\ldots,\varphi_m] + \sum_i \varepsilon_i D_{\varepsilon_i} J[\varphi_1,\ldots,\varphi_m]\,, \tag{6.57}$$

where

$$D_{\varepsilon_i} = \langle \varepsilon_i | \frac{\delta}{\delta\langle\varphi_i|} + \frac{\delta}{\delta|\varphi_i\rangle}|\varepsilon_i\rangle\,, \tag{6.58}$$

denote Gâteaux derivatives. Extrema are obtained by requiring the first order terms to vanish and by doing so one obtains m eigenvalue equations of the form

$$\begin{aligned}\left(\frac{\partial f}{\partial\Delta^2 p_i}\hat{p}^2 + \frac{\partial f}{\partial\Delta^2 q_i}\hat{q}^2 + \frac{1}{2}\frac{\partial f}{\partial C_{p_iq_i}}\left(\hat{q}\hat{p}+\hat{p}\hat{q}\right)\right)|\xi_i\rangle& \\ = \left(\frac{\partial f}{\partial\Delta^2 p_i}\Delta^2 p_i + \frac{\partial f}{\partial\Delta^2 q_i}\Delta^2 q_i + \frac{\partial f}{\partial C_{p_iq_i}}C_{p_iq_i}\right)|\xi_i\rangle\,,&\end{aligned} \tag{6.59}$$

where $|\xi_i\rangle = \hat{T}^\dagger_{\alpha_i}|\varphi_i\rangle$ and $i=1,\ldots,m$. Again, notation becomes more concise if we set

$$x_i = \Delta^2 p_i\,,\quad y_i = \Delta^2 q_i\,,\quad w_i = C_{p_iq_i}\,, \tag{6.60}$$

and the previous eigenvalue equations become

$$\left(f_{x_i}\hat{p}^2 + f_{y_i}\hat{q}^2 + \frac{f_{w_i}}{2}\left(\hat{q}\hat{p}+\hat{p}\hat{q}\right)\right)|\xi_i\rangle = \left(x_i f_{x_i} + y_i f_{y_i} + w_i f_{w_i}\right)|\xi_i\rangle\,, \tag{6.61}$$

the solutions of which are given by

$$|\varphi_i\rangle = \hat{T}^{(i)}_{\alpha_i}\hat{G}^{(i)}_{b_i}\hat{S}^{(i)}_{\gamma_i}|n_i\rangle\,. \tag{6.62}$$

For each state there will be three equations of the form of Eqs. (5.41)-(5.43), for a total of $3m$ conditions:

$$\begin{aligned} x_i f_{x_i} &= y_i f_{y_i}\,,\\ 2w_i f_{y_i} &= -x_i f_{w_i}\,,\\ x_i y_i - w_i^2 &= c_{n_i}^2\hbar^2\,. \end{aligned} \tag{6.63}$$

Note that the c_{n_i} refer to the eigenvalues of a harmonic oscillator *of the same* degree of freedom, but there is no reason to demand that every solution $|\varphi\rangle$ of the eigenvalue equation occupy the same number state (suitably symplectically transformed), $|n\rangle$.

Observe that the consistency conditions in this case are effectively the same as the ones of a functional of one state but with multiple degrees of freedom and no correlation, as in Sec. 6.3.2. In matrix notation, they are given by

$$\mathbf{F}_{ms}\mathbf{C} = \mathbf{N}\,, \tag{6.64}$$

where

$$\mathbf{F}_{ms} = \begin{pmatrix} |\mathbf{F}_1|^{-1/2}\mathbf{F}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & |\mathbf{F}_m|^{-1/2}\mathbf{F}_m \end{pmatrix}, \tag{6.65}$$

$\mathbf{C}$ is the block matrix formed by the covariance matrices in each state on the diagonal, i.e. symbolically $\mathbf{C} = \mathrm{diag}(\mathbf{C}_1, \ldots, \mathbf{C}_m)$ and $\mathbf{N} = \mathrm{diag}\left(c_{n_1}, c_{n_1}, \ldots, c_{n_m}, c_{n_m}\right)$.

As a result, we see that there is a mathematical equivalence between the case of functionals of one degree of freedom and many states and the one of functionals in more degrees of freedom but no correlations. The potential extremal values of the second moments are the same for functionals that are structurally identical in both cases, but with a different interpretation of the symbols. That is not true for the extremal states, where in the first case we have m vectors in the same Hilbert space, while in the second they live in the tensor product of m Hilbert spaces.

6.4.1.1 Inequalities for two states

Following the considerations at the end of last section and the equivalence between functionals with non-correlated degrees of freedom and functionals in one dimension but many states, it is obvious that inequalities in the two cases have the same form but the symbols have different meanings: e.g. in the former case, $\Delta^2 q_1$ would denote the variance of position of one conjugate pair, while in the latter it would denote the variance of position in the first state. Thus, all inequalities of Sec. 6.3.2 hold with a different interpretation of the symbols. Owing to that fact, the inequality

$$\frac{1}{2}\left(\Delta^2 p_1\, \Delta^2 q_2 + \Delta^2 p_2\, \Delta^2 q_1\right) - C_{p_1 q_1}\, C_{p_2 q_2} \geq \left(\frac{\hbar}{2}\right)^2 , \tag{6.66}$$

is Trifonov's *state extended uncertainty relation* [74], in the context of this section. Variations and generalisations of it can be easily obtained, in a similar manner as in Sec. 6.3.2. E.g., one possible three state generalisation would be

$$\frac{1}{3}\left(\Delta^2 p_1\, \Delta^2 p_2\, \Delta^2 p_3 + \Delta^2 q_1\, \Delta^2 q_2\, \Delta^2 q_3\right) - C_{p_1 q_1}\, C_{p_2 q_2}\, C_{p_3 q_3} \geq \left(\frac{\hbar}{2}\right)^3 . \tag{6.67}$$

6.4.2 Functionals for a quantum system in multiple dimensions in more than one states

For completeness, we conclude this chapter by considering the most general case possible: functionals in one or more degrees of freedom, depending on one or more states.

Let us assume that the functional depends on a total of m states, in k degrees of freedom and denote the two variances and the covariance of a canonical pair of operators of the j-th degree of freedom in the i-th state with the symbols x_{ij}, y_{ij}, w_{ij}. For example, x_{ij} is explicitly given by:

$$x_{ij} = \Delta_i^2 p_j = \langle \varphi_i | \hat{p}_j^2 | \varphi_i \rangle - \langle \varphi_i | \hat{p}_j | \varphi_i \rangle^2 , \tag{6.68}$$

with $i = 1, \ldots, m$ and $j = 1, \ldots, k$ and similarly for the rest. Then, in the absence of correlations between different degrees of freedom, the uncertainty functional is explicitly given by

$$J[\varphi_1, \ldots, \varphi_m] = f\left(x_{11}, y_{11}, w_{11}, \ldots, x_{mk}, y_{mk}, w_{mk}; \varphi_1, \ldots, \varphi_m\right) - \sum_i \lambda_i (\langle \varphi_i | \varphi_i \rangle - 1) . \tag{6.69}$$

The extremal values of the uncertainty functional are obtained for the extremal states, which are the solutions to the m equations:

$$\sum_j \left(f_{x_{ij}} \hat{p}_j{}^2 + f_{y_{ij}} \hat{q}_j{}^2 + \frac{f_{w_{ij}}}{2} \left(\hat{q}_j \hat{p}_j + \hat{p}_j \hat{q}_j\right) \right) |\varphi_i\rangle = \sum_j \left(x_{ij} f_{x_{ij}} + y_{ij} f_{y_{ij}} + w_{ij} f_{w_{ij}} \right) |\varphi_i\rangle \,, \tag{6.70}$$

which are explicitly given by the states

$$|\varphi_i\rangle = \prod_j \left(\hat{T}^{(ij)}_{\alpha_{ij}} \hat{G}^{(ij)}_{b_{ij}} \hat{S}^{(ij)}_{\gamma_{ij}} \right) \prod_j |n_j\rangle \,. \tag{6.71}$$

For each degree of freedom and for each state there is a triple of consistency conditions, for a total of $3 \times m \times k$ conditions:

$$\begin{aligned} x_{ij} f_{x_{ij}} &= y_{ij} f_{y_{ij}} \,, \\ 2 w_{ij} f_{y_{ij}} &= -x_{ij} f_{w_{ij}} \,, \\ x_{ij} y_{ij} - w_{ij}^2 &= \left(n_{ij} + \frac{1}{2} \right)^2 \hbar^2 \,. \end{aligned} \tag{6.72}$$

If we allow for correlations, then for each one of the m states, we find a set of consistency conditions as in Eq. (6.13), i.e.

$$\mathbf{C}_i = \mathbf{S}_i^{-1} \mathbf{N}_i (\mathbf{S}_i^{-1})^\top \,, \quad i = 1, \ldots, m \,. \tag{6.73}$$

These are the consistency conditions for the most general functional one can consider and the main result of this section.

6.4.2.1 Inequalities for a two dimensional system involving two states

We conclude this chapter by mentioning a few simple examples of inequalities in more that one states and more degrees of freedom.

One of the simplest functionals of this type is

$$f = \alpha x_{12} y_{21} + \beta x_{21} y_{12} \,, \tag{6.74}$$

with $\alpha, \beta > 0$, that leads to the inequality

$$\alpha\, \Delta_1^2 p_2\, \Delta_2^2 q_1 + \beta\, \Delta_2^2 p_1\, \Delta_1^2 q_2 \geq \frac{\hbar^2}{2} \sqrt{\alpha\beta}\,, \tag{6.75}$$

which however, along with the minimising states, cannot be obtained from known inequalities in a trivial way. Another similar example is given by

$$\alpha\, \Delta_1^2 p_1\, \Delta_2^2 q_1 + \beta\, \Delta_1^2 p_2\, \Delta_2^2 q_2 + \gamma\, \Delta_2^2 p_1\, \Delta_1^2 q_1 + \delta\, \Delta_2^2 p_2\, \Delta_1^2 q_2 \geq \frac{\left(\sqrt{\alpha\gamma} + \sqrt{\beta\delta}\right)}{2} \hbar^2\,. \tag{6.76}$$

6.5 Discussion

In this chapter we extended the results of chapter 5 and considered functionals of one or more spatial degrees of freedom and states.

The extremisation of such a functional led to an eigenvalue equation quadratic in the positions and the momenta, and under certain assumptions we identified the set of solutions. The *consistency conditions* select the extrema for a given function f of the second moments, if they exist, from the set of all possible solutions. Applying these results, we derived a number of inequalities for the case of two spatial degrees of freedom.

We also found that there is a certain equivalence between the extremisation problem of functionals with one degree of freedom and multiple states, and the one of functionals in more than one degrees of freedom but without correlations. The resulting inequalities are the same with a different interpretation of the symbols. We extended the results to the most general case of functionals in multiple degrees of freedom and multiple states.

Chapter 7

The Arthurs-Kelly process for more than two observables

7.1 Introduction

In the previous chapters of this thesis we have been exclusively dealing with preparational inequalities, statements concerning limitations in preparing a quantum system in a state where certain incompatible observables attain arbitrarily precise values. These inequalities are in the same spirit as the one proven by Kennard [41] and generalise it along different directions: bounds for more observables, for other smooth functions apart from the product and sum as well as uncertainty relations for more degrees of freedom. All these inequalities do not require any reference to a measurement on the system since they only describe the intrinsic uncertainty in a quantum state. As it was mentioned in the introduction, preparational inequalities are inadequate to capture the full physical content of the uncertainty principle and Heisenberg's original ideas, one aspect of which concern limitations of the act of measurement on a quantum system. Recently error-disturbance inequalities have been derived [19, 20, 17, 52], independent of the specific model of the measuring process. In this chapter, however, we will mostly discuss measurement inequalities within the context of a proposed generalisation of the *Arthurs-Kelly* model of joint position and momentum measurement, in a similar fashion as Appleby [3, 4].

The Arthurs-Kelly model [6] is the first model describing a joint measurement of incompatible observables, or specifically position and momentum. It appeared in 1965 and since then, along with variations, it has been studied by various authors. It generalises von Neumann's model of measurement of position and describes, with a specific interaction Hamiltonian, the effect of the coupling of the measured system to a measuring apparatus. In its original form, a quantum system with one spatial degree of freedom is coupled with an apparatus consisting of two parts, the probes. Both probes are taken to be quantum mechanical systems characterised by a pair of canonical observables, i.e. operators obeying the standard commutation relations. In the form discussed by Arthurs and Kelly, the position and momentum of the system under investigation are coupled with the momenta of the probes and the interaction is assumed to be impulsive with the coupling constant large enough so that the free evolution of the individual systems can be ignored for as long as the measurement lasts. The main result of Arthurs and Kelly was an inequality for the product of the standard deviations of the probe pointer observables, the lower bound of which was shown to be twice as large as the one of the preparational Heisenberg inequality. This increase in noise or reduction in accuracy is attributed to the fact that the probes are quantum mechanical systems.

