Majorization theoretical approach to quantum uncertainty

From Wigner entropy to Gaussian bosonic channels

Author
Zacharie Van Herstraeten

Supervisor
Nicolas Cerf

Centre for Quantum Information and Communication
École polytechnique de Bruxelles
Université libre de Bruxelles

Zacharie Van Herstraeten
Université libre de Bruxelles
École polytechnique de Bruxelles
50 av. F.D. Roosevelt - CP165/59
1050 Bruxelles
Belgique
Email: zvherstr@ulb.ac.be

Thèse présentée en vue de l'obtention du grade de Docteur en Sciences de l'Ingénieur et Technologie.

Défense publique le 10 décembre 2021 à l'Université libre de Bruxelles.

Jury : **Stephan De Bièvre**, Université de Lille, France

Anthony Leverrier, INRIA Paris, France

Ognyan Oreshkov, Secrétaire du jury, ULB

Stefano Pironio, Président du jury, ULB

Karol Życzkowski, Jagiellonian University, Pologne

Nicolas J. Cerf, Superviseur, ULB

Acknowledgments

Si c'était possible, j'aimerais recommencer une thèse, avec le même promoteur, sur le même sujet, et dans le même service. Ces quatre années ont été pour moi des années enrichissantes, stimulantes, pleines de découvertes, et surtout pleines de liberté. Je regrette déjà qu'elles soient passées si vite !

Merci sincèrement Nicolas, car je pense que c'est avant tout grâce à toi que ces années de doctorat ont été si agréables pour moi. Merci de m'avoir guidé tout au long de cette thèse. Ton enthousiasme scientifique est communicatif et a régulièrement regonflé ma motivation. Je pense que le sujet de thèse que tu m'as proposé est très riche, et je ne manquerai pas de continuer à y réfléchir dans le futur. Les fonctions de Wigner finiront un jour par livrer leur secret ! Merci également de m'avoir lancé sur les rails de ce postdoc outre-atlantique.

J'adresse ensuite un merci particulier à Michael. Merci Michael d'avoir été présent lorsque je débutais ma thèse. Heureusement que tu étais là pour répondre à mes questions lorsque j'étais encore trop intimidé par Nicolas ! J'espère avoir la chance de continuer à collaborer avec toi dans de futurs projets.

Je remercie mes camarades de bureau Matthieu, Célia, et plus récemment Antoine, pour leur bonne humeur. Je regrette que le covid ait mis sur pause la vie de service pendant une si longue période. Je vous souhaite bon courage pour terminer vos thèses, et une bonne route pour la suite.

Je remercie Leonardo pour sa bienveillance et sa gentillesse. Je remercie Timothée pour son humour de haute voltige, et Philippe pour son optimisme. Je suis ravi de remercier Ravi. Je remercie Pascale pour sa sympathie et son efficacité. Je remercie Jérémie et Ognyan d'être des membres permanents faciles d'accès et abordables. Je souhaite une bonne continuation à tous les autres membres du service, Julian, Lin-Qing, Hamed, Enea, Joseph et Benoit. Je suis également heureux d'avoir eu la chance de croiser la route de Shan, Atul, Levon, Mathieu, Anaelle, Evgueni, Stefan, Luc, Raúl, Krishna, Zoé, Simon, Uttam, Siddhartha, Chrysoula et Esteban.

Je remercie le FNRS pour la bourse FRIA qui m'a permis de travailler dans un grand confort et une grande liberté.

Je remercie les membres du jury pour leur intérêt. Merci pour vos remarques et commentaires qui ont significativement amélioré cette thèse.

Enfin, plus personnellement, je remercie mes parents. Je remercie ma mère qui m'a transmis le goût des études et mon père qui est à l'origine de mon intérêt pour les maths. Je remercie ma soeur, Zoé, ainsi que le Professeur Vermisseau. Je remercie Rose qui, par intermittence, a su pimenter mes années de doctorat. Je remercie mes amis et mes colocs, et les renvoie vers les présent manuscrit afin de justifier ma disparition des radars au cours des deux derniers mois.

Abstract

This thesis is centered on a novel approach to quantum uncertainty based on applying the theory of continuous majorization to quantum phase-space distributions. Majorization theory is a powerful mathematical framework that is aimed at comparing distributions with respect to intrinsic disorder. It is particularly significant in the sense that establishing a majorization relation between two distributions amounts to proving that every (Shur-concave) measure of disorder will categorize one distribution as more ordered than the other. Although this is less known, the distributions here do not need to be normalized nor positive for majorization theory to apply, so the latter even extends beyond probability distributions. Further, a majorization relation can rigorously be defined for both discrete and continuous distributions over a finite-size domain, as well as for (discrete and continuous) distributions that are positive over an infinite-size domain.

The central thrust of this thesis is to characterize quantum uncertainty in phase space by applying the tools of majorization theory to the Wigner function, which is the most common (quasi)distribution that embodies a quantum state in phase space. Wigner functions are in general positive and negative, putting them beyond the reach of most information-theoretical measures but perfect candidates for the theory of majorization. We start our manuscript with a succinct overview of the basics of quantum optics in phase space, which are a prerequisite for the characterization of disorder in phase space. This gives us the occasion to present a secondary achievement of the thesis consisting in establishing a resource theory for local Gaussian work extraction, which exploits the symplectic formalism within quantum thermodynamics. In this context, work can be defined as the difference between the trace and symplectic trace of the covariance matrix of the state, and it displays a number of interesting properties. Back to our primary interest, our first contribution is to construct an extended formulation of majorization with the applicability to Wigner functions as our main objective. It must be stressed that when relaxing the positivity condition, majorization relations have not been addressed in the literature for (discrete or continuous) distributions defined over an infinite-size domain. Here, we write extended majorization relations that equally apply to discrete or continuous, positive or negative distributions defined over a finite-size or infinite-size domain. Then, applying these to Wigner functions (which can be negative and have an infinite-size domain), our first finding is that all pure states are either incomparable or equivalent as regards majorization. Hence, any pure state cannot be objectively deemed more disordered than any other one in phase space, so that majorization appears to be better suited to compare mixed states.

A large part of the thesis is then concerned with the convex set of quantum states that have a positive Wigner function, seeking a better understanding of its structure. As a consequence of Hudson theorem, this set only contains mixed states, with the notable exception of Gaussian pure states. We highlight a large subset of Wigner-positive states that can be prepared using a balanced beam-splitter and show that these states play a key role in the geometry of the Wigner-positive set as they are extremal states. Restricting ourselves to Wigner-positive states, we formulate the conjecture that the Gaussian pure states (most notably, the vacuum state) are the states of least disorder, which is expressed via a fundamental majorization relation. This conjecture is supported by numerical simulations and is analytically proven over the subset of phase-invariant states containing up to two photons. It can be viewed as a precursor to all uncertainty relations in phase space.

Our next contribution pertains to the usage of information theory in quantum phase space. The conjectured fundamental majorization relation for Wigner-positive states implies in turn an infinite number of inequalities relating Schur-concave functionals, most notably the Rényi entropy and the Shannon differential entropy of the Wigner function. We define the latter as the Wigner entropy of a state, and conjecture that it is lower bounded by $\ln \pi + 1$, namely the value it takes for Gaussian pure states. We then prove that this bound holds over the thermodynamically-relevant set of passive states in arbitrary dimension. The Wigner entropy is itself a lower bound on the sum of the entropies of the

marginal distributions in phase space, which makes a clear connection with the entropic uncertainty relations of Białynicki-Birula and Mycielski (which are implied by our conjecture). Interestingly, the Wehrl-Lieb conjecture is also a consequence of our conjecture (while it does not imply it). Turning back to uncertainty measures that could also be applicable to Wigner-negative states, we also investigate which Schur-concave functionals can be considered as relevant uncertainty measures in phase space.