Although the original Arthurs-Kelly model and some of its subsequent generalisations refer to a quantum particle to be measured, the conclusions drawn find application in the field of quantum optics. In that context, the conjugate pair of position and momentum refer to the quadrature components of the field which are to be jointly measured, and the statistics obtained match those of heterodyne or homodyne detection [45, 88, 87, 44, 86].

In this chapter we will propose and study generalisations of the Arthurs-Kelly process so that it is not restricted to the joint measurement of position and momentum. These generalisations will allows us to investigate the impact of a joint measurement of more than two observables. We will derive *joint-measurement* inequalities and compare them with their preparational counterparts obtained in Chapter 4.

Furthermore, following the analysis of Appleby, we will use certain definitions of error and disturbance and derive a number of inequalities that generalise the notion of error and disturbance from two to three observables. These state dependent noise measures are of the Ozawa type and suffer from a number of known shortcomings. Regardless of this observation, they are still mathematically well defined and one can derive statements as a first step in generalising error-disturbance inequalities to the case of a joint measurement of three observables.

Finally, we will review the analysis of Arthurs and Goodman [7] in the case of three incompatible observables and with the aid of the definition of suitable noise operators, we will derive a general inequality that needs to be obeyed in a triple joint measurement. This model independent analysis provides a bound which agrees with our findings for the Arthurs-Kelly measurement of three canonical observables but is extended to arbitrary operators.

7.1.1 Preliminaries

There have been various closely related variants of the model in the literature but in this work we will consider generalisations of the following versions:

- the original *Arthurs-Kelly* model, characterised by an interaction Hamiltonian of the form (in suitable units),

$$\hat{H} = k(\hat{q} \otimes \hat{P}_1 \otimes \mathbb{I}_2 + \hat{p} \otimes \mathbb{I}_1 \otimes \hat{P}_2)\,, \quad k \in \mathbb{R}\,. \tag{7.1}$$

- The generalisation due to Busch [15], with interaction Hamiltonian:

$$\hat{H} = \lambda \hat{q} \otimes \hat{P}_1 \otimes \mathbb{I}_2 - \mu \hat{p} \otimes \mathbb{I}_1 \otimes \hat{Q}_2 + \frac{\lambda\mu}{2}\kappa \mathbb{I}_s \otimes \hat{P}_1 \otimes \hat{Q}_2\,, \quad \kappa\,,\lambda\,,\mu \in \mathbb{R}\,. \tag{7.2}$$

In both cases the interaction between the system being measured and the apparatus is thought to be impulsive in nature, which can be mathematically modelled by multiplying with a delta function, and that the Hamiltonians describing the evolution of each subsystem can be ignored for the duration of the measurement. The operators without an index correspond to the position and momentum of the system under investigation, while operators with indices correspond to the two probe systems of the measuring device. They act on the corresponding Hilbert spaces $\mathscr{H}_s, \mathscr{H}_1, \mathscr{H}_2$ associated with the system and the two probes respectively; with $\mathbb{I}$ we denote the identity and the appropriate subscripts indicate the corresponding Hilbert space. From now on, the tensor product symbols will be dropped for notational simplicity.

The first of the two models has been studied in some detail in the literature [6, 3, 4] and is the one that is more straightforward to generalise to more observables: one will couple the momentum of additional probes with each extra system operator. The second is more general in that it allows for the description of different couplings and that it incorporates the case of sequential measurements for special values of them; however, it is not entirely obvious how to generalise it to more than two observables and slightly more complicated to work with. A quite obvious difference is that in the first case, one is always coupling the momenta of the probes with the system's observables while in the second, one couples the position of the second probe with the momentum of the system under investigation and the momentum of the first probe with the position of the system. This, however, has not any impact on our considerations.

7.2 Joint measurement inequalities for the statistics of the probes

7.2.1 Uncorrelated probes

Our initial analysis assumes that no correlation exist between the probes and under this hypothesis, we will review the original analysis of Arthurs and Kelly and discuss extensions of it to N observables. A discussion of the effect of correlated probes is postponed until Sec. 7.2.2.

7.2.1.1 The original Arthurs-Kelly process

Let us briefly review the original model by Arthurs and Kelly before we study in more detail the generalisation for a canonical triple. We first assume that the system and probes are prepared in uncorrelated pure states and following the analysis of Appleby [3, 4], we work in the Heisenberg picture for the evolution of the system, in contrast with Arthurs and Kelly who derive the evolved wavefunction of the system in the Schrödinger picture. We will add a subscript "f" to denote the operators after the interaction and leave with no subscript the ones before.

We assume there is a Hilbert space $\mathscr{H}_s$ associated with the system to be measured and a Hilbert space $\mathscr{H}_i$ with $i = 1, 2$, for each probe of the measuring apparatus. The state space of the total system is represented by the tensor product $\mathscr{H} = \mathscr{H}_s \otimes \mathscr{H}_1 \otimes \mathscr{H}_2$.

Assume we have the system operators $\hat{p}, \hat{q}$ and the probe operators $\hat{P}_1, \hat{P}_2, \ldots$. In the Hilbert space of the total system, some of these are given by $\hat{p} \otimes \mathbb{I}_1 \otimes \mathbb{I}_2, \ldots$ and $\mathbb{I}_s \otimes \hat{P}_1 \otimes \mathbb{I}_2, \ldots$ and so on, where $\mathbb{I}_s, \mathbb{I}_i$ are the identity operators in the Hilbert spaces of the system and the probes, respectively. From now on the identities and tensor product symbols will be suppressed.

The unitary evolution associated with the impulsive interaction Hamiltonian, Eq. (7.1) , is given by $\hat{U}_2 = \exp\left(-\frac{i\hat{H}_2}{\hbar}\right)$ or explicitly

$$\hat{U}_2 = \exp\left(-\frac{i}{\hbar}(\hat{p}\hat{P}_1 + \hat{q}\hat{P}_2)\right). \tag{7.3}$$

According to the model, the readings that capture the measured values of $\hat{p}, \hat{q}$, are those of the evolved probe observables conjugate to the momenta, $\hat{P}_1$, $\hat{P}_2$, or specifically $\hat{Q}_{1f}, \hat{Q}_{2f}$, obtained by a projective measurement on the pointer observables after the interaction. In the Heisenberg picture, observables evolve according to $\hat{Q}_{jf} = \hat{U}_2^\dagger \hat{Q}_j \hat{U}_2$ and after a short Baker-Campbell-Hausdorff calculation we find

$$\begin{aligned} \hat{Q}_{1f} &= \hat{p} + \hat{Q}_1 + \frac{1}{2}\hat{P}_2 = \hat{p} + \hat{w}_1 \,, \\ \hat{Q}_{2f} &= \hat{q} + \hat{Q}_2 - \frac{1}{2}\hat{P}_1 = \hat{q} + \hat{w}_2 \,, \end{aligned} \tag{7.4}$$

where $[\hat{w}_1, \hat{w}_2] = \hbar/i$ and we have used the identity

$$e^X Y e^{-X} = Y + [X, Y] + \frac{1}{2!}[X, [X, Y]] \frac{1}{3!}[X, [X, [X, Y]]] + \ldots . \tag{7.5}$$

With no correlations, the variances are given by

$$\begin{aligned} \Delta^2 Q_{1f} &= \Delta^2 Q_1 + \Delta^2 p + \frac{\Delta^2 P_2}{4} \,, \\ \Delta^2 Q_{2f} &= \Delta^2 Q_2 + \Delta^2 q + \frac{\Delta^2 P_1}{4} \,. \end{aligned} \tag{7.6}$$

Let us first examine what happens to the sum inequality for the system,

$$\Delta^2 p + \Delta^2 q \geq \hbar\,. \tag{7.7}$$

The sum of the variances of the positions of the two probes after the interaction are given by

$$\Delta^2 Q_{1f} + \Delta^2 Q_{2f} = \Delta^2 Q_1 + \Delta^2 Q_2 + \Delta^2 p + \Delta^2 q + \frac{\Delta^2 P_1}{4} + \frac{\Delta^2 P_2}{4}. \tag{7.8}$$

To obtain the minimum, we use the framework developed in Chapter 6, specifically the case of product states, and we define the following uncertainty functional

$$f = y_1 + y_2 + x + y + \frac{x_1}{4} + \frac{x_2}{4}, \tag{7.9}$$

where we switched to the notation of Chapter 6; x, y denote the variances of momentum and position, respectively, and the indices differentiate between system and probe variances. For example, $x = \Delta^2 p$, or $x_1 = \Delta^2 P_1$ and so on. In the absence of correlations between the probes, the extrema of this uncertainty functional correspond to the consistency conditions, (6.34), which we repeat:

$$\begin{aligned}
x f_x &= y f_y \;\;, & xy &= \left(n + \frac{1}{2}\right)^2 \hbar^2 \;\;, & n &\in \mathbb{N}\,, \\
x_1 f_{x_1} &= y_1 f_{y_1}\,, & x_1 y_1 &= \left(n_1 + \frac{1}{2}\right)^2 \hbar^2\,, & n_1 &\in \mathbb{N}\,, \\
x_2 f_{x_2} &= y_2 f_{y_2}\,, & x_2 y_2 &= \left(n_2 + \frac{1}{2}\right)^2 \hbar^2\,, & n_2 &\in \mathbb{N}\,.
\end{aligned} \tag{7.10}$$

The extrema are easily found to be

$$x = y = \left(n + \frac{1}{2}\right)\hbar\,, \quad x_1 = 4y_1 = (2n_1 + 1)\,\hbar\,, \quad x_2 = 4y_2 = (2n_2 + 1)\,\hbar\,. \tag{7.11}$$

The extremal values of the functional are given by

$$f_{\mathrm{extr}} = (2n + n_1 + n_2 + 2)\,\hbar\,, \tag{7.12}$$

and the minimum is obviously attained for $n = n_1 = n_2 = 0$.

As a result, we obtain the sum Arthurs-Kelly equivalent inequality for a joint measurement of a canonical pair:

$$\Delta^2 Q_{1f} + \Delta^2 Q_{2f} \geq 2\hbar \,. \tag{7.13}$$

We find that the usual bound has doubled. A similar analysis for the product, which differs only in the definition of the functional, verifies the main result by Arthurs and Kelly [6],

$$\Delta^2 Q_{1f} \Delta^2 Q_{2f} \geq \hbar^2 \,, \tag{7.14}$$

which shows that the usual Heisenberg bound for the product is higher.

Both inequalities demonstrate the claim of the introduction that the joint measurement of position and momentum introduces extra noise, which should be attributed to the fact that the measuring apparatus is itself a quantum system with intrinsic uncertainty.

In the following section, we generalise these two inequalities for a canonical triple.

7.2.1.2 A generalisation for a joint measurement of three observables

We first generalise the original Hamiltonian for a canonical triple by adding another probe system to our measuring device; then, the total interaction Hamiltonian is given by

$$\hat{H}_3 = \hat{p}\hat{P}_1 + \hat{q}\hat{P}_2 + \hat{r}\hat{P}_3 \,, \tag{7.15}$$

while the unitary operator effecting the coupling of the system with the apparatus is given by

$$\hat{U}_3 = \exp\left(-\frac{i}{\hbar}(\hat{p}\hat{P}_1 + \hat{q}\hat{P}_2 + \hat{r}\hat{P}_3)\right) . \tag{7.16}$$

In accordance with the original model, the readings representing the values of $\hat{p}, \hat{q}, \hat{r}$ are those of the evolved probe observables $\hat{Q}_1, \hat{Q}_2, \hat{Q}_3$ after the interaction, obtained by a projective measurement on the pointer observables. After the interaction of the system with the measuring apparatus, we find

$$\hat{Q}_{1f} = \hat{p} + \hat{Q}_1 + \frac{1}{2}\left(\hat{P}_3 - \hat{P}_2\right) = \hat{p} + \hat{v}_1\,, \tag{7.17}$$

and cyclic permutation of the indices $(1,2,3)$. Note that operators $\hat{v}_1, \hat{v}_2, \hat{v}_3$ form a triple with the pairwise commutators equal to the value $i\hbar$. However, due to the fact that each operator is "composite" and that the third is not equal to minus the sum of the other two, we cannot apply the triple sum inequality, Eq. (5.79).

However, under the assumption that no correlations exist between the different probes, the variances of the pointer observables are $\hat{Q}_{1f}, \hat{Q}_{2f}, \hat{Q}_{3f}$, are

$$\begin{aligned}
\Delta^2 Q_{1f} &= \Delta^2 Q_1 + \Delta^2 p + \frac{1}{4}\left(\Delta^2 P_2 + \Delta^2 P_3\right) \\
\Delta^2 Q_{2f} &= \Delta^2 Q_2 + \Delta^2 q + \frac{1}{4}\left(\Delta^2 P_3 + \Delta^2 P_1\right) \\
\Delta^2 Q_{3f} &= \Delta^2 Q_3 + \Delta^2 r + \frac{1}{4}\left(\Delta^2 P_1 + \Delta^2 P_2\right),
\end{aligned} \tag{7.18}$$

which follow from (7.17) by taking the square and then the average with respect to an arbitrary product state, $|\Psi\rangle = |\psi\rangle \otimes |\varphi_1\rangle \otimes |\varphi_2\rangle \otimes |\varphi_3\rangle$; $|\psi\rangle$ refers to the initial state of the measured system, while $|\varphi_i\rangle$ to the states of the probes.

We first examine the sum inequality. As in the last section, we are looking for the minima of the uncertainty functional

$$f = \sqrt{3} + y_1 + y_2 + y_3 + \frac{1}{2}\left(x_1 + x_2 + x_3\right), \tag{7.19}$$

where, once again, we switched to the notation of Chapter 6. For simplicity we have assumed that the measured system resides in a state of minimal triple uncertainty where the variances of the three operators associated with the system in question, $\hat{p}, \hat{q}, \hat{r}$ take their lowest values, according to [39] and Chapter 3; we would have re-derived the minimal triple state if we had not made that assumption. The minimum of the functional is contained in the solutions of the consistency conditions, which give the extrema

$$
\begin{aligned}
&x_i = 2y_i = \sqrt{2}\left(n_i + \frac{1}{2}\right)\hbar\,, \\
&i = 1,2,3\,, \quad n_i \in \mathbb{N}_0\,,
\end{aligned}
\tag{7.20}
$$

from which one can derive the inequality

$$
\Delta^2 Q_{1f} + \Delta^2 Q_{2f} + \Delta^2 Q_{3f} \geq \left(\sqrt{3} + \frac{3}{\sqrt{2}}\right)\hbar\,. \tag{7.21}
$$

We find that the bound for the sum of the three variances has increased and it is more than twice larger. We also see that to reach this bound the probes have to be prepared in the same squeezed state each, in which the variance in position is twice that of the variance in momentum.

We finally turn to the product of the variances. Having established the sum inequality, Eq. (7.21), one can obtain the triple product inequality

$$
\Delta^2 Q_{1f} \Delta^2 Q_{2f} \Delta^2 Q_{3f} \geq \left(\frac{1}{\sqrt{3}} + \frac{1}{\sqrt{2}}\right)^3 \hbar^3. \tag{7.22}
$$

Again, preparing the probes in the same squeezed states as for the sum, we can achieve the lower bound. For the product inequality, the increase is of an order of magnitude. However, it should be noted that in both cases, the *increase per observable* is the same.

7.2.1.3 The case of N observables in phase space

In Chapter 4 we derived the sum and product preparational inequalities for N observables for one spatial degree of freedom, in terms of their commutators. The sum inequality was found to be (4.27), while the bound for the product was given by (4.29).

The bound after the interaction The model of joint measurement is a straightforward generalisation of the triple case: we couple each observable with the momentum of a probe, and as a result our measuring apparatus consists of N probe subsystems.

The unitary operator that models the measurement is explicitly given by

$$\hat{U}_N = \exp\left(-\frac{i}{\hbar}\sum_{k=o}^{N-1}\hat{r}_k\hat{P}_k\right), \tag{7.23}$$

where lower-case letters denote the operators of the measured system, while capitals correspond to the N probes.

As in the treatment of the last section, in the Heisenberg picture, the position operators of the probe after the interaction are found to be

$$\hat{U}_N^\dagger \hat{Q}_j \hat{U}_N = \hat{Q}_j + \hat{r}_j + \frac{i}{2\hbar}\sum_k \hat{P}_k\left[\hat{r}_k, \hat{r}_j\right] , \tag{7.24}$$

and since $\hat{r}_j = \alpha_j\hat{p} + \beta_j\hat{q}$, one finds that the commutators between the system operators are equal to

$$\left[\hat{r}_k, \hat{r}_j\right] = \ldots = -\left(\alpha_k\beta_j - \alpha_j\beta_k\right) i\hbar\,. \tag{7.25}$$

The variances of pointer positions in the absence of correlations are found to be

$$\Delta^2 Q_{jf} = \Delta^2 Q_j + \Delta^2 r_j + \frac{1}{4\hbar^2}\sum_k \Delta^2 P_k\left|\langle\left[\hat{r}_k, \hat{r}_j\right]\rangle\right|^2 , \tag{7.26}$$

and the bound for their sum is obtained by looking for the minimum of the expression

$$\sum_j \Delta Q_{jf}^2 = \sum_j \Delta Q_j^2 + \sum_j \Delta r_j^2 + \sum_{j,k}\Delta^2 P_k\left|\langle\left[\hat{r}_k, \hat{r}_j\right]\rangle\right|^2 \ , \tag{7.27}$$

or equivalently, of the uncertainty functional

$$f = B + \sum_k y_k + \sum_{k,j}\gamma_{kj}x_k\,, \tag{7.28}$$

where B denotes the preparational bound

$$B = \sqrt{\sum_{k>j}\left|\langle\left[\hat{r}_k, \hat{r}_j\right]\rangle\right|^2}\,, \tag{7.29}$$

and γ_{kj} are given by

$$\gamma_{kj} = \frac{\left|\langle[\hat{r}_k, \hat{r}_j]\rangle\right|^2}{4\hbar^2}\,. \tag{7.30}$$

After the solutions of consistency conditions, the extremal values of the functional are given by

$$f^{(e)} = B + 2\hbar \sum_k \left(n_k + \frac{1}{2}\right)\left(\sum_j \gamma_{kj}\right)^{1/2}, \tag{7.31}$$

and it is obvious that the minimal value is obtained when all n_k attain their lowest values $n_k = 0$. Thus, we can finally deduce the inequality

$$\sum_j \Delta^2 Q_{jf} \geq \sqrt{\sum_{k>j} \left|\langle[\hat{r}_k, \hat{r}_j]\rangle\right|^2} + \frac{1}{2}\sum_k \sqrt{\sum_j \left|\langle[\hat{r}_k, \hat{r}_j]\rangle\right|^2}\,. \tag{7.32}$$

which is the main result of this section. In terms of the coefficients of the $\hat{r}_k$ it can be rewritten as

$$\begin{aligned}\sum_j \Delta^2 Q_{jf} \geq \hbar \left(\left(\sum_i \alpha_i^2\right)\left(\sum_i \beta_i^2\right) - \left(\sum_i \alpha_i \beta_i\right)\right)^{1/2} \\ + \frac{\hbar}{2}\sum_i \left(\sum_j \left(\alpha_i\beta_j - \alpha_j\beta_i\right)^2\right)^{1/2}.\end{aligned} \tag{7.33}$$

The product version of the above inequality can be obtained by looking for the minimum of the product of variances under the constraint of the sum inequality.

N **canonical operators** We now study the inequality we derived for the case of N *canonical observables*. As we described in Chapter 4, the coefficients of such N observables define a regular polygon and the symmetry of the structure simplifies our results. We considered such structures in Chapter 4 and derived preparational inequalities for the sum and product of their variances, which we will contrast with the results of this section.

Substituting the particular expressions for the coefficients of the N canonical operators in (7.32), we obtain the inequality

$$\sum_j \Delta^2 Q_{jf} \geq \frac{N\hbar}{2\sin\left(\frac{2\pi}{N}\right)}\left(1+\sqrt{\frac{N}{2}}\right). \tag{7.34}$$

Comparing with the preparational bound, Eq.(4.5), we find that the sum of the variances is increased by a factor of $\left(1+\sqrt{\frac{N}{2}}\right)$.

Solving the KKT conditions for the product given the sum, we find the lower bound on the product of N canonical operators in a joint measurement :

$$\prod_j \Delta^2 Q_{jf} \geq \frac{\hbar^N}{2^N \sin\left(\frac{2\pi}{N}\right)^N}\left(1+\sqrt{\frac{N}{2}}\right)^N, \tag{7.35}$$

again to be contrasted with its preparational counterpart, Eq. (4.7).

Note that if we were not dealing with canonical structures but just rotated observables of unit "length" forming a regular polygon, then the difference in the expressions for the sum and product would be the absence of the factors $\sin\left(\frac{2\pi}{N}\right)$ and $\sin\left(\frac{2\pi}{N}\right)^N$ from the denominators, respectively.

7.2.2 The effect of correlations between the probes

So far we have been assuming that no correlations exist between the probes; in this section we relax this assumption and study the effect of initial correlations for a joint measurement of two and three observables. Although correlations can now exist between the subsystems of the probes, this is not the case between the system and the measuring device, i.e. the state of the total system is of the form $|\Psi\rangle = |\psi_s\rangle \otimes |\varphi_m\rangle$.

Recall that the pointer observables after the interaction are given by Eq. (7.4), with

$$[\hat{p},\hat{q}] = [\hat{w}_1,\hat{w}_2] = \frac{\hbar}{i}, \tag{7.36}$$

and the general sum inequality

$$\Delta^2 A + \Delta^2 B \geq \left|\langle[\hat{A},\hat{B}]\rangle\right|, \tag{7.37}$$

that follows from the inequality of Robertson for the product [56],

$$\Delta^2 A\, \Delta^2 B \geq \frac{\left|\langle[\hat{A},\hat{B}]\rangle\right|^2}{4}\,, \tag{7.38}$$

if one looks for the minimum of the sum given the product. Using the former, we find that the evolved pointer observables $\hat{Q}_{1f}, \hat{Q}_{2f}$ obey

$$\Delta^2 Q_{1f} + \Delta^2 Q_{2f} \geq 2\hbar\,, \tag{7.39}$$

while using the latter we obtain the bound for the product

$$\Delta^2 Q_{1f}\, \Delta^2 Q_{2f} \geq \hbar^2\,, \tag{7.40}$$

which are identical to the inequalities in the absence of correlations. Thus for the original Arthurs-Kelly process, allowing for correlations between the probes does not improve the statistics of the measurement.

Let us now compare these findings to the case of joint measurement of a canonical triple. We will use the general sum inequality for three operators that we derived in Chapter 4, Eq. (4.51), which we restate here:

$$\Delta^2 A + \Delta^2 B + \Delta^2 C \geq \frac{1}{\sqrt{3}}\left(|\langle[A,B]\rangle| + |\langle[B,C]\rangle| + |\langle[C,A]\rangle|\right)\,. \tag{7.41}$$

We will now apply this inequality to the probes after the interaction. The state of the total system is once again taken to be of the form $|\Psi\rangle = |\psi_s\rangle \otimes |\varphi_m\rangle$. Using Eq. (7.17) along with its cyclic permutations, the commutation relations of the canonical triple $(\hat{p},\hat{q},\hat{r})$ and the triple $(\hat{v}_1,\hat{v}_2,\hat{v}_3)$, and substituting in the general triple sum inequality, Eq. (4.51), we obtain the bound

$$\Delta^2 Q_{1f} + \Delta^2 Q_{2f} + \Delta^2 Q_{3f} \geq 2\sqrt{3}\,\hbar\,. \tag{7.42}$$

In contrast with the joint measurement inequality for position and momentum, we find that allowing correlations between the probes does have an impact on the lower bound for the sum inequality. Comparing the values of the right hand sides of (7.21) and (7.42), which are approximately equal to $3.85\hbar$ and $3.46\hbar$, we find that the bound in the presence of correlations can be lower by approximately 10%.

We can also deduce a product inequality from (4.51), by looking for the minimum product of the variances of three operators under the constraint of the sum. This turns out to be

$$\Delta^2 A\, \Delta^2 B\, \Delta^2 C \geq \left(\frac{|\langle [A,B] \rangle| + |\langle [B,C] \rangle| + |\langle [C,A] \rangle|}{3\sqrt{3}} \right)^3 , \tag{7.43}$$

and using the expressions for $\hat{Q}_{1f}, \hat{Q}_{2f}, \hat{Q}_{3f}$ from Eqs. (7.17), we find the triple product inequality

$$\Delta^2 Q_{1f}\, \Delta^2 Q_{2f}\, \Delta^2 Q_{3f} \geq \left(\frac{2}{\sqrt{3}}\, \hbar \right)^3 . \tag{7.44}$$

Not unexpectedly, the lower bound for the product of the variances similarly to the sum, can be lower in the presence of correlations. The numerical values give a decrease of about 27%. However, this is just an artifact of the different function used, sum versus product, and in both cases the *decrease per observable* is the same and equal to approximately 10%.

This finding suggests an important difference between the two cases considered, which we expect to hold for the joint measurement of N observables as well. To give an intuitive explanation why this must hold in general, we mention that the consistency conditions in the uncorrelated case only force minimisation of the sum of variances for each canonical pair of the same degree of freedom, ignoring all the other terms in the total sum. In the presence of correlations one minimises all the contributions at the expense of increasing the uncertainty of each pair; it is only a numerical coincidence that in the case of two observables the two bounds coincide.

Following this argument, measuring N canonical operators while allowing correlations between the probes, we conjecture the lower bound for the sum of the variances to be equal twice the preparational one, i.e. $N\hbar/\sin\left(\frac{2\pi}{N}\right)$ which should be contrasted with the bound $\left(1+\sqrt{N/2}\right)N\hbar/2\sin\left(\frac{2\pi}{N}\right)$, Eq. (7.34), that we obtained in last section. However, to our knowledge there do not exist generalisations of (4.51) to the case of N operators and as a result we can not prove this statement, which we leave as a conjecture.