A last contribution of the thesis concerns (discrete) majorization in state space, which is the usual way of applying majorization theory to quantum physics as described in the literature. Here, we focus on Gaussian phase-invariant bosonic channels and demonstrate the existence of a majorization ladder for Fock states at the input of the channel. This extends a prior analysis that was restricted to special (quantum-limited) channels. We also find a simple relation to determine whether such channels produce a Wigner-positive state for any input state, echoing the condition for entanglement-breaking channels. We hope that these considerations may help resolving some of the pending questions regarding entropies in bosonic systems and channels.

List of publications

- *Quantum thermodynamics in a multipartite setting: A resource theory of local Gaussian work extraction for multimode bosonic systems*
  Uttam Singh, Michael G. Jabbour, Zacharie Van Herstraeten and Nicolas J. Cerf
  Physical Review A **100**, 042104 (2019)
  arXiv:1905.02948, DOI:10.1103/PhysRevA.100.042104

- *Quantum Wigner entropy*
  Zacharie Van Herstraeten and Nicolas J. Cerf
  Physical Review A **104**, 042211 (2021)
  arXiv:2105.12843, DOI:10.1103/PhysRevA.104.042211

- *Continuous majorization in quantum phase space*
  Zacharie Van Herstraeten, Michael G. Jabbour and Nicolas J. Cerf
  Submitted to Quantum
  arXiv:2108.09167

- *Majorization ladder in Gaussian bosonic channels*
  Zacharie Van Herstraeten, Michael G. Jabbour and Nicolas J. Cerf
  In preparation

Contents

Introduction

What is disorder? When you think about it, the notion of disorder is harder to grasp than it seems at first sight. Generally speaking, disorder is opposed to order. Something that is ordered is something that is well arranged, when everything is in its place, with no room for uncertainty. In contrast, disorder appears as soon as something is in a place where it was not expected to be, that is, as soon as uncertainty enters into the picture. Hence, when we try to define disorder, we quickly come to the theory of probabilities. Indeed, what could be more adequate than a probability distribution in order to represent something that is uncertain? Disorder thus becomes a property of the probability distribution underlying the system under study.

From a mathematical point of view, there exist a large variety of measures that can be used in order to quantify disorder, as we are going to see throughout this thesis. All these measures generally quantify a particular aspect of disorder, and the entropy is of course the paradigm of such a measure of disorder. Equipped with a measure, it becomes possible, for example, to distinguish a narrow from a wide distribution, or a peaked from a flattened distribution. Naturally, the wide distribution will be considered as more disordered than the narrow distribution. Similarly, the flattened distribution will be considered as more disordered than the peaked distribution. However, the comparison may be more subtle than it seems, as illustrated on Figure 1. Let us consider the comparison between distributions f_1 and f_2. In this example, distribution f_1 is more peaked than distribution f_2, but it has a wider support. Then, what can we say then about their respective disorder? The answer depends on the actual measure that is used, which itself depends on the context, but it so happens that f_1 may be more ordered than f_2 according to some measures and more disordered than f_2 according to other measures. (The situation is easier for f_3 which is more disordered than both f_1 and f_2.) From this example, we learn that the measures of disorder have inherent limitations in that they cannot always be used to compare disorder in an unambiguous way. In this thesis, we will see that there exits a well-defined mathematical tool, called majorization theory, which allows us to go one step further and approach the notion of disorder from a more global viewpoint than any simple measure can do.

The playground on which we will apply these notions of disorder is the phase space of coordinates of a physical system. The phase space is an abstract space which originates from statistical physics. In classical mechanics, it is the locus of all possible configurations of the system. There is one dimension of phase space for the coordinate associated with each degree of freedom of the system, as well as another dimension for the momentum that is canonically conjugate to each of these coordinates. For example, the spatial coordinate and its conjugate momentum form the two-dimensional phase space

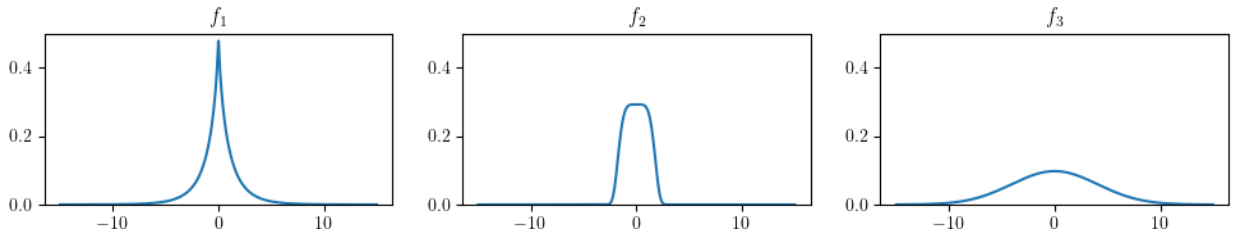

Figure 1: Qualitative examples of distributions. The distribution f_1 appears to be more peaked than f_2, however it is also more widespread. Hence, it is difficult to decide which one is the more disordered between f_1 and f_2. The case of f_3 is easier to decide, as it appears to be more widespread and more flattened than both f_1 and f_2.

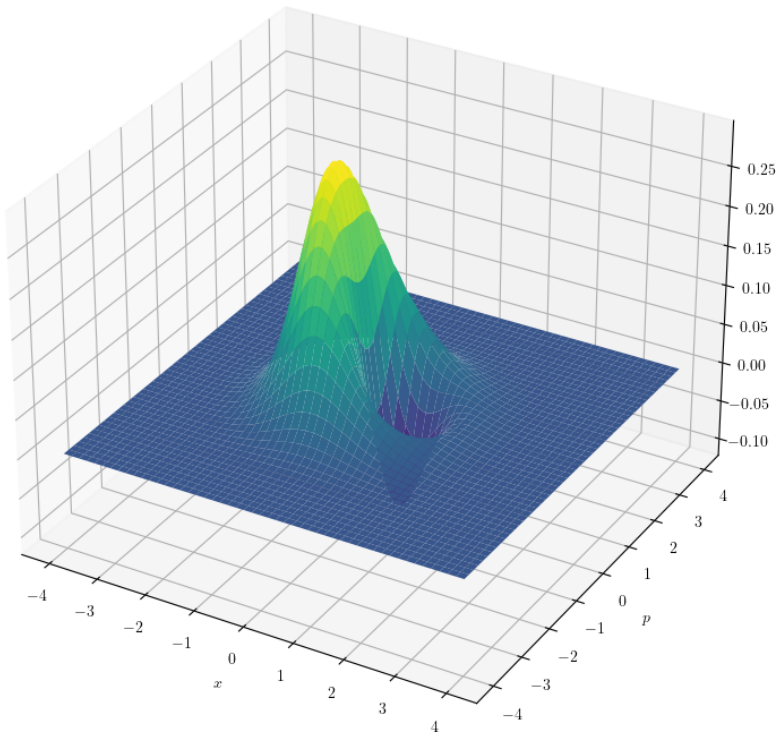

Figure 2: Illustration of the Wigner function of a quantum state. The function is defined over phase space and takes in general both positive and negative values.

needed to describe a single particle in one dimension. The laws of motion are completely determined if the set of all coordinates and momenta are known, so that a point in phase space follows a perfectly determined trajectory. One then understands that phase space gives rise to a privileged formulation of the evolution of the state of a system. In this context, the notion of disorder appears naturally when considering physical systems that are not perfectly determined or controlled, but are defined according to a probability distribution in phase space, as can for example be the case for a system at finite temperature. The probability distribution will then evolve in phase space following a Liouville equation, in accordance with the laws of motion.