To conclude this section, we found that correlations do have an impact when going from a joint measurement of two to the joint measurement of three. More specifically, the statistics of the pointer observables of our measuring apparatus allow for a lower triple sum and product bound, compared to the ones in the uncorrelated case. We expect this behaviour to extend to a joint measurement of N observables and conjecture that the preparational bound, Eq. (4.5), is doubled, in contrast with our findings for the absence of correlations, where the bound is given by Eq. (7.34). If the above observation is correct, the increase per observable (or the increase for the sum) is given by a factor

$$\kappa = \frac{\sqrt{N} - \sqrt{2}}{\sqrt{N} + \sqrt{2}}\,. \tag{7.45}$$

Since the lengths ℓ cancel, the increase is the same with the case of just rotated position or momentum ($\ell = 1$). The limit $N \to \infty$ that corresponds to probing all directions in phase space, is equal to one, which suggests that the reduction of the noise due to correlations becomes maximal.

7.2.3 The analysis of Arthurs and Goodman

In this section we generalise the analysis of Arthurs and Goodman [7] for a measurement of two operators to the case of three. This analysis does not assume a specific interaction and is thus model independent but it does rely on defining certain noise operators.

As in the Arthurs-Kelly process, the system to be measured is associated with a Hilbert space $\mathscr{H}_s$ and the measuring apparatus, with a Hilbert space $\mathscr{H}_m$. The total system is represented by the tensor product $\mathscr{H} = \mathscr{H}_s \otimes \mathscr{H}_m$.

Assume we have the system operators $\hat{A}, \hat{B}, \ldots$ and the probe operators $\hat{X}, \hat{Y}, \ldots$. In the Hilbert space of the total system these are represented by $\hat{A} \otimes \mathbb{I}_m, \ldots$ and $\mathbb{I}_s \otimes \hat{X}, \ldots$, where $\mathbb{I}_m, \mathbb{I}_s$ are the identity operators in the Hilbert spaces of the measuring apparatus and the system respectively; from now on the identities will be suppressed. Also, assume that the system plus apparatus are prepared into a state $\hat{\rho} = \hat{\rho}_s \otimes \hat{\rho}_m$, with $\mathrm{Tr}\,\hat{\rho} = 1$.

For each observable to be measured, we assume there is some process correlating it with an observable of the apparatus. A *noise operator* is then defined for each one, e.g.

$$\hat{N}_X = \hat{X} - g_X \hat{A}, \tag{7.46}$$

for the case where $\hat{X}$ is *tracking* $\hat{A}$ and g_X is some positive coupling constant, which can be identified as the *amplification gain* [7]. In addition, it is assumed that the measuring observables match on average the observables they measure, i.e. $\mathrm{Tr}(\hat{\rho}\hat{N}_X) = 0$, from which it follows that the noise operator is uncorrelated with all system observables, i.e. $\mathrm{Tr}(\hat{\rho}\hat{N}_X\hat{O}) = 0$, for all $\hat{O}$ acting on $\mathscr{H}_s$.

Assume we make a joint measurement of three observables of a system $\hat{A}, \hat{B}, \hat{C}$, tracked by $\hat{X}, \hat{Y}, \hat{Z}$, such that

$$\mathrm{Tr}(\hat{\rho}\hat{N}_X) = \mathrm{Tr}(\hat{\rho}\hat{N}_Y) = \mathrm{Tr}(\hat{\rho}\hat{N}_Z) = 0. \tag{7.47}$$

From the definitions of the noise operators one can show that the variances of the observables of the probes are given by,

$$\begin{aligned} \Delta^2 X &= g_X^2 \Delta^2 A + \Delta^2 N_X \\ \Delta^2 Y &= g_Y^2 \Delta^2 B + \Delta^2 N_Y \\ \Delta^2 Z &= g_Z^2 \Delta^2 C + \Delta^2 N_Z \end{aligned} \tag{7.48}$$

and similarly for the average commutators

$$\begin{aligned} \mathrm{Tr}(\hat{\rho}\,[\hat{N}_X, \hat{N}_Y]) &= \mathrm{Tr}(\hat{\rho}\,[\hat{X}, \hat{Y}]) - g_X g_Y\, \mathrm{Tr}(\hat{\rho}\,[\hat{A}, \hat{B}]) \\ \mathrm{Tr}(\hat{\rho}\,[\hat{N}_Y, \hat{N}_Z]) &= \mathrm{Tr}(\hat{\rho}\,[\hat{Y}, \hat{Z}]) - g_Y g_Z\, \mathrm{Tr}(\hat{\rho}\,[\hat{B}, \hat{C}]) \\ \mathrm{Tr}(\hat{\rho}\,[\hat{N}_Z, \hat{N}_X]) &= \mathrm{Tr}(\hat{\rho}\,[\hat{Z}, \hat{X}]) - g_Z g_X\, \mathrm{Tr}(\hat{\rho}\,[\hat{C}, \hat{A}]). \end{aligned} \tag{7.49}$$

Moreover, observe that $\mathrm{Tr}\left(\hat{\rho}\hat{N}_X\hat{Y}\right) = 0$ and $\mathrm{Tr}\left(\hat{\rho}\hat{N}_Y\hat{X}\right) = 0$ imply that

$$\mathrm{Tr}\,\hat{\rho}\,[\hat{X}, \hat{Y}] = 0. \tag{7.50}$$

As a result, we obtain the three conditions

$$
\begin{aligned}
|\langle[\hat{N}_X, \hat{N}_Y]\rangle| &= |g_X g_Y \langle[\hat{A}, \hat{B}]\rangle| \\
|\langle[\hat{N}_Y, \hat{N}_Z]\rangle| &= |g_Y g_Z \langle[\hat{B}, \hat{C}]\rangle| \\
|\langle[\hat{N}_Z, \hat{N}_X]\rangle| &= |g_Z g_X \langle[\hat{C}, \hat{A}]\rangle| .
\end{aligned} \tag{7.51}
$$

Applying this to the observables of the three probes, Eqs. (7.48), we find

$$
\begin{aligned}
\Delta^2 X + \Delta^2 Y + \Delta^2 Z &= \Delta^2 N_X + \Delta^2 N_Y + \Delta^2 N_Z + g_X^2 \Delta^2 A + g_Y^2 \Delta^2 B + g_Z^2 \Delta^2 C \\
&\geq \frac{1}{\sqrt{3}} \left(|\langle[N_X, N_Y]\rangle| + \ldots\right) + \frac{1}{\sqrt{3}} \left(|g_X g_Y \langle[A, B]\rangle| + \ldots\right)
\end{aligned} \tag{7.52}
$$

or by using (7.51), we finally obtain

$$
\Delta^2 X + \Delta^2 Y + \Delta^2 Z \geq \frac{2}{\sqrt{3}} \left(g_X g_Y \, |\langle[A, B]\rangle| + g_Y g_Z \, |\langle[B, C]\rangle| + g_Z g_X \, |\langle[C, A]\rangle|\right) . \tag{7.53}
$$

The case of $g_X = g_Y = g_Z = 1$, confirms the results of Sec. 7.2.2. We do not mention the product inequality but it is easily obtained from the last result by the usual optimisation method.

We conclude this section by summarising our results: similar to the analysis of Arthurs and Goodman, by defining appropriate noise operators we find a general, model-independent bound for a measurement of three observables. Whenever they form a canonical triple the minimum value is found to be double the preparational one, which agrees with our findings for the model-dependent analysis based on the generalised Arthurs-Kelly model.

7.3 Inequalities for the error and disturbance

7.3.1 Error and disturbance inequalities according to Appleby

7.3.1.1 Review of the position-momentum case

So far we had been interested in the statistics of the meter readings of the probes after the interaction and we saw that they obey a Heisenberg-type inequality with a bound higher than the preparational one; this we attributed to the fact that for each measured observable there is a probe coupled to it, which is also a quantum system with intrinsic uncertainty. We first follow a similar analysis to Appleby in [3], where he uses the root mean square of suitably defined noise operators.

With the evolution defined by (7.3), one can evaluate the effect of the interaction to all system and measuring apparatus operators through $\hat{\mathscr{O}}_f = \hat{U}_2^\dagger \hat{\mathscr{O}} \hat{U}_2$:

$$\begin{aligned} &\hat{P}_{1\mathrm{f}} = \hat{P}_1\,, \;\; \hat{Q}_{1\mathrm{f}} = \hat{Q}_1 + \hat{p} - \frac{\hat{P}_2}{2}, \\ &\hat{P}_{2\mathrm{f}} = \hat{P}_2\,, \;\; \hat{Q}_{2\mathrm{f}} = \hat{Q}_2 + \hat{q} + \frac{\hat{P}_1}{2}, \\ &\hat{p}_\mathrm{f} = \hat{p} - \hat{P}_2\,, \;\; \hat{q}_\mathrm{f} = \hat{q} + \hat{P}_1. \end{aligned} \tag{7.54}$$

One can subsequently define the error operators of retrodiction [3]

$$\hat{\epsilon}_{\mathrm{pi}} = \hat{Q}_{1\mathrm{f}} - \hat{p}\,, \quad \hat{\epsilon}_{\mathrm{qi}} = \hat{Q}_{2\mathrm{f}} - \hat{q}\,, \tag{7.55}$$

the error operators of prediction

$$\hat{\epsilon}_{\mathrm{pf}} = \hat{Q}_{1\mathrm{f}} - \hat{p}_\mathrm{f}\,, \quad \hat{\epsilon}_{\mathrm{qf}} = \hat{Q}_{2\mathrm{f}} - \hat{q}_\mathrm{f}\,, \tag{7.56}$$

and the disturbance operators

$$\hat{\delta}_\mathrm{p} = \hat{p}_\mathrm{f} - \hat{p}\,, \quad \hat{\delta}_\mathrm{q} = \hat{q}_\mathrm{f} - \hat{q}\,. \tag{7.57}$$

If one thinks in classical terms, then these definitions are intuitively justified; the errors of retrodiction are an estimate of how well the measurements capture the initial state of the system, the ones of prediction compare the measurement to the new state of the system *after* the measurement, while the disturbance operators give the difference between the system operators before and after the measurement has been performed. Of course these definitions can not be interpreted as easily in the quantum context but regardless of their "quantum" meaning, they can still be defined mathematically. As to whether they are good measures of error and disturbance, this has been debated and in many cases it appears that this is not the case. Regardless, we consider them as a first step towards inequalities of error and disturbance for a triple and we provide a number of inequalities based on them. A more detailed discussion on the topic can be found in [3, 4]. An analysis based on other measures such as the Wasserstein-2 distance [19], for example, has not been given in this thesis and will be the focus of future investigations.

One can now define the *mean-squared errors* of retrodiction (or the root-mean-squared by taking the square root):

$$\Delta^2_{\mathrm{ei}} p = \langle \Psi | \hat{\epsilon}^2_{\mathrm{pi}} | \Psi \rangle \ , \ \ \Delta^2_{\mathrm{ei}} q = \langle \Psi | \hat{\epsilon}^2_{\mathrm{qi}} | \Psi \rangle, \tag{7.58}$$

the mean-squared errors of prediction,

$$\Delta^2_{\mathrm{ef}} p = \langle \Psi | \hat{\epsilon}^2_{\mathrm{pi}} | \Psi \rangle \ , \ \ \Delta^2_{\mathbf{ef}} q = \langle \Psi | \hat{\epsilon}^2_{\mathrm{qi}} | \Psi \rangle, \tag{7.59}$$

and the mean-squared disturbances

$$\Delta^2_{\mathrm{d}} p = \langle \Psi | \hat{\delta}^2_{\mathrm{p}} | \Psi \rangle \ , \ \ \Delta^2_{\mathrm{d}} q = \langle \Psi | \hat{\delta}^2_{\mathrm{q}} | \Psi \rangle, \tag{7.60}$$

where here we assume that the initial state of the system is of the form $|\Psi\rangle = |\psi\rangle \otimes |\varphi_{\mathrm{ap}}\rangle$. Under the Arthurs-Kelly evolution, one finds that the error and disturbance operators are explicitly given by,

$$\begin{aligned}
\hat{\epsilon}_{\mathrm{pi}} &= \hat{Q}_1 - \frac{1}{2}\hat{P}_2 \ \ , \ \ \hat{\epsilon}_{\mathrm{qi}} = \hat{Q}_2 + \frac{1}{2}\hat{P}_1, \\
\hat{\epsilon}_{\mathrm{pf}} &= \hat{Q}_1 + \frac{1}{2}\hat{P}_2 \ \ , \ \ \hat{\epsilon}_{\mathrm{qf}} = \hat{Q}_2 - \frac{1}{2}\hat{P}_1, \\
\hat{\delta}_{\mathrm{p}} &= -\hat{P}_2 \ \ , \ \ \hat{\delta}_{\mathrm{q}} = \hat{P}_1.
\end{aligned} \tag{7.61}$$

One can verify that the following commutation relations hold

$$\left[\hat{\epsilon}_{\mathrm{pi}}, \hat{\epsilon}_{\mathrm{qi}}\right] = \left[\hat{\epsilon}_{\mathrm{qf}}, \hat{\epsilon}_{\mathrm{pf}}\right] = \left[\hat{\delta}_{\mathrm{p}}, \hat{\epsilon}_{\mathrm{qi}}\right] = \left[\hat{\delta}_{\mathrm{p}}, \hat{\epsilon}_{\mathrm{qf}}\right] = \left[\hat{\epsilon}_{\mathrm{pi}}, \hat{\delta}_{\mathrm{q}}\right] = \left[\hat{\epsilon}_{\mathrm{pf}}, \hat{\delta}_{\mathrm{q}}\right] = i\hbar \tag{7.62}$$

which immediately imply the retrodictive and predictive error relations

$$\Delta^2_{\mathrm{ei}} p\, \Delta^2_{\mathrm{ei}} q \geq \frac{\hbar^2}{4} \;\; , \;\; \Delta^2_{\mathrm{ef}} p\, \Delta^2_{\mathrm{ef}} q \geq \frac{\hbar^2}{4}, \tag{7.63}$$

and the four error disturbance relations

$$\begin{aligned} \Delta^2_{\mathrm{ei}} p\, \Delta^2_{\mathrm{d}} q \geq \frac{\hbar^2}{4} \;\; , \;\; \Delta^2_{\mathrm{ei}} q\, \Delta^2_{\mathrm{d}} p \geq \frac{\hbar^2}{4}, \\ \Delta^2_{\mathrm{ef}} p\, \Delta^2_{\mathrm{d}} q \geq \frac{\hbar^2}{4} \;\; , \;\; \Delta^2_{\mathrm{ef}} q\, \Delta^2_{\mathrm{ef}} p \geq \frac{\hbar^2}{4}. \end{aligned} \tag{7.64}$$

7.3.1.2 Error and disturbance for a triple joint measurement

The definitions of last subsection extend to the triple in a straightforward manner. Once again, after the unitary evolution modelling the measurement, the system and probe operators are given by,

$$\begin{aligned} \hat{Q}_{1\mathrm{f}} &= \hat{Q}_1 + \hat{p} + \frac{\left(\hat{P}_3 - \hat{P}_2\right)}{2} \; , \; \hat{Q}_{2\mathrm{f}} = \hat{Q}_2 + \hat{q} + \frac{\left(\hat{P}_1 - \hat{P}_3\right)}{2} \; , \\ \hat{Q}_{3\mathrm{f}} &= \hat{Q}_3 + \hat{r} + \frac{\left(\hat{P}_1 - \hat{P}_2\right)}{2} \; , \; \hat{P}_{1\mathrm{f}} = \hat{P}_1 \; , \; \hat{P}_{2\mathrm{f}} = \hat{P}_2 \; , \; \hat{P}_{3\mathrm{f}} = \hat{P}_3 \\ \hat{p}_{\mathrm{f}} &= \hat{p} + \hat{P}_3 - \hat{P}_2 \; , \; \hat{q}_{\mathrm{f}} = \hat{q} + \hat{P}_1 - \hat{P}_3 \; , \; \hat{r}_{\mathrm{f}} = \hat{q} + \hat{P}_2 - \hat{P}_1. \end{aligned} \tag{7.65}$$

One can then, in accordance with last subsection, define the errors of retrodiction from the three operators:

$$\begin{aligned} \hat{\epsilon}_{\mathrm{pi}} = \hat{Q}_{1\mathrm{f}} - \hat{p} = \hat{Q}_1 + \frac{\left(\hat{P}_3 - \hat{P}_2\right)}{2} \;\; , \;\; \hat{\epsilon}_{\mathrm{qi}} = \hat{Q}_{2\mathrm{f}} - \hat{q} = \hat{Q}_2 + \frac{\left(\hat{P}_1 - \hat{P}_3\right)}{2} \;\; , \\ \hat{\epsilon}_{\mathrm{ri}} = \hat{Q}_{3\mathrm{f}} - \hat{r} = \hat{Q}_3 + \frac{\left(\hat{P}_1 - \hat{P}_2\right)}{2} \;\; , \end{aligned} \tag{7.66}$$

the errors of prediction using the operators

$$\begin{aligned} \hat{\epsilon}_{\mathrm{pf}} = \hat{Q}_{1\mathrm{f}} - \hat{p}_{\mathrm{f}} = \hat{Q}_1 + \frac{\left(\hat{P}_2 - \hat{P}_3\right)}{2} \;\; , \;\; \hat{\epsilon}_{\mathrm{qf}} = \hat{Q}_{2\mathrm{f}} - \hat{q}_{\mathrm{f}} = \hat{Q}_2 + \frac{\left(\hat{P}_3 - \hat{P}_1\right)}{2} \;\; , \\ \hat{\epsilon}_{\mathrm{rf}} = \hat{Q}_{3\mathrm{f}} - \hat{q}_{\mathrm{f}} = \hat{Q}_3 + \frac{\left(\hat{P}_2 - \hat{P}_1\right)}{2} \;\; , \end{aligned} \tag{7.67}$$

and the three disturbance operators

$$\hat{\delta}_{\mathrm{p}} = \hat{p}_{\mathrm{f}} - \hat{p} = \hat{P}_3 - \hat{P}_2 \ , \quad \hat{\delta}_{\mathrm{q}} = \hat{q}_{\mathrm{f}} - \hat{q} = \hat{P}_1 - \hat{P}_3 \ , \quad \hat{\delta}_{\mathrm{r}} = \hat{r}_{\mathrm{f}} - \hat{r} = \hat{P}_2 - \hat{P}_1. \tag{7.68}$$

By observing that any two disturbance operators commute, we see that for the triple we can only have *error-error-error* (EEE) and *error-error-disturbance* (EED) relations. Regarding the former, we have the two EEE inequalities,

$$\Delta^2_{\mathrm{ei}} p \, \Delta^2_{\mathrm{ei}} q \, \Delta^2_{\mathrm{ei}} r \geq \left(\frac{\tau\hbar}{2} \right)^3 \quad , \quad \Delta^2_{\mathrm{ef}} p \, \Delta^2_{\mathrm{ef}} q \, \Delta^2_{\mathrm{ef}} r \geq \left(\frac{\tau\hbar}{2} \right)^3 \tag{7.69}$$

of retrodiction and prediction, respectively. These follow from the fact that the error operators constitute canonical triples. In order to obtain the bound we have to be cautious in that we cannot immediately apply the result in [39] since this inequality was derived under the constraint that the third operator in the triple is equal to minus the sum of the other two; we can, however, use the inequality (7.43).

Moreover, note that we also have the following triples: $(\hat{\epsilon}_{\mathrm{pi}}, \hat{\epsilon}_{\mathrm{qi}}, \hat{\delta}_r)$, $(\hat{\epsilon}_{\mathrm{pi}}, \hat{\delta}_q, \hat{\epsilon}_{\mathrm{ri}})$, $(\hat{\delta}_p, \hat{\epsilon}_{\mathrm{qi}}, \hat{\epsilon}_{\mathrm{ri}})$, $(\hat{\epsilon}_{\mathrm{pf}}, \hat{\epsilon}_{\mathrm{qf}}, \hat{\delta}_r)$, $(\hat{\epsilon}_{\mathrm{pf}}, \hat{\delta}_q, \hat{\epsilon}_{\mathrm{rf}})$, $(\hat{\delta}_p, \hat{\epsilon}_{\mathrm{qf}}, \hat{\epsilon}_{\mathrm{rf}})$, from which one can derive six EED inequalities of the form

$$\Delta^2_{\mathrm{ei}} p \, \Delta^2_{\mathrm{ei}} q \, \Delta^2_{\mathrm{d}} r \geq \left(\frac{\tau\hbar}{2} \right)^3 \tag{7.70}$$

for each triple. Notice that not all triples can be ordered into a canonical form so that they obtain cyclic symmetry; nevertheless this has no impact on the derived inequalities.

Inequality (7.70), relates the error of the position and momentum measurement with the disturbance caused on the third operator $\hat{r}$; identical conclusions follow for the other five cases. All these relations should be considered as a generalisation of the error-disturbance relations for a canonical pair. Finally, note that all the error-error and error-disturbance relations that were derived for a pair, hold in the triple case as well.

7.4 Extending the model of Busch

In this section we review some of the previous results of this chapter to allow for different coupling strengths and to incorporate the case of sequential measurements, which can be thought of as a type of joint measurement. The generalised version that we will study is due to Busch [15] and it is described by the unitary operator,

$$\begin{aligned}\hat{U}_2 &= e^{-\frac{i}{\hbar}\left(\alpha\hat{p}\hat{P}_1+\beta\hat{q}\hat{P}_2+\kappa\frac{\alpha\beta}{2}\hat{P}_1\hat{P}_2\right)} = e^{-\frac{i}{\hbar}\alpha\hat{p}\hat{P}_1}e^{-\frac{i}{\hbar}\beta\hat{q}\hat{P}_2}e^{-\frac{i}{\hbar}\frac{\alpha\beta}{2}(\kappa+1)\hat{P}_1\hat{P}_2} \\ &= e^{-\frac{i}{\hbar}\beta\hat{q}\hat{P}_2}e^{-\frac{i}{\hbar}\alpha\hat{p}\hat{P}_1}e^{-i\frac{\alpha\beta}{2}(\kappa-1)\hat{P}_1\hat{P}_2}\,, \end{aligned} \tag{7.71}$$

with $\alpha, \beta \geq 0$ and $\kappa \in \mathbb{R}$. In the original [15], the unitary considered had the momentum of the system coupled to the position of the probe but in our considerations we will always couple observables of the system with the momenta of the probes. Working in units where $[p] = [q] = [\sqrt{\hbar}]$, this does not make any difference for the results.

Whenever $|\kappa| = 1$ we are dealing with a strictly sequential measurement, otherwise we have a joint measurement of position and momentum with different coupling strengths. Note that whenever the order of the exponentials of position and momentum is changed, there follows a change of sign in the exponential of $\hat{P}_1\hat{P}_2$, that is, $\kappa + 1$ goes to $\kappa - 1$; this observation will be useful for the generalisations to more than two observables.

This model has been studied in detail, in the context of error-disturbance uncertainty relations and it has been shown that Heisenberg's error disturbance uncertainty relation is obeyed for definitions of error and disturbance related to the Wasserstein distance.

To incorporate the measurement of a canonical triple this can be generalised to

$$\begin{aligned}\hat{U}_3 &= e^{-\frac{i}{\hbar}\left(\alpha\hat{p}\hat{P}_1+\beta\hat{q}\hat{P}_2+\gamma\hat{r}\hat{P}_3+\kappa\frac{\alpha\beta}{2}\hat{P}_1\hat{P}_2+\lambda\frac{\beta\gamma}{2}\hat{P}_2\hat{P}_3+\mu\frac{\gamma\alpha}{2}\hat{P}_3\hat{P}_1\right)} \\ &= e^{-\frac{i}{\hbar}\alpha\hat{p}\hat{P}_1}e^{-\frac{i}{\hbar}\beta\hat{q}\hat{P}_2}e^{-\frac{i}{\hbar}\gamma\hat{r}\hat{P}_3}e^{\frac{i}{\hbar}\left(\frac{\alpha\beta}{2}(\kappa+1)\hat{P}_1\hat{P}_2+\frac{\beta\gamma}{2}(\lambda+1)\hat{P}_2\hat{P}_3+\frac{\gamma\alpha}{2}(\mu-1)\hat{P}_3\hat{P}_1\right)}\,, \end{aligned} \tag{7.72}$$

with the positive coupling strengths $\alpha, \beta, \gamma \geq 0$ and the real numbers $\kappa, \lambda, \mu \in \mathbb{R}$ that distinguish the different types of measurements. We call the *canonical order*, the decomposition where the exponentials of $\hat{p}, \hat{q}, \hat{r}$ appear in this order, as shown in the second line of Eq. (7.72). Different decompositions can be obtained from the canonical order by flipping signs according to the following simple rule: for each exchange of two terms in the canonical order, there is a flip of the ± 1, in the corresponding term of the last exponential. For example, the decomposition where the order of the first three exponentials is in equivalence with $\hat{r}, \hat{q}, \hat{p}$, is equal to

$$\hat{U}_3 = e^{\frac{i}{\hbar}\gamma\hat{r}\hat{P}_3} e^{\frac{i}{\hbar}\beta\hat{q}\hat{P}_2} e^{\frac{i}{\hbar}\alpha\hat{p}\hat{P}_1} e^{-\frac{i}{\hbar}\left(\frac{\alpha\beta}{2}(\kappa-1)\hat{P}_1\hat{P}_2 + \frac{\beta\gamma}{2}(\lambda+1)\hat{P}_2\hat{P}_3 + \frac{\gamma\alpha}{2}(\mu-1)\hat{P}_3\hat{P}_1\right)}, \tag{7.73}$$

since the sign in the $\hat{P}_3\hat{P}_1$ term flips twice, while the one of $\hat{P}_1\hat{P}_2$ once.