Things get tougher when it comes to quantum mechanics. Indeed, in the quantum formalism, the observable quantities associated with the coordinates and conjugate momenta become linear operators in a Hilbert space. On top of this, pairs of conjugate operators do not commute, which implies that a state can never have a well-determined coordinate and, at the same time, a well-determined associated momentum. This observation implies that a perfectly determined state – a point in phase space, as one could have classically – no longer exists in quantum mechanics. While in classical mechanics, disorder as witnessed by a spreading in phase space necessarily originates from some statistical uncertainty, quantum mechanics holds, in contrast, an intrinsic source of uncertainty. This raises the question of whether the phase space formalism could be used in quantum mechanics. Representing the state of a quantum system in phase space seems unsuitable at first sight. Indeed, how could it result from some distribution of "point states" that are physically forbidden by quantum mechanics? This apparent impossibility can however be overcome if we allow ourselves to consider distributions that can be negative in some regions of phase space. This is the role played by the Wigner function in quantum mechanics, which behaves just like a probability distribution in phase space except that it can be partly negative as illustrated in Figure 2. This discussion, though superficial, clearly indicates that studying disorder in quantum phase space will necessarily require special tools that go beyond the standard framework of probability theory.

Let us now introduce the central mathematical theory that we will use throughout this thesis, namely the theory of majorization. It is a very powerful tool, which is designed to compare distributions in terms of their disorder. It comes in two complementary facets, namely discrete majorization, which deals with discrete probability distributions (probability vectors), and continuous majorization,

which deals with continuous distributions (probability density functions). When a majorization relation is established between two distributions, it implies in turn an infinite number of inequalities on a large class of functions, called Shur-concave functions, which can be viewed as measures of disorder. One can then understand the majorization relation as the fact that one distribution is objectively more disordered than the other, in the sense that every measure of disorder will compare the two distributions in the same way. Another strength of majorization theory is that it goes beyond the scope of probability distributions. Indeed, majorization enables comparing non-normalized distributions (in this case, one speaks of weak majorization) and even functions that can be both positive and negative. The theory of continuous majorization appears therefore perfectly suitable to tackle the characterization of disorder among Wigner functions, which is the central thrust of this thesis.

Majorization theory is also closely linked to information theory. Indeed, in information theory, one is concerned with the disorder or uncertainty contained in a probability distribution, which can alternatively be viewed as the information that must be transmitted in order to acquire one instance of the associated random variable. This information content is generally measured with the Shannon entropy, named after the seminal paper by Shannon which is often considered as the birth certificate of information theory [60]. This particular measure, though it is not unique, enjoys numerous interesting properties and is the most suitable tool to investigate the communication rates in channels. The link with majorization theory is straightforward as the Shannon entropy is nothing else but one particularly useful and predominant measure of disorder. Hence, a majorization relation between two probability distributions simply determines which one has a larger Shannon entropy than the other.

Information theory has evolved, over the past quarter century, into a quantum form, which will actually be our main concern in this thesis. This was triggered by the realization that information, whatever it is, must necessarily be carried by a physical system, which in turn must obey the laws of physics. In other words, information is physical [40]. Hence, the laws of quantum physics have brought some brand new perspectives on the field of information theory, and, as such, quantum information theory is now considered as a research field in its own right. Among the many outcomes of this field, let us mention just a single one, namely quantum cryptography. More precisely, quantum key distribution is a procedure which, unlike its classical analog, allows us to establish a secret communication channel whose confidentiality relies on the laws of physics, without the need for any computational hypotheses. This quantum advantage originates from the impossibility to measure a variable encoded in a quantum state without disturbing its conjugate variable, which is reminiscent of the uncertainty principle. On a more fundamental side, let us mention that quantum information theory is responsible for a regained interest in uncertainty relations, in particular entropic uncertainty relations which allow us to formulate the uncertainty principle in very general terms and will be touched upon in this thesis.

Quantum information theory has been developed in two related flavors, discrete or continuous. We speak about discrete-variable quantum information when the considered states are described by a quantum degree of freedom that takes discrete values. The paradigmatic objects of discrete-variable quantum information are the qubits, which can be materialized, for example, by electrons with a spin $1/2$ or atoms with two discrete energy levels. In contrast, we speak about continuous-variable quantum information when considering quantum states defined over a continuous degree of freedom. This could for example be a component of the electromagnetic field, in which case we speak of continuous-variables bosonic quantum information. It is mostly this formalism that we will use throughout this thesis, as it is in this context that the Wigner function takes all its meaning.

Majorization theory is known to play an essential role in quantum information theory. More precisely, the discrete version of majorization theory has been very successfully used in the context of entanglement transformations [49, 51], for the discrimination of entangled states [50, 30], and to generalize quantum uncertainty relations [53, 57, 15]. In other words, the potential of discrete majorization in quantum information theory is well established. Surprisingly enough, the continuous version of majorization has remained unexplored to date in this field. Yet, as argued earlier, the quantum phase space seems to lend itself very well to applying this tool, so the primary objective of this thesis is to initiate the application of continuous majorization to quantum information theory.

The central objective that we pursue throughout this work is to shape the theory of majorization to quantum phase space in order to derive useful applications. We conduct our research using the strategy

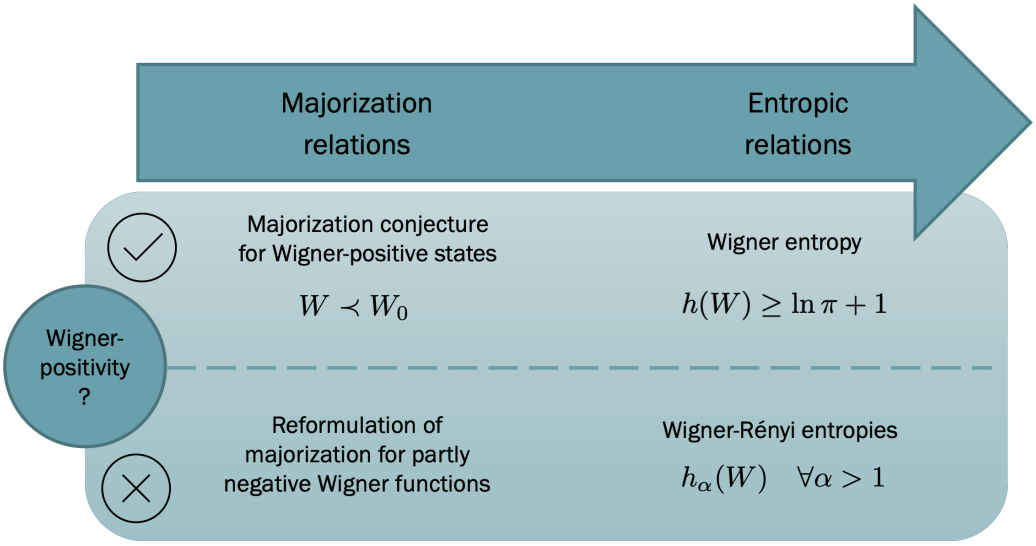

Figure 3: This figure schematizes the research flow of this thesis. We aim to establish strong majorization relations, which will in turn imply entropic relations that can be used in practice. We divide our approach in two parts depending whether we consider Wigner-positive states or not. In the case of Wigner-positive states, we already have the required mathematical background for our work. We thus focus our attention on a continuous majorization conjecture, and on a conjecture over their Wigner entropy. When dealing with partly negative Wigner functions, we have to build the appropriate mathematical tools for our work. This leads us to an extended reformulation of continuous majorization, and to the definition of Wigner-Rényi entropies.