Again, we are dealing with a strictly sequential measurement whenever $|\kappa| = |\lambda| = |\mu| = 1$. For the triple, in contrast with the two observables case, there are more interesting possibilities that appear: if only one of the κ, λ, μ is not equal to plus or minus one, λ say, then this constitutes a sequential measurement of the system's position followed or preceded by a joint measurement of the momentum and the third observable in the canonical triple, $\hat{r}$; the case when only one of the couplings is plus or minus one, can be interpreted as an in between case, something like a joint measurement of the position and momentum of the quantum system, followed or preceded by a joint measurement of momentum and $\hat{r}$.

7.4.1 Joint measurement inequalities

Joint measurement of a canonical pair Let us first write down joint measurement inequalities in this more general setting in the case of a joint measurement of position and momentum and then generalise to a triple. Observe that we can rewrite the unitary evolution operator as

$$\hat{U}_2 = e^{-\frac{i}{\hbar}\left(\alpha\hat{p}\hat{P}_1 + \beta\hat{q}\hat{P}_2 + \kappa\frac{\alpha\beta}{2}\hat{P}_1\hat{P}_2\right)} = e^{-\frac{i}{\hbar}\left(\hat{r}_1\hat{P}_1 + \hat{r}_2\hat{P}_2\right)} e^{-i\kappa\frac{\alpha\beta}{2\hbar}\hat{P}_1\hat{P}_2} \tag{7.74}$$

and thus the commutators become

$$\begin{aligned}\hat{Q}_{1f} &= \alpha\hat{p} + \hat{Q}_1 + \frac{\alpha\beta}{2}(\kappa - 1)\,\hat{P}_2\,,\\ \hat{Q}_{2f} &= \beta\hat{q} + \hat{Q}_2 + \frac{\alpha\beta}{2}(\kappa + 1)\,\hat{P}_1\,,\end{aligned} \tag{7.75}$$

whose variances in the absence of correlations are equal to

$$\begin{aligned}\Delta^2 Q_{1f} &= \alpha^2\Delta^2 p + \Delta^2 Q_1 + \frac{\alpha^2\beta^2}{4}(\kappa - 1)^2\,\Delta^2 P_2\,,\\ \Delta^2 Q_{2f} &= \beta^2\Delta^2 q + \Delta^2 Q_2 + \frac{\alpha^2\beta^2}{4}(\kappa + 1)^2\,\Delta^2 P_1\,.\end{aligned} \tag{7.76}$$

A bound for their sum can be obtained by defining the functional

$$f = y_1 + \alpha^2 x + \frac{\alpha^2\beta^2}{4}(\kappa - 1)^2\,x_2 + y_2 + \beta^2 y + \frac{\alpha^2\beta^2}{4}(\kappa + 1)^2\,x_1. \tag{7.77}$$

At the extrema, we find

$$\begin{aligned}f^{(e)} &= \alpha\beta|\kappa + 1|\left(n_1 + \frac{1}{2}\right)\hbar + \alpha\beta|\kappa - 1|\left(n_2 + \frac{1}{2}\right)\hbar + 2\alpha\beta\left(n + \frac{1}{2}\right)\hbar\\ &\geq \frac{\alpha\beta}{2}\left(|\kappa + 1| + |\kappa - 1| + 2\right) \geq 2\alpha\beta\hbar\,,\end{aligned} \tag{7.78}$$

which is obeyed for probes prepared in non-entangled states. A straightforward calculation, however, using Eq. (7.37), gives the same result for arbitrary states of the probes. Observe that the last bound, $2\alpha\beta\hbar$, is achieved for any $\kappa \in [-1, 1]$, while it grows rapidly with κ for values outside that range. Supported by that observation, it seems a reasonable physical restriction to only allow κ to take values between minus and plus one, which covers all possibilities: the two cases of sequential measurement, with the order determined by the sign of ± 1, a joint measurement for $\kappa = 0$ and in between cases for the remaining values in $(-1, 1)$. Regardless, we will allow κ to be an arbitrary real number.

Joint measurement of a canonical triple Let us know re-examine these results for the case of a triple joint-measurement, as the one effected by the unitary given in

(7.72). The time evolved observables after the interaction, are found to be

$$\hat{Q}_{1f} = \alpha\hat{p} + \hat{Q}_1 + \frac{\alpha\beta}{2}(\kappa-1)\,\hat{P}_2 + \frac{\gamma\alpha}{2}(\mu+1)\,\hat{P}_3\,,$$
$$\hat{Q}_{2f} = \beta\hat{q} + \hat{Q}_2 + \frac{\beta\gamma}{2}(\lambda-1)\,\hat{P}_3 + \frac{\alpha\beta}{2}(\kappa+1)\,\hat{P}_1\,,$$
$$\hat{Q}_{3f} = \gamma\hat{r} + \hat{Q}_3 + \frac{\alpha\gamma}{2}(\mu-1)\,\hat{P}_1 + \frac{\beta\gamma}{2}(\lambda+1)\,\hat{P}_2\,, \tag{7.79}$$

with variances equal to

$$\Delta^2 Q_{1f} = \alpha^2\Delta^2 p + \Delta^2 Q_1 + \frac{\alpha^2\beta^2}{4}(\kappa-1)^2\,\Delta^2 P_2 + \frac{\gamma^2\alpha^2}{4}(\mu+1)^2\,\Delta^2 P_3\,,$$
$$\Delta^2 Q_{2f} = \beta^2\Delta^2 q + \Delta^2 Q_2 + \frac{\beta^2\gamma^2}{4}(\lambda-1)^2\,\Delta^2 P_3 + \frac{\alpha^2\beta^2}{4}(\kappa+1)^2\,\Delta^2 P_1\,,$$
$$\Delta^2 Q_{3f} = \gamma^2\Delta^2 r + \Delta^2 Q_3 + \frac{\alpha^2\gamma^2}{4}(\mu-1)^2\,\Delta^2 P_1 + \frac{\beta^2\gamma^2}{4}(\lambda+1)^2\,\Delta^2 P_2\,, \tag{7.80}$$

which hold for uncorrelated states of the probes. A bound for their sum is easily found to be

$$\begin{aligned}\Delta^2 Q_{1f} + \Delta^2 Q_{2f} + \Delta^2 Q_{3f} &\geq \hbar\sqrt{\alpha^2\beta^2+\beta^2\gamma^2+\gamma^2\alpha^2} + \hbar\frac{a}{2}\sqrt{\beta^2(\kappa+1)^2+\gamma^2(\mu-1)^2} \\ &\quad + \hbar\frac{\beta}{2}\sqrt{\alpha^2(\kappa-1)^2+\gamma^2(\lambda+1)^2} + \hbar\frac{\gamma}{2}\sqrt{\alpha^2(\mu+1)^2+\beta^2(\lambda-1)^2} \\ &\geq \hbar\frac{\alpha\beta}{2\sqrt{2}}\left(|\kappa+1|+|\kappa-1|+\frac{2\sqrt{2}}{\sqrt{3}}\right) \\ &\qquad + \hbar\frac{\beta\gamma}{2\sqrt{2}}\left(|\lambda+1|+|\lambda-1|+\frac{2\sqrt{2}}{\sqrt{3}}\right) \\ &\qquad\quad + \hbar\frac{\gamma\alpha}{2\sqrt{2}}\left(|\mu+1|+|\mu-1|+\frac{2\sqrt{2}}{\sqrt{3}}\right)\,,\end{aligned} \tag{7.81}$$

where we have used the inequality $\sqrt{x_1^2+\ldots+x_n^2} \geq (|x_1|+\ldots|x_n|)/\sqrt{n}$ proved in Chapter 4. In the case $|\kappa|\,,|\lambda|\,,|\mu|<1$, the last bound reduces to

$$\Delta^2 Q_{1f} + \Delta^2 Q_{2f} + \Delta^2 Q_{3f} \geq \frac{\hbar}{\sqrt{3}}(\alpha\beta+\beta\gamma+\gamma\alpha)\left(1+\sqrt{\frac{3}{2}}\right)\,. \tag{7.82}$$

For $\alpha = \beta = \gamma = 1$, this correctly reproduces inequality (7.21). Using inequality (4.51), we can compare the last result to the case of correlated probes, which turns out to be

$$\Delta^2 Q_{1f} + \Delta^2 Q_{2f} + \Delta^2 Q_{3f} \geq \frac{2\hbar}{\sqrt{3}} \left(\alpha\beta + \beta\gamma + \gamma\alpha \right) . \tag{7.83}$$

Whenever $\alpha = \beta = \gamma = 1$ this reduces to twice the preparational bound of the canonical triple $(\hat{p}, \hat{q}, \hat{r})$.

7.5 Discussion

We derived a number of joint-measurement uncertainty relations and inequalities for the error and disturbance within generalisations of the Arthurs-Kelly model.

Our results suggest that in a joint-measurement of N observables, correlated probes can lead to a lower bound for the sum and product of the variances of the pointer observables, in contrast to the case of position and momentum. We proved that in the case of three, the lower bound is twice the preparational one and we argued that this result should hold for N as well. A model-independent analysis similar to Arthurs and Goodman [7] for the joint measurement of three observables, was found to be in agreement with those within the generalised Arthurs-Kelly process.

Following Appleby [3, 5], we derived a number of inequalities that attempt to characterise the error and disturbance due to the measurement. In the case of a canonical triple, apart for the error-disturbance relations that one can formulate for each canonical pair, we find that possible relations come in two varieties: these can be expressed as error-error-error and error-error-disturbance inequalities.

Chapter 8

Conclusions

The investigations reported in this thesis can be separated into two main categories: the first, longer part, was on *preparational uncertainty*, while the second was about the indeterminacy of *measurement*. Both express the different type of limitations imposed by incompatible properties of one or more particles, according to quantum theory. They are different aspects of the *uncertainty principle*, the former dealing with intrinsic uncertainty of a system and the latter with the uncertainty present in the act of measurement.

In Chapter 3 we considered a triple of canonical observables for a quantum particle, Eq. (3.3), unique up to unitary transformations and we proved a lower bound for the product of their standard deviations. Stemming from the question of why are most statements regarding limitations and incompatibility usually only expressed for pairs of incompatible properties of a system, we showed through the triple uncertainty relation that the incompatibility of three observables is not a direct consequence of the incompatibility of each pair alone. We identified the unique state up to phase space translations that minimises the triple product, the triple sum and other symmetric functions of the three variances. We discussed the threefold symmetry associated with the triple, for which in the symmetric case the operator that cycles through its elements is a fractional Fourier transform. We finally conjectured an entropic version of the triple inequality, only proved for Gaussian states.

The focus of Chapter 4 was to take this idea one step further, as no indications suggested that this idea can not extend to more than three observables. Although a canonical triple can not be extended to a canonical quadruple or more, due to the impossibility to make all pair commutators equal to $\hbar/i$, we can however define closely related algebraic structures, if we relax the assumption that all commutators are the same and demand that only neighbouring ones are pairwise canonical. The coefficients of these rotated observables form regular polygons in $\mathbb{R}^2$, and the uncertainty relations that we derive for their variances, can be expressed in terms of geometric quantities of the associated polygon. We also prove inequalities for N observables which are arbitrary linear combinations of position and momentum and show that the lower bound can be expressed in terms of all the pairwise commutators. In both cases the lower bound can alternatively be geometrically associated with the area of the parallelogram spanned by two vectors in $\mathbb{R}^N$, constructed from the momentum and position "coefficients" of each observable. We also proved an inequality for the integral of rotated observables through an angle φ, for which a number of observations were made: for angles π and 2π the lower bounds agree with the areas of the associated disks but in all other cases the area of the disk is larger than the bound of the integral inequality. In addition, we derived a bound for the variances of three operators in more than one degree of freedom and subsequently extended the result to inequalities for N operators, which however are valid for product states only. Finally, we demonstrated how comparison of the different bounds in the case of three observables could be used for entanglement detection.

In Chapter 5, we considered a quantum particle in one spatial dimension and provided a set of equations that any function of its second moments must obey to attain a lower bound. These consistency conditions allowed us to derive a number of new inequalities along with already existing ones, unified under a common perspective. We discovered that the Robertson-Schrödinger inequality plays a special role for a quantum particle and that its extremal states are universal: any other expression of the second moments can be extremised only by a subset of them. The results can be represented geometrically in the space of second moments, and extrema correspond to points on the surfaces of nested one-sheeted hyperboloids, while the minima define the boundary of the uncertainty region. Its privileged position among other relations is due to the role of symplectic transformations for this problem and its solutions: it is the only inequality (and functions of it) that is invariant under any symplectic transformations, $Sp(2, \mathbb{R})$.

In Chapter 6 we studied the extrema of arbitrary functions of the second moments in N degrees of freedom, extending the framework of Chapter 5. We demonstrated that the extrema of any uncertainty functional are solutions to an eigenvalue equation quadratic in the operators of position and momentum in the N degrees of freedom, which we are able to solve by employing symplectic transformations. We found that there is a universal boundary for the uncertainty region, in agreement with the one-dimensional case. We showed that the solutions lead to a set of consistency conditions in analogy to the one-dimensional problem and we provided specific examples that utilise the method. We finally examined functionals of second moments in more than one states and derived the conditions for extrema, along with a number of examples.

The last chapter was concerned with measurement inequalities. First, we proposed a generalisation of the original Arthurs-Kelly model to allow for the description of a joint measurement of three and, subsequently, N observables. In this context, we proved a number of *joint-measurement* uncertainty relations for N operators and, by comparing with the preparational results of Chapter 4, we showed that there is always an increase in the product and sum of their variances. We found that allowing correlations between the probes improves significantly the lower bound with increasing number N of jointly measured observables, which is in contrast with the case of two, where no improvement is achieved. Based on this observation, we conjectured that in the presence of correlations the limit imposed by the preparational inequality is doubled, a bound significantly lower than the one for uncorrelated probes. By defining suitable noise operators, we also derived a number of inequalities for the errors and disturbances to the system due to the measurement. A straightforward generalisation of the considerations for a canonical pair to a triple allowed us to formulate the fundamental restrictions of measurement as statements in the form of *error-error-error* and *error-error-disturbance* uncertainty relations. In all cases, the lower bound was found to match the preparational one, in agreement with existing inequalities for the pair of position and momentum.

Thus, we provided answers to the question of whether a smooth function of the second moments of a quantum particle with one or more spatial degrees of freedom has a lower bound. Our analysis led to a simple characterisation in terms of the *consistency conditions*, whose solutions contain any potential minima. Structurally, the results for the one- and N-dimensional case are similar and the main differences stem from the structural differences of the symplectic groups between $n = 1$ and $n \neq 1$. The admissible region in the space of second moments for $N = 1$ is described by the Robertson-Schrödinger inequality, while for $N \geq 1$ it is described by the matrix inequality $\mathbf{C} + i\frac{\mathbf{\Omega}}{2} \geq 0$, with $\mathbf{C}$ being the covariance matrix.

For a particle in one dimension, using the general linear inequality for the second moments, (5.74), we derived a number of inequalities involving more than one observable, as an attempt to break the tradition of formulating uncertainty relations only for pairs. In the context of optics, these inequalities acquire a natural meaning, as they bound products and sums of the variances of the rotated quadratures [45, 44, 87].

There are a number of open questions or potential extensions of this work. It would be interesting to know whether the method developed in this thesis can be applied to systems associated with a finite-dimensional Hilbert space. In the continuous variables case, looking for the extrema of the functional leads to an eigenvalue equation involving a quadratic operator of the position and momentum. By employing a series of symplectic transformations, we showed that it is unitarily equivalent to the harmonic oscillator Hamiltonian, whenever a non-trivial bound exists. In the finite-dimensional case, the quadratic eigenvalue equation that we get is not solvable with simple transformations, in general, and as a result it appears to be impossible to diagonalise and obtain the analogous consistency conditions. However, for linear functionals with real constant coefficients, it is still possible to numerically obtain the eigenvalues and eigenstates of the resulting equation, and look for a global minimum among its solutions. We have not pursued this prospect yet and we are not certain to what extent this can lead to useful results.

Another interesting potential generalisation is the case of higher moments where it appears that our method can be used but only to some extent. After the variation of the functional, the resulting eigenvalue equations is no longer quadratic but of higher degree, dependent on the moments being considered. For example, for the product of the α-spread in position times the β-spread in momentum, we get an operator of the form $|\hat{q}|^\alpha + |\hat{p}|^\beta$, which one needs to diagonalise. In some cases the lower bound can be found exactly, but for the majority one can only look for the minimum numerically [16]. It seems plausible but has not been pursued further that expressions other than products and sums can be treated in this case, even if only with the aid of numerical calculations.

In Chapter 6 we considered measurement uncertainty in the context of an extended Arthurs-Kelly model and derived a number of joint measurement inequalities. We only briefly considered inequalities for the error and disturbance using specific noise operators, following a similar analysis as Appleby. Regarding the latter, more work is needed to clarify some of our findings. More specifically, it would be of importance to investigate error-disturbance uncertainty relations in terms of state independent error measures, such as the Wasserstein-2 distance.

We also stated a number of conjectures that need to be dealt with. In Chapter 4 we proposed an inequality for the entropies based on the observation that the inequality for pairs is based on an inequality for the (p, q)-norm of the Fourier transform. We showed that this holds for Gaussian states, owing to the equivalence of the Shannon entropies and the variances in that case, but a general proof eludes us at the moment. Finally, in Chapter 6 we conjectured a number of inequalities for more than three observables.

Appendices

Appendix A

Appendices of Chapter 2

A.1 A Baker-Campbell-Hausdorff identity

The relation $\hat{G}_b \hat{S}_\gamma = \hat{S}(\xi)\hat{R}(\chi)$ in Eq. (5.33) can be shown by requiring that both products map the annihilation operator $\hat{a} = (\hat{q} + i\hat{p}) \,/\, \sqrt{2\hbar}$ to the same operator. We obtain

$$\hat{G}_b \hat{S}_\gamma \hat{a} \hat{S}_\gamma^\dagger \hat{G}_b^\dagger = \hat{a}\left(\cosh\gamma - i\frac{b}{2}e^\gamma\right) + \hat{a}^\dagger\left(\sinh\gamma + i\frac{b}{2}e^\gamma\right) \tag{A.1}$$

and

$$\hat{S}(\xi)\hat{R}(\chi)\hat{a}\hat{R}^\dagger(\chi)\hat{S}^\dagger(\xi) = \hat{a}e^{-i\chi}\cosh r - \hat{a}^\dagger e^{-i\chi}e^{i\theta}\sinh r\,, \tag{A.2}$$

respectively. Equating the coefficients of the operators $\hat{a}$ and $\hat{a}^\dagger$ leads to two equations

$$\cosh\gamma - i\frac{b}{2}e^\gamma = e^{-i\chi}\cosh r\,, \tag{A.3}$$

$$\sinh\gamma + i\frac{b}{2}e^\gamma = -e^{i(\theta-\chi)}\sinh r\,, \tag{A.4}$$

which we need to solve for the variables $\xi \equiv re^{i\theta}$ and χ. Separating the real and imaginary parts of the first equation, one finds that

$$\chi = \arctan\left(\frac{b}{1+e^{-2\gamma}}\right) \in (-\frac{\pi}{2}, \frac{\pi}{2})\,. \tag{A.5}$$

In a similar way, the second equation allows one to solve for the function $\tan(\theta - \chi)$ which, upon using (A.5), leads to

$$\theta = \arctan\left(\frac{b}{1-e^{-2\gamma}}\right) + \arctan\left(\frac{b}{1+e^{-2\gamma}}\right) \in (-\pi\,,\pi)\,. \tag{A.6}$$

The case of $\gamma = 0$ needs to be treated separately leading to the relation

$$\theta = \pm\frac{\pi}{2} + \arctan\left(\frac{b}{2}\right) \in (-\pi, \pi). \tag{A.7}$$

Finally, the condition $\cosh r\, \cos\chi = \cosh\gamma$ results in the expression

$$r = \operatorname{arcosh}\left(\cosh^2\gamma + \frac{b^2}{4}e^{2\gamma}\right)^{1/2} \in [0,\infty)\,, \tag{A.8}$$

which establishes the desired identity (5.33).

The number states $|n\rangle, n \in \mathbb{N}$, are eigenstates of phase-space rotations $\hat{R}(\chi)$. Therefore, the product $\hat{G}_b\hat{S}_\gamma$ acts on those states according to

$$\hat{G}_b\hat{S}_\gamma|n\rangle \cong \hat{S}(\xi)|n\rangle\,, \tag{A.9}$$

where an irrelevant phase has been suppressed and the symbol "$\cong$" indicates equality up to overall constant phases. Thus, the operator $\hat{S}(\xi)$ generates all squeezed states from $|0\rangle$ when the parameter ξ runs through the points of the complex plane.

A.2 Convexity of the uncertainty region

Given two mixed quantum states described by density matrices $\hat{\rho}_1$ and $\hat{\rho}_2$, their convex combinations $\hat{\rho}_t = t\hat{\rho}_1 + (1-t)\hat{\rho}_2$, $t \in [0,1]$, are also quantum states. We will now show that the uncertainty region inherits convexity from the body of density matrices: mixing two states $\hat{\rho}_1$ and $\hat{\rho}_2$ with moment triples (x_k, y_k, w_k), $k = 1,2$, inside the uncertainty region results in another state $\hat{\rho}_t$ with a moment triple also in that region.

The moments $x_k = \mathrm{Tr}(\hat{x}^2\hat{\rho}_k)$, $k = 1,2$, etc., satisfy the RS inequality,

$$x_k y_k - w_k^2 \geq \frac{\hbar^2}{4} \equiv e_0^2\,, \quad k = 1,2\,, \tag{A.10}$$

and the moments of the mixture are given by

$$\sigma_t = t\sigma_1 + (1-t)\sigma_2\,, \qquad \sigma = x, y, w\,. \tag{A.11}$$

Writing $\bar{t} = 1 - t$, the variances of the convex combination satisfy

$$x_t y_t - w_t^2 \geq \left(t^2 + \bar{t}^2\right) e_0^2 + t\bar{t}\left(x_1 y_2 + x_2 y_1 - 2w_1 w_2\right) \tag{A.12}$$

using A.10. Since

$$\begin{aligned} x_1 y_2 + x_2 y_1 - 2w_1 w_2 &\geq e_0^2 \left(\frac{y_2}{y_1} + \frac{y_1}{y_2}\right) + \left(w_1\sqrt{\frac{y_2}{y_1}} - w_2\sqrt{\frac{y_1}{y_2}}\right)^2 \\ &\geq 2e_0^2\,, \end{aligned}$$

which implies that the moment triple of the convex combination $\hat{\rho}_t$ must also be contained in the uncertainty region, i.e.

$$x_t y_t - w_t^2 \geq \frac{\hbar^2}{4}\,. \tag{A.13}$$

The minimum is obtained only if either $t = 0$ or $t = 1$, so that the resulting density matrix must describe a state on the boundary of the uncertainty region, i.e. a Gaussian state.

A.3 Extrema of functionals of various general forms

A.3.1 Functions of the form $f = f(xy, w)$

Some general statements can be made for functionals corresponding to a function $f = f(xy, w)$; it is here argued that whenever solutions exist, they have to be at least a one parameter family.

The first partial derivatives of the function f are

$$f_x = y\frac{\partial f}{\partial(xy)} \equiv y f_{xy} \equiv y g_1(xy, w)\,, \quad f_y = x f_{xy}\,, \quad f_w = g_2(xy, w). \tag{A.14}$$

Then the first of the consistency conditions, Eq. (5.45), is trivially satisfied, which is essentially what guarantees that if there are solutions they have to lead to at least a one-parameter family. The second consistency condition gives:

$$\begin{aligned} 2wf_y = -xf_w \qquad &\Rightarrow \quad 2wxf_{xy} = -xg_2(xy,w) \quad \Rightarrow \\ 2wf_{xy} + g_2(xy,w) = 0 \quad &\Rightarrow \quad 2wg_1(xy,w) + g_2(xy,w) = 0\,. \end{aligned} \tag{A.15}$$

The last equation either has no solutions, is trivially satisfied, or has a solution of the form $w = h_1(xy)$.

If it is trivially satisfied then it leads to the RS extrema. If it has a solution of the form $w = h_1(xy)$, then the third equation gives

$$xy - w^2 = \left(n + \frac{1}{2}\right)^2 \hbar^2 \;\Rightarrow\; xy - h_1^2(xy) = \left(n + \frac{1}{2}\right)^2 \hbar^2 \tag{A.16}$$

The last equation will either have no solution or lead to an inconsistency, or assume a solution of the form

$$xy = h_2\left((n + 1/2)^2\, \hbar^2 \right) \equiv K(n). \tag{A.17}$$

Thus, if there are no inconsistencies for the solutions obtained, i.e. $F > 0$ for all the extrema, then we have found that the potential solution is of the form

$$y_e = K(n)/x\,,\;\; w_e = h_1(xy) = h_1\left(K(n)\right) \equiv H(n), \tag{A.18}$$

which obviously leads to a one parameter family of states.

To sum up, we examined functionals of the form $f = f(xy, w)$ and showed that there are no solutions that lead to isolated points in the b, γ plane: there will either be no solutions at all, we will get the full RS extremal set, or we will get a one parameter family.

A.3.2 Functions of the form $f = f(ax^m + by^k, w)$

As in the case of last section, some general statements about functionals of the form $f = f(ax^m + by^k, w)$ can be made and whenever solutions exist, they now have to be a one parameter family or isolated points.