that we describe hereafter and schematize in Figure 3. We aim to construct strong majorization relations for Wigner functions, which will then in turn imply a wide amount of inequalities, and in particular entropic relations. In order to do so, we proceed differently whether we consider quantum states with non-negative Wigner functions or quantum states with partly negative WIgner functions. In the case of Wigner-positive states, we have at our disposal the usual tools of probability theory, and we focus on a continuous majorization conjecture. That conjecture implies then a conjectured lower-bound on the Shannon differential entropy of their Wigner function, which we define as the Wigner entropy. After dealing with Wigner-positive states, we turn our attention to quantum states with partly negative Wigner functions. This leads us to present a reformulation of continuous majorization which encompass the case of partly negative distributions defined over an infinite-size domain. Finally, we study which entropic measures are relevant in the case of partly negative Wigner functions, which leads us to define the Wigner-Rényi entropies.

This thesis is structured into 9 chapters. In Chapter 1, we quickly describe the path from classical to quantum phase space. We first recall notions of statistical physics in order to illustrate the classical meaning of a phase-space distribution and then introduce the Wigner function as the core mathematical object in quantum phase space. We define it from a very general perspective going beyond quantum optics, as it applies to any quantum system associated with a continuous degree of freedom. We consider continuous-variable pure and mixed states, and introduce the Weyl transform to move from density operators in a state-space view to Wigner functions in a phase-space view. We then go over the properties that make the Wigner function an object that is comparable to a probability distribution, as well as the properties that make it an intrinsically quantum object. Overall, Chapter 1 portrays the general landscape of this thesis, especially the Wigner function which plays a central role. The landscape and goals being set, Chapters 2-4 cover the material upon which the thesis is built (information theory, majorization theory, and quantum optics). The original results are then reported in Chapters 5-9, followed by a general conclusion.

Chapter 2 is devoted to the characterization of probability distributions using the tools of information theory. We treat discrete and continuous probability distributions separately, and define the Shannon entropy accordingly. We then focus on the Shannon differential entropy for continuous

probability distributions, and give an overview of its main properties. After that, we pay a particular attention to Gaussian probability distributions, which leads us to introduce the entropy power inequality. We then present a family of relevant information-theoretic measures, namely p-norms and Rényi entropies. We conclude this chapter with some simple quantum mechanical applications of these notions.

Once the probability distributions and uncertainty measures have been introduced, we address the theory of majorization in Chapter 3. We provide a mathematical overview of majorization relations for the different possible types of distributions considered. We do not limit ourselves to probability distributions, but consider very general distributions that simply have to satisfy absolute convergence conditions. Such distributions can thus be both positive and negative. We start by discussing discrete and continuous majorization for distributions defined on a finite domain, while preparing the grounds for distributions defined on an infinite domain. We then turn to discrete and continuous infinite-domain distributions, and expose which majorization conditions can still be applied to such distributions. Note that the reader who is familiar with the theory of majorization can be satisfied with a quick overview of this chapter. Moreover, we will recall these tools in the context of Wigner functions at the beginning of Chapter 6.

In Chapter 4, we introduce the physics background on which the thesis relies, namely quantum optics. We define several families of quantum optical states, which are based on the quadrature basis, the Fock basis, and the coherent basis. Then, we turn to linear optical transformations and define them in terms of mode operators as well as Wigner functions. We present the common single-mode transformations, followed by the two-mode transformations effected by a beam-splitter and two-mode squeezer. The last part of this Chapter is devoted to the broad family of quantum states called Gaussian states. This leads us to define the covariance matrix, which, together with the displacement vector, fully characterizes any Gaussian state. We consider Gaussian pure and mixed states, in particular thermal states. At this point, we have defined the main mathematical and physical notions that we will use throughout this work, so we move on to the outcomes of our thesis.

In Chapter 5, we address the long-standing problem of characterizing the set of quantum states with a non-negative Wigner function. This problem is of particular interest to us here because Wigner-positive states constitute states for which every classical uncertainty measure is well defined. We show that, among the Wigner-positive states, a large subset of such states that can be constructed by using a balanced beam-splitter. We then choose to focus our interest on phase-invariant quantum states that contain up to two photons, and provide an in-depth analysis of the corresponding subset. This allows us to draw conclusions about the structure of the Wigner-positive quantum set.

In Chapter 6, we tackle the primary objective of this thesis, namely the application of continuous majorization in phase space. With continuous majorization extended to infinite-domain distributions, we address the question of determining the state of least disorder in phase-space. First, we carry out a reasoning implying that pure states are, in general, either incomparable or equivalent as far as majorization is concerned. We then observe that the negative Wigner volume in phase space is a Schur-convex measure, which can reach arbitrarily high values. This leads us to formulate a majorization conjecture restricted to the Wigner-positive set, stating that pure Gaussian states majorize any Wigner-positive state. We present some numerical simulations supporting the conjecture, and prove it over the subset of phase-invariant states containing up to two photons that was characterized in Chapter 5. Finally, we conclude this Chapter by defining two special extensions of majorization, namely radial-majorization and square-majorization.

In Chapter 7, we come back to the question of comparing states in terms of measures of disorder, exploiting our findings on majorization in phase space as exposed in the previous Chapter. We first determine which measures of disorder are consistent with Wigner functions that are partly negative. In the case of Wigner-positive states, we define the Wigner entropy as the Shannon differential entropy of their Wigner function. We establish its properties and conjecture a lower-bound which follows on the majorization conjecture laid down in the previous Chapter. We then provide a in proof of this lower-bound on the subset of passive states. Finally, we define Wigner-Rényi entropies and illustrate their interest in the context of inequalities for beam-splitters and two-mode squeezers.

Then, Chapter 8 is devoted to the application of majorization theory for quantum bosonic channels.

In particular, we consider Gaussian phase-insensitive channels, which are quite common in continuous-variable quantum information and quantum optics. First, we show that for some particular values of its parameters, such a channel always yields a Wigner-positive output. We then prove a continuous majorization relation between the input and output Wigner functions for channels with a gain parameter greater than 1. Finally, our main contribution is to prove a ladder of (discrete) majorization relations between the (infinite) vector of eigenvalues associated with the output for several possible inputs. Namely, we show that the vector of eigenvalues associated with the output of the n^{th} Fock states majorizes the one associated with the output of the $n + 1^{\text{th}}$ Fock state. This very general property of all Gaussian phase-insensitive channels gives us information on the entropy (or disorder) that is produced by such channels.

Finally, Chapter 9 is devoted to a resource theory for Gaussian work extraction in a multipartite setting in quantum thermodynamics. This result heavily relies on symplectic transformations in phase space, so we first introduce the symplectic formalism which allows us to describe the evolution of the covariance matrix through linear canonical transformations. We also present Williamson's theorem and the Bloch-Messiah decomposition. Then, we exploit these notions to approach the work that can be extracted from a multipartite system by use of a local Gaussian unitary assisted with a global energy-preserving Gaussian unitary (i.e., a linear interferometer). We show how to take advantage of the symplectic formalism in order to develop a resource theory for the extractable work. Most notably, the underlying resource is associated with a monotone that finds a very natural expression in the symplectic formalism as the difference between the trace and symplectic trace of the covariance matrix.

Chapter 1

From classical to quantum phase-space

This first chapter is intended to sketch the basics of quantum mechanics used in this work and to present the general guideline of this work. As explained in the introduction, our ambition is to focus on the notion of disorder in quantum phase space. First, we will present the phase space from a completely classical approach. In this way we will introduce the coordinates and the generalized momenta associated to each degree of freedom of a classical system. This will lead us to present the phase space as the privileged place to represent the evolution of a physical system.

Following this, we will move to quantum mechanics, and introduce states with continuous variables defined according to degrees of freedom taking a continuum of values. We take here a general approach, and do not specify the nature of these degrees of freedom. We introduce mixed states and the density matrix formalism.

We will then introduce quantum phase space as the quantum counterpart of classical phase space. In this context, we will define the Wigner functions that constitute the focus of our research. We will present their properties as well as their quantum peculiarity.

1.1 Statistical mechanics perspective

In a classical vision of mechanics, the configuration of a physical system is fully described by a vector of real numbers, each of which corresponding to the configuration of its elements. They can for example corresponds to their position along some axis. Let us consider a physical system with N continuous degrees of freedom defined over the real axis. All the information about the physical system is encoded in its generalized coordinates $\mathbf{x}$:

$$\mathbf{x} = (x_1, x_2, ..., x_N) \tag{1.1}$$

where each component x_i corresponds for example to the position of one of its component along some axis. Any vector $\mathbf{x} \in \mathbb{R}$ corresponds to a precise configuration of the system.

Newtonian mechanics tells us that the movement of any body is governed by the relation $F = ma$. As a consequence, it suffices to know the exact position and momentum of a particle at some time to know it at any future and past time. If we know the vector of generalized velocities $\dot{\mathbf{x}} = (\dot{x}_1, \dot{x}_2, ..., \dot{x}_N)$, we can predict the evolution of the system for all time, where we use the notation $\dot{x} = \partial x/\partial t$.

The equations of motion of the system can be formulated in an elegant way by introducing two quantities. The Lagrangian L of the system is defined as the difference between the kinetic energy T and potential energy V, so that we have $L = T - V$. The Hamiltonian H of the system is the total energy of the system, defined as $H = T + V$. The generalized momentum are defined as $p_i = \partial L/\partial \dot{x}_i$, which enables us to define the momentum vector $\mathbf{p}$:

$$\mathbf{p} = (p_1, p_2, ..., p_N) \tag{1.2}$$

When considering a free particle of mass m, the generalized momentum is simply $p = m\dot{x}$. Any vector $\mathbf{p} \in \mathbb{R}$ is an acceptable momentum vector. In general, the coordinates vector $\mathbf{x}$ and the momentum

vector $\mathbf{p}$ define the evolution of the system at all time. This appears from the equations of motion of Hamilton which links the respective evolution of $\mathbf{x}$ and $\mathbf{p}$.

$$\begin{cases} \dot{x}_i = \dfrac{\partial H}{\partial p_i} \\ \dot{p}_i = -\dfrac{\partial H}{\partial x_i} \end{cases} \tag{1.3}$$

The laws of motion are fully characterized by these equations. The phase space is defined as the set of possible configurations of $(\mathbf{x}, \mathbf{p})$. As it evolves, the system traces its path in the phase space. If we know the initial position and momentum, we will know all the future and all the past of the system. Any point in phase space is associated to a precise trajectory. We define the configuration vector $\mathbf{q}$ as the concatenation of the coordinate vector $\mathbf{x}$ and the momentum vector $\mathbf{p}$, and we write $\mathbf{q} = (\mathbf{x}, \mathbf{p})$.

In statistical mechanics, it is common that we don't know exactly the state of the system, but rather we know that it is in a macrostate that corresponds to different possible microstates. A microstate is a precise value of $\mathbf{x}$ and $\mathbf{p}$. A macrostate is characterized by a density function $\rho(\mathbf{x}, \mathbf{p})$. The density functions is a true probability distribution over phase-space, as it is non-negative and normalized to 1, namely

$$\rho\left(\mathbf{x}, \mathbf{p}\right) \geq 0 \qquad\qquad \iint \rho\left(\mathbf{x}, \mathbf{p}\right) \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p} = 1. \tag{1.4}$$

The density function ρ describes a system that is in a statistical mixture. We have a statistical uncertainty about the exact configuration of the state, so that we can only approximate it with a probability distribution. Hence, one can recover the probability that the system is in a particular subset of phase space by integrating the density function over that region. The resulting quantity corresponds to the probability that the state is in that configuration. Let Q be a subset of phase-space, then the probability that the macrostate is in that subset can be computed as:

$$\Pr\left[\mathbf{q} \in Q\right] = \int\limits_{\mathbf{q} \in Q} \rho(\mathbf{q}) \mathrm{d}\mathbf{q} \tag{1.5}$$

As a consequence, one can recover the probability distribution for the coordinates and momentum as the integration over the other variables:

$$\rho_x(\mathbf{x}) = \int \rho(\mathbf{x}, \mathbf{p}) \mathrm{d}\mathbf{p}, \qquad\qquad \rho_p(\mathbf{p}) = \int \rho(\mathbf{x}, \mathbf{p}) \mathrm{d}\mathbf{x}. \tag{1.6}$$

As such, the physical quantities associated with this system can be computed as their statistical average over each microstate. Let A be a quantity that only depends on the coordinates and momentum of the system. The mean value of any property A fo the state can then be evaluated as the mean of the value it takes on any microstate $\mathbf{q} = (\mathbf{x}, \mathbf{p})$, so that we have:

$$\langle A \rangle = \iint A\left(\mathbf{x}, \mathbf{p}\right) \rho\left(\mathbf{x}, \mathbf{p}\right) \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p}. \tag{1.7}$$

Finally, let us mention that the evolution of the density function is governed by Liouville's theorem:

$$\frac{\partial}{\partial t}\rho = (\{H, \rho\}) \tag{1.8}$$

where $(\{\bullet, \bullet\})$ denotes the Poisson brackets. This equation implies that the volume of the density function is conserved over time. It can be understood as the fact that each infinitesimal volume of probability in the density function is following its own trajectory.

1.2 Continuous-variable quantum states

The theory of quantum mechanics introduces operators to represent physically observable quantities. In that manner, an operator $\hat{x}_i$ is associated to each of the N degrees of freedom of the quantum system. Conversely, each momentum is associated to a quantum operator $\hat{p}_i$. Each momentum operator $\hat{p}_i$ is canonically conjugated to its corresponding coordinate $\hat{x}_i$. We have the relation $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$. We define the coordinate operator vector $\hat{\mathbf{x}}$ and momentum operator vector $\hat{p}$ as follows:

$$\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, ..., \hat{x}_N), \qquad \hat{\mathbf{p}} = (\hat{p}_1, \hat{p}_2, ..., \hat{p}_N). \tag{1.9}$$

1.2.1 Pure states

A continuous-variable pure state is then described by a complex wave-function $\psi(\mathbf{x})$, which associates to each coordinate a complex amplitude of probability. The wave-function ψ is normalized in the sense that the integral of its squared modulus gives 1. In general, any normalized complex-valued function that is continuously differentiable (of class C^1) is a physically acceptable wave-function.

Similarly, a continuous-variable pure state possesses a momentum wave-function $\phi(\mathbf{p})$ that associate a complex amplitude of probability to each momentum. However, these two wave-function cannot be chosen independently from each other. These two wave-functions are linked through a Fourier transformation.

$$\phi(\mathbf{p}) = \frac{1}{(2\pi\hbar)^{\frac{N}{2}}} \int \psi(\mathbf{x}) \exp\left(-\frac{i}{\hbar}\mathbf{p}\mathbf{x}\right) \mathrm{d}\mathbf{x} \tag{1.10}$$

$$\psi(\mathbf{x}) = \frac{1}{(2\pi\hbar)^{\frac{N}{2}}} \int \phi(\mathbf{p}) \exp\left(\frac{i}{\hbar}\mathbf{x}\mathbf{p}\right) \mathrm{d}\mathbf{p} \tag{1.11}$$

The probability distribution associated to each coordinate is then linked to the squared of the absolute value of $\psi(\mathbf{x})$, while the probability distribution associated to each momentum is linked to the squared absolute value of $\phi(\mathbf{p})$:

$$\rho_x(\mathbf{x}) = |\psi(\mathbf{x})|^2, \qquad \rho_p(\mathbf{p}) = |\phi(\mathbf{p})|^2. \tag{1.12}$$

This is very different compared to the classical case. Indeed, in the classical case the probability densities ρ_x and ρ_p are completely disconnected from each other. One could imagine a state with any coordinate probability density ρ_x and any momentum probability density ρ_p. In quantum mechanics, these two probability densities are intrinsically linked, as they arise from the same object. Indeed, the coordinate wave-function and the momentum wave-function are two sides of a same coin, since they are related by a Fourier transform.

The Fourier transformation originates from the fact that coordinate and momentum are canonically conjugate. Indeed, in quantum mechanics, two observables that do not commute cannot take simultaneously a determined value.

Furthermore, there is an additional degree of freedom since the wave-function itself is a complex object. The phase of the wave-function is not accessible in itself, as probability densities are evaluated by taking the square modulus, but it plays a role in the Fourier transformation.

Here, both the position and momentum information are encoded in the same object, which is a complex wave-function. We will use more generally the Dirac notation, where a quantum state is represented by a vector $|\psi\rangle$. The vector $|\psi\rangle$ belongs to a infinite-dimensional Hilbert space $\mathcal{H}$. The coordinate wave-function and the momentum wave-function are then respectively the projection of $\langle\mathbf{x}|\psi\rangle$ and $\langle\mathbf{p}|\psi\rangle$. The states $|\mathbf{x}\rangle$ and $|\mathbf{p}\rangle$ designate a non-physical state with an infinitely focused coordinate and momentum, respectively.

The expectation value of an operator $\hat{A}$ over a pure state $|\psi\rangle$ is then written as

$$\langle\hat{A}\rangle = \langle\psi|\,\hat{A}\,|\psi\rangle\,. \tag{1.13}$$

It can be formulated in both coordinate and momentum representation, using the relation $\langle x|\,\hat{p}\,|\psi\rangle = -i\hbar(\partial/\partial x)\,\langle x|\psi\rangle$ and $\langle p|\,\hat{x}\,|\psi\rangle = i\hbar(\partial/\partial p)\,\langle p|\psi\rangle$. Notice that $\rho_x(\mathbf{x}) = \langle\psi|\,\hat{\mathbf{x}}\,|\psi\rangle$ and $\rho_p(\mathbf{p}) = \langle\psi|\,\hat{\mathbf{p}}\,|\psi\rangle$.

The evolution of a quantum state is governed by the Hamiltonian of the system, which is an operator $\hat{H}$. The Schrödinger equation define the time evolution of $|\psi\rangle$ as follows:

$$i\hbar\frac{\partial}{\partial t}|\psi\rangle = \hat{H}|\psi\rangle . \tag{1.14}$$

1.2.2 Mixed states

The states we have presented until now are labeled as pure states, as they are defined by a unique wave-function, which is according to quantum theory the utmost precise description we can have of a quantum state. Sometimes there is a statistical uncertainty about the quantum state, which can be in several pure states in a probabilistic fashion. Such states are called mixed states and are appropriately described in the density matrix formalism. The density operator $\hat{\rho} : \mathcal{H} \mapsto \mathcal{H}$ is defined as follows:

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle \langle\psi_i| \tag{1.15}$$

where $\{|\psi_i\rangle\}$ is a set of pure states, such that our system is in state $|\psi_i\rangle$ with probability p_i. In general, any quantum operator is a physically acceptable density operator if it satisfies three physicality conditions. It should be normalized, hermitian and positive semi-definite:

$$\mathrm{Tr}\left[\hat{\rho}\right] = 1 \tag{1.16}$$

$$\hat{\rho}^\dagger = \hat{\rho} \tag{1.17}$$

$$\hat{\rho} \geq 0 \tag{1.18}$$

Note that (1.18) can also be expressed as $\langle\psi|\hat{\rho}|\psi\rangle \geq 0 \ \forall |\psi\rangle$. Equivalently, the density operator $\hat{\rho}$ is positive semi-definite if and only if $\mathrm{Tr}\left[\hat{\rho}\hat{\psi}\right] \geq 0$ for any pure state $\hat{\psi} = |\psi\rangle\langle\psi|$. The three physicality conditions can be reduced to conditions on the eigenvalues of $\hat{\rho}$, which should sum up to 1, be real-valued and non-negative. These three properties can be understood as the fact that $\hat{\rho}$ derives from a true probability distribution.

The expectation value of an observable $\hat{A}$ over a quantum state $\hat{\rho}$ can be computed as the weighted average of the mean of $\hat{A}$ over each pure states $\{|\psi_i\rangle\}$. This expression can be written in a simple writing as follows:

$$\langle\hat{A}\rangle_{\hat{\rho}} = \mathrm{Tr}\left[\hat{A}\hat{\rho}\right] \tag{1.19}$$

In practice, we will sometimes simply write $\langle\hat{A}\rangle$ when the state over which we compute the expected value is clear from the context.

The time evolution of a mixed states is governed by von Neumann equation, which reads as follows:

$$i\hbar\frac{\partial}{\partial t}\hat{\rho} = \left[\hat{H}, \hat{\rho}\right], \tag{1.20}$$

where the notation $[\bullet, \bullet]$ denotes the commutator. This equation can directly be derived from (1.14).

1.3 Wigner functions in quantum phase-space

In classical physics, the existence of a phase-space probability distribution does not pose any conceptual problem since the particles have simulatenously well determined coordinates and momenta. In that case indeed, it is nothing more than the joint probability distribution for each coordinate and momentum, which contains the information about each parameter and their correlations. On the quantum side this interpretation completely falls apart. Indeed, we know that observables that do not commute cannot be simultaneously defined. The existence of a joint distribution appears then as a non-sense. However, we are going to show that it is possible to construct an object with properties that are surprisingly close to a joint probability distribution.

The Wigner function of a state with N continuous degrees of freedom is a $2N$-dimensional function $W : \mathbb{R}^{2N} \mapsto \mathbb{R}$. It can be equivalently defined from the coordinate basis or the momentum basis. The Wigner function of a pure quantum state with coordinate wave-function $\psi(\mathbf{x})$ and momentum wave-function $\phi(\mathbf{p})$ is defined as follows:

Definition 1.1 (Wigner function of a pure state)**.** *The Wigner function of a N-dimensional continuous-variable quantum state with coordinate wave-function $\psi(\mathbf{x})$ and momentum wave-function $\phi(\mathbf{p})$ is the distribution $W : \mathbb{R}^{2N} \mapsto \mathbb{R}$ defined as:*

$$W(\mathbf{x},\mathbf{p}) = \frac{1}{(2\pi\hbar)^N} \int \psi^* \left(\mathbf{x} + \frac{\mathbf{y}}{2}\right) \psi \left(\mathbf{x} - \frac{\mathbf{y}}{2}\right) \exp\left(\frac{i}{\hbar}\mathbf{p}\mathbf{y}\right) \mathrm{d}\mathbf{y} \tag{1.21}$$

$$= \frac{1}{(2\pi\hbar)^N} \int \phi^* \left(\mathbf{p} + \frac{\mathbf{q}}{2}\right) \phi \left(\mathbf{p} - \frac{\mathbf{q}}{2}\right) \exp\left(-\frac{i}{\hbar}\mathbf{x}\mathbf{q}\right) \mathrm{d}\mathbf{q} \tag{1.22}$$

The Wigner function comes as close as quantum mechanics allows to a joint probability distribution for the coordinates and momentum observables. To illustrate this, let us notice that the marginal of the Wigner function corresponds to true probability distributions

$$\int W(\mathbf{x},\mathbf{p})\mathrm{d}\mathbf{p} = |\psi(\mathbf{x})|^2, \qquad \int W(\mathbf{x},\mathbf{p})\mathrm{d}\mathbf{x} = |\phi(\mathbf{p})|^2. \tag{1.23}$$

However, there are also intrinsic quantum features to the Wigner function, which make it a very particular object. The most striking feature is that it is in general partly negative. The negativity of the Wigner function, however, is in some sense limited. Indeed, the marginal should always result in non-negative distribution, as these are true probability densities. Also, the negative regions cannot exceed some extent in phase space. Typically they cannot exceed a region of area $\hbar$. This phenomenon will appear clearer later with the following observation. The squared modulus of the scalar poduct of two pure states is directly related to the integral over all phase space of the product of their Wigner functions:

$$|\langle\psi_1|\psi_2\rangle|^2 = (2\pi)^N \iint W_1(\mathbf{x},\mathbf{p}) W_2(\mathbf{x},\mathbf{p}) \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p} \tag{1.24}$$

As their exist quantum states with positive Wigner function which is concentrated over a region of $\hbar$ in phase space, it appears that a state which is negative over that region would violate Eq. (1.24). The negativity of the Wigner function have to be localized.

Another very non-classical features is that the Wigner function is upper and lower-bounded. This can be seen by applying the Cauchy-Riemann inequality on Eq. (1.22), which yields the inequality:

$$|W(\mathbf{x},\mathbf{p})| \leq \frac{1}{(\pi\hbar)^N}. \tag{1.25}$$

As a consequence, the Wigner function cannot be arbitrarily peaked. Completely determined distribution such as it was possible to imagine classically are thus forbidden.

Figure 1.1 illustrates the observations that we have mentioned for a random pure state. We see that the Wigner function has negative pockets with limited extend, and such that the marginal distributions are non-negative.

Until now, we have limited our observations to pure states. In order to give a complete description of quantum mechanics in phase-space, we need to define how are represented quantum operators in phase-space. That step is taken thanks to the Weyl transfrom, which maps any quantum operator $\hat{A}$ to a phase-space distribution $A(\mathbf{x},\mathbf{p})$.

Definition 1.2 (Weyl transform)**.** *The Weyl transform is the map $\mathcal{T}_\mathrm{w} : \mathcal{H}^\mathcal{H} \mapsto \mathbb{C}^{\mathbb{R}^{2N}} : \hat{A} \mapsto A$ where $A(\mathbf{x},\mathbf{p})$ is defined as follows:*

$$A(\mathbf{x},\mathbf{p}) = \frac{1}{(2\pi\hbar)^N} \int \left\langle \mathbf{x} - \frac{\mathbf{y}}{2} \right| \hat{A} \left| \mathbf{x} + \frac{\mathbf{y}}{2} \right\rangle \exp\left(\frac{i}{\hbar}\mathbf{p}\mathbf{y}\right) \mathrm{d}\mathbf{y} \tag{1.26}$$

$$= \frac{1}{(2\pi\hbar)^N} \int \left\langle \mathbf{p} - \frac{\mathbf{q}}{2} \right| \hat{A} \left| \mathbf{p} + \frac{\mathbf{q}}{2} \right\rangle \exp\left(-\frac{i}{\hbar}\mathbf{x}\mathbf{q}\right) \mathrm{d}\mathbf{q} \tag{1.27}$$

We will write the Weyl transformation operator as $\mathcal{T}_\mathrm{w}$ and the inverse map as $\mathcal{T}_\mathrm{w}^{-1}$.

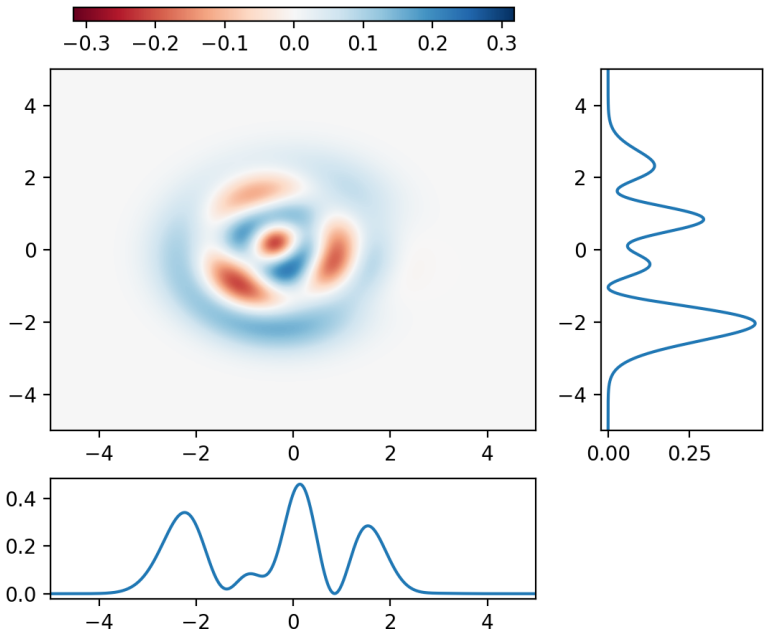

Figure 1.1: Wigner function of a random pure state. The Wigner function exhibits positive and negative values. The blue parts correspond to positive areas while the red parts correspond to negative areas. However, its marginals are always non-negative.

The Weyl transform has the useful property to translate the trace of a product of operators into the integral of a product of phase-space distributions. That property is sometimes called the overlap formula [42] and is the translation of Eq. (1.24). For any two operators $\hat{A}_1, \hat{A}_2$ with respective Weyl transforms A_1, A_2 it reads as follows:

$$\mathrm{Tr}\left[\hat{A}_1\hat{A}_2\right] = (2\pi)^N \iint A_1(\mathbf{x},\mathbf{p})A_2(\mathbf{x},\mathbf{p})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p} \tag{1.28}$$

The overlap formula captures the essence of what makes the Weyl transform so special. It translates a trace product of operators into a phase-space product. The former is ubiquitous in the operators formulation of quantum mechanics, while the latter recalls statistical physics.

The Wigner function of the pure state $|\psi\rangle$ is the Weyl transform of the projector $|\psi\rangle\langle\psi|$. By linearity of the Weyl transform, we see that the Wigner function of a mixed states is the weighted average of the Wigner functions of the pure states from which it is built. We can simply write $W = \sum p_i W_i$, where $\{W_i\}$ are the Wigner functions of the pure states $\{|\psi_i\rangle\}$ which is the statistical set associated to $\hat{\rho}$ with respective probabilities $\{p_i\}$. In a more direct way, one can simply find the Wigner function of a mixed state by computing the Weyl transform of its density operator, as $W = \mathcal{T}_{\mathrm{w}}[\hat{\rho}]$.

From that observation, it follows that the observation we had laid for pure states extend to mixed states. In general, any distribution $W(\mathbf{x},\mathbf{p})$ defined over phase-space is physically acceptable as a quantum state if it satisfies the three following conditions:

$$\iint W(\mathbf{x},\mathbf{p})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p} = 1 \tag{1.29}$$

$$W(\mathbf{x},\mathbf{p}) \in \mathbb{R} \tag{1.30}$$

$$\iint W(\mathbf{x},\mathbf{p})W_\psi(\mathbf{x},\mathbf{p})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p} \geq 0 \qquad \forall\,|\psi\rangle \tag{1.31}$$

Conditions (1.29), (1.30) and (1.31) are the analog of respectively (1.16), (1.17) and (1.18). Note that in phase-space, it appears to be difficult to determine whether a distribution corresponds to a positive semi-definite operator. Indeed, Condition (1.31) should be checked for any pure Wigner function W_ψ,

and the task is difficult in general as the Hilbert space is infinite dimensional. Let us compute the expectation value of an operator $\hat{A}$ over a state described by a Wigner function W. This can be done directly from the overlap formula (1.28) and we find:

$$\langle \hat{A} \rangle = (2\pi)^N \iint A(\mathbf{x}, \mathbf{p}) W(\mathbf{x}, \mathbf{p}) \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p} \tag{1.32}$$

That equation should be compared with Eq. (1.7). Both equation are almost identical, which explains why the Wigner function is such a powerful tool. It gives a description of quantum states in a formalism that uses the same tools as statistical mechanics, and provides as a consequence a powerful and intuitive interpretation of quantum mechanics. However, this interpretation is done at the cost of partly negative phase distributions.

Let us finally present the equation describing the time evolution of the Wigner function evolving according to a given Hamilatonian.

$$\frac{\partial}{\partial t} W = \{\{H, W\}\} \tag{1.33}$$

where $H = \mathcal{T}_{\mathrm{w}}\left[\hat{H}\right]$ is the Weyl transform of the Hamiltonian and $\{\{\bullet, \bullet\}\}$ denotes the Moyal brackets. We will not dwell on the above expression. Moyal brackets rely on the Moyal product which is cumbersome to use. We mention Equation (1.33) to show that a complete self-consistent formulation of quantum mechanics in phase space exists. We will see that for usual Hamiltonians, the time evolution of the Wigner function is easily described and corresponds to simple operations in phase space.

Chapter 2

Elements of information theory

with a touch of quantum

In the previous chapter we introduced the quantum distribution that is the Wigner function. As we have seen, it has properties that make it intrinsically different from a probability distribution. However, the Wigner function also shares some properties with the latter, and in fact plays a role in quantum mechanics very similar to that of a probability distribution. It is for this reason that in this chapter we are interested in probability distributions and their properties. In this context, information theory is full of particularly powerful tools to measure the information content of distributions.

First, we will look at discrete and continuous random variables, and we will present how they are represented from a mathematical point of view. In this way, we define the probability vectors and the probability density distributions.

Then, we introduce the central tool in information theory that is Shannon entropy. We define the Shannon discrete entropy of a probability vector and the Shannon differential entropy of a probability density function. We then highlight several particularly useful properties of Shannon differential entropy.

After that, we focus our interest on Gaussian distributions and highlight several of their properties. This leads us to introduce the entropy power inequality. We then define some other information-theoretical measures. Finally, we conclude this chapter by applying some of the notion we have defined to quantum mechanics.

2.1 Random variables and probability distributions

Depending on the situation they describe, random variables can take different forms. A random variable is said to be discrete if takes a discrete number of values. A discrete random variable X is then completely described by its probability mass function p_X, which associate to any outcome its corresponding probability of occurrence. The probability mass function is defined as follows:

Definition 2.1 (Probability mass function)**.** *The probability mass function of a random variable X is the function p_X that associates to each outcome the corresponding probability of occurrence.*

$$p_X(x_i) = \Pr\left[X = x_i\right]. \tag{2.1}$$

The probability mass function is the appropriate tool to characterize discrete random variables.

When X is a discrete variable, p_X is non-negative only for a discrete number of x_i, namely the set $\{x_i\}$ which is the set of possible outcome of the discrete variable X.

Some situations, however, are more appropriately described by a continuous random variable. The outcome of such variables belongs to a continuous set. The probability mass function is then a poorly efficient tool to describe them, as the probability to measure a precise occurrence is vanishingly small. We will rather use the probability density function f_X to describe them. It is defined as follows:

Definition 2.2 (Probability density function)**.** *The probability density function of a random variable X is the function f_X that associate to each outcome the corresponding density of probability.*

$$\Pr\left[a \leq X \leq b\right] = \int_a^b f_X(x)\mathrm{d}x \tag{2.2}$$

The probability density function is the appropriate tool to characterize continuous random variables.

The probability mass function is an appropriate tool to characterize discrete random variable, while the probability density function is more convenient for continuous random variable. Note however that from an abstract mathematical point of view, one could associate a probability density function f_X with a discrete random variable X as:

$$\int_x^{x+\Delta} f_X(y)\mathrm{d}y \;\approx\; p_X(x) \;\approx\; \Delta \cdot f_X(x) \tag{2.3}$$

where $\Delta \to 0^+$ is a non-negative vanishingly small quantity. From that relation one understands that these two distribution mutually define each other, and that only one of the two is relevant.

One can understand that a continuous random variable is intrinsically more disordered than a discrete random variable. Indeed, the outcome of a discrete random variable belongs to a discrete set, which means that we already have information about the outcome. However, a continuous random variable belongs to a continuous set, which is a different type of infinity.

In practice, we will use vectors to represent discrete probability distributions, and functions to represent continuous probability distributions. We introduce the probability mass vector as a vectorized version of the probability mass function. We define $\mathbf{p}_X = (p_X(x_1), p_X(x_2), ...)$. We will see that for discrete random variables, the probability mass vector $\mathbf{p}_X$ encloses what we need to characterize its information content. A vector $\mathbf{x}$ is an acceptable probability distribution if it is normalized and non-negative. We will consider vectors of dimension N, which is possibly infinite.

Definition 2.3 (Probability vector)**.** *A probability vector is a vector $\mathbf{p} \in \mathbb{R}^N$ which satisfies the normalization and non-negativity conditions:*

$$\sum_{i=1}^{N} p_i = 1, \qquad \textit{and} \qquad p_i \geq 0 \qquad \forall i. \tag{2.4}$$

This generalizes to vectors $\mathbf{p} \in \mathbb{R}^{\mathbb{N}}$ in which case the summation is carried until ∞.

Similarly, a function f is an acceptable probability density function if it is normalized and non-negative. We will use probability distribution and probability density function as synonyms. Note that we will sometimes be confronted to multidimensional probability distribution, so that to be more general we will consider function of a n-dimensional