Let $r = ax^m + by^k$. Then, the first partial derivatives are

$$f_x = amx^{m-1} f_r \equiv amx^{m-1} g_1(r, w) \,, \quad f_y = bky^{k-1} g_1(r, w) \,, \quad f_w = g_2(r, w) \,. \tag{A.19}$$

Then the first of the consistency conditions gives

$$\begin{aligned} x f_x = y f_y \;\Rightarrow\; amx^m &= bky^k \;\Rightarrow \\ y = \left(\frac{am}{bk}\right)^{1/k} & x^{m/k} \,. \end{aligned} \tag{A.20}$$

Thus, one can use the last relation to reduce from three to two variables; since there is a relation between them then the extremal set can already be of at most one free parameter. Then,

$$r(x, y) \equiv r(x) = a \left(1 + \frac{m}{k}\right) x^n. \tag{A.21}$$

The second consistency condition gives:

$$\begin{aligned} 2w f_y = -x f_w \;\Rightarrow\; 2wbk \left(\frac{am}{bk}\right)^{k-1/k} x^{m(k-1)/k} f_r(r, w) &= -x g_2(r, w) \\ 2w g_3(x, w) &= -x g_4(x, w) \,. \end{aligned} \tag{A.22}$$

The last equation will either have no solutions, will be trivially satisfied, or have a solution of the form $w = h_1(x)$. If it is trivially satisfied then it will lead to a one parameter set, while if it has a solution of the form $w = h_1(x)$, then the third equation gives

$$xy - w^2 = \left(n + \frac{1}{2}\right)^2 \hbar^2 \;\Rightarrow\; \left(\frac{am}{bk}\right) x^{m+k/k} - h_1^2(x) = \left(n + \frac{1}{2}\right)^2 . \tag{A.23}$$

The last equation will either have no solution or lead to an inconsistency, or assume a solution of the form

$$x = h_2 \left((n + 1/2)^2 \hbar^2 \right) \equiv K(n). \tag{A.24}$$

Thus, if there are no inconsistencies for the obtained solutions, i.e. $F > 0$ for all the extrema, then we have found that the potential solution is of the form

$$x_e = K(n) = \left(\frac{am}{bk}\right)^{1/m} y_e^{k/m}\,, \;\; w_e = h_1(x) = h_1\left(K(n)\right) \equiv H(n), \tag{A.25}$$

and obviously corresponds to a point in the (b, γ) space.

Table of Notation

$\mathscr{H}$	Hilbert space
$\hat{A}$, $\hat{B}$	Self-adjoint operators
$\mathbb{I}$	Identity operator
$\hat{U}$, $\mathscr{U}$	Unitary operators
$\hat{T}_\alpha$	Unitary operator effecting phase space translations
$\hat{S}_\gamma$, $\hat{S}(\xi)$	Unitary operator effecting real/complex squeezing
$\lvert\psi\rangle$, $\lvert\varphi\rangle$	Unit vectors in $\mathscr{H}$
$\Delta^2 A$	Variance of self-adjoint operator $\hat{A}$
C_{AB}	Covariance of self-adjoint operators $\hat{A}, \hat{B}$
S_r	Shannon entropy of the probability distribution of $\hat{r}$
$J[\psi]$	Uncertainty functional for a pure state $\lvert\psi\rangle$
$dF_\phi(x)$, D_ϕ	Gâteaux differential
$Sp(2n, \mathbb{R})$	The n-dimensional symplectic group
$\mathbf{\Omega}$	n-symplectic form
$\mathbf{C}$	Covariance matrix
$W(p, q)$	Wigner quasi-probability distribution
$f : V \to W$	A function from V to W
$\mathscr{F}_\alpha[f]$	Fractional Fourier transform of function f of angle α
$\mathbb{N}_0$, $\mathbb{R}_+$	Natural numbers including zero, positive real numbers
$\in$	"belongs to"
$\forall$	"for all"
$\setminus$	"excluding"

List of References

[1] G. Adesso, and F. Illuminati, J. Phys. A **40**, 7821 (2007).

[2] G. Adesso, S. Ragy, and A.R. Lee, Open Syst. Inf. Dyn. **21**, 1440001 (2014).

[3] D. M. Appleby, Int. J. Theor. Phys. **37**, 1491 (1998).

[4] D. M. Appleby, Int. J. Theor. Phys. **37**, 2557 (1998).

[5] D.M. Appleby, J. Phys. A Math. Gen. **31**, 6419 (1998).

[6] E. Arthurs and J.L. Kelly, *Bell Syst. Tech. J.* **44**, 725 (1965).

[7] E. Arthurs and M .S. Goodman, Phys. Rev. Lett. **60**, 2447 (1988).

[8] Arvind, B. Dutta, N. Mukunda, and R. Simon, Pramana **45**, 471 (1995).

[9] K. I. Babenko, Izv. Akad. Nauk SSSR. Set. Mat. 25, 531 (1961).

[10] D. Bartlett, T. Rudolph and R. W. Spekkens, Phys. Rev. A **86**, 012103 (2012).

[11] W. Beckner, Proc. Nat. Acad. Sci. U.S.A. 72, 638 (1975).

[12] A. Bednorz, W. Belzig, Phys. Rev. A **83**, 052113 (2011).

[13] M. Bellini, A. S. Coelho, S. N. Filippov, V. I. Man'ko, and A. Zavatta, Phys. Rev. A **85**, 052129 (2012).

[14] I. Bialynicki-Birula and J. Mycielski, Comm. Math. Phys. **44**, 129 (1975).

[15] P. Busch, Int. J. Theor. Phys. **24**, 63 (1985).

[16] P. Busch, P. Lahti and R. F. Werner, Rev. Mod. Phys. **86**, 1261 (2014).

[17] P. Busch, P. Lahti, R. F. Werner, J. Math. Phys. **55**, 042111 (2014).

[18] P. Busch, P. Lahti, R. F. Werner, *Measurement Uncertainty: Reply to Critics*, *arXiv:1402.3102* (2014).

[19] P. Busch, P. Lahti and R. F. Werner, Phys. Rev. Lett. **111**, 160405 (2013).

[20] P. Busch, P. Lahti, R. F. Werner, Phys. Rev. A **89**, 012129 (2014).

[21] P. Busch, T. P. Schonbeck and F. Schroeck Jr., J. Math. Phys. **28**, 2866 (1987).

[22] P. Busch, T. Heinonen, and P. Lahti, Phys. Rep. *452*, 155 (2007).

[23] P. Busch, M. Grabowski, P.J. Lahti, *Operational Quantum Physics* (Springer, Berlin, 1997).

[24] A. Cabello, *Interpretations of quantum theory: A map of madness*, *arXiv:1509.04711* (2015).

[25] L. Dammeier, R. Schwonnek, and R. F. Werner, New J. Phys. **17**, 093046 (2015).

[26] G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, Adv. Imaging Electron Phys. **128**, 205 (2003).

[27] V. V. Dodonov, E. Y. Kurmyshev and V. I. Man'ko, Phys. Lett. A **79**, 150 (1980).

[28] V. V. Dodonov, J. Opt. B **4**, R1 (2002).

[29] J. Dressel and F. Nori, Phys. Rev. A **89**, 022106 (2014).

[30] L.-M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **84**, 2722 (2000).

[31] J. Erhart, S. Sponar, G. Sulyok, G. Badurek, M. Ozawa and Y. Hasegawa, Nat. Phys. **8**, 185 (2012).

[32] C. A. Fuchs and A. Peres, Phys. Today **53** (3), 70 (2000).

[33] C. A. Fuchs, N. D. Mermin, and R. Schack, Am. J. Phys. **82**, 749 (2014).

[34] C. R. Givens and R. M. Shortt, Michigan Math. J., **31**, 231 (1994).

[35] W. Heisenberg, Z. Phys. **43**, 172 (1927).

[36] R. Heese and M. Freyberger, Phys. Rev. A **89**, 052111 (2014).

[37] I. I. Hirschman, Jr., Am. J. Math. **79**, 152 (1957).

[38] R. Jackiw, J. Math. Phys. **9**, 339 (1968).

[39] S. Kechrimparis and S. Weigert, Phys. Rev. A **90**, 062118 (2014).

[40] S. Kechrimparis and S. Weigert, *Universality in Uncertainty Relations for a Quantum Particle*, *arXiv:1509.02146* (2015).

[41] E. H. Kennard, Z. Phys. **44**, 326 (1927).

[42] E. Kot, N. Grønbech-Jensen, B. M. Nielsen, J. S. Neergard-Nielsen, E. S. Polzik, A. S. Sørensen, Phys. Rev. Lett. **108**, 233601 (2012).

[43] K. Korzekwa, D. Jennings, and T. Rudolph, Phys. Rev. A **89**, 052108 (2014).

[44] U. Leonhardt and H. Paul, Prog. Quant. Electr. **19**, 89-130 (1995).

[45] U. Leonhardt and H. Paul, J. Mod. Optics **40**, 1745-1751 (1993).

[46] L. Maccone, and A. K. Pati, Phys. Rev. Lett. **113** (26), 260401 (2014).

[47] V. I. Man'ko, G. Marmo, A. Simoni and F. Ventriglia, Adv. Sci. Lett. **2**, 517 (2009).

[48] K. B. Møller, T. G. Jørgensen, and J. P. Dahl, Phys. Rev. A **54**, 5378 (1996).

[49] B. R. Mollow and R. J. Glauber, Phys. Rev. **160** 1076 (1967).

[50] V. Namias, *IMA. J. Appl. Math.* **25**, 241-264 (1980).

[51] T. Norsen and S. Nelson, *Yet Another Snapshot of Foundational Attitudes Toward Quantum Mechanics*, *arXiv:1306.4646* (2013).

[52] M. Ozawa, Phys. Rev. A **67**, 042105 (2003).

[53] M. Ozawa, Ann. Phys. (N.Y), **311**, 350 (2004).

[54] J. Plebanski, Phys. Rev. **101** 1825 (1956).

[55] R. Raymer, Am. J. Phys. **62**, 986 (1994).

[56] H. P. Robertson, Phys. Rev. **34**, 163-164 (1929).

[57] H. P. Robertson, Phys. Rev. **46**, 794 (1934).

[58] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar and A. M. Steinberg, Phys. Rev. Lett. **109**, 100404 (2012).

[59] A. Ruszczyński, *Nonlinear Optimization*, (Princeton Univ. Press., 2006).

[60] H. Sagan, *Introduction to the Calculus of Variations*, (Mc-Graw-Hill, New York, 1969).

[61] M. Schlosshauer, J. Kofler, and A. Zeilinger, *A Snapshot of Foundational Attitudes Toward Quantum Mechanics*, *arXiv:1301.1069* (2013).

[62] E. Schrödinger, Sitzungsber. Preuss. Akad. Wiss. (Phys.-Math. Klasse) **19**, 296 (1930).

[63] A. Serafini, Phys. Rev. Lett. **96**, 110402 (2006).

[64] J.-P. Serre, *A Course on Arithmetic* (Springer-Verlag, New York, 1973).

[65] M. I. Shirokov, *Interpretation of Uncertainty Relations for Three or More Observables*, *arXiv:quant-ph/0404165*.

[66] R. Simon, Phys. Rev. Lett. **84**, 2726 (2000).

[67] R. Simon, N. Mukunda, and B. Dutta, Phys. Rev. A **49**, 1567 (1994).

[68] N. J. A. Sloane and Simon Plouffe, *The Encyclopedia of Integer Sequences*, (Academic Press, 1995).

[69] Q. Song, C. Qiao, *Stronger Schrödinger-like uncertainty relations*, *arXiv:1504.01137* (2015).

[70] S. Stenholm, 1992, Ann. Phys. (N.Y.) **218**, 233.

[71] D. Stoler, Phys. Rev. D **1** 3217 (1970).

[72] A. Sudbery, *Quantum Mechanics and the Particles of Nature*, (Cambridge Press, Cambridge, 1986) .

[73] D. A. Trifonov, *Schrödinger uncertainty relation and its minimization states*, *arXiv:physics/0105035* (2001).

[74] D. A. Trifonov, J. Phys. A: Math. Gen. **33** L299 (2000).

[75] S. Weigert and M. Wilkinson, Phys. Rev. A **78**, 020303(R) (2008).

[76] C. Weedbrook, S. Pirandola, R. Garcia-Patron, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Rev. Mod. Phys. **84**, 621 (2012).

[77] S. Weigert (unpublished).

[78] S. Weigert, Phys. Rev. A **53**, 2084-2088 (1996).

[79] S. Weigert, J. Phys. A: Math. Gen. **35** 4169 (2002).

[80] D. G. Welsch, W. Vogel, T. Opatrny, in *Progress in Optics*, *Vol. XXXIX*, edited. E. Wolf (Elsevier, Amsterdam, 1999), pp. 63–211.

[81] H. Weyl, *Gruppentheorie und Quantenmechanik* (Hirzel, Leipzig, 1928) [English translation, H. P. Robertson, *The Theory of Groups and Quantum Mechanics* (Dover, New York, 1931)].

[82] J. A. Wheeler and W. H. Zurek, *Quantum Theory and Measurements* (Princeton University Press, Princeton,1983).

[83] R. M. Wilcox, J. Math. Phys. **8**, 962 (1967).

[84] J. Williamson, Am. J. Math. **58**, 141 (1936).

[85] E. Wigner, Phys. Rev. **40**, 749 (1932).

[86] K. Wódkiewicz, Phys. Lett. A **124**, 207 (1987).

[87] Y. Yamamoto and H. Haus, Rev. Mod. Phys. **58**, 1001 (1986).

[88] H. P. Yuen, in: *Quantum Optics, Experimental Gravitation and Measurement Theory*, eds. P. Meystre and M. O. Scully (Plenum, New York, 1983) p. 249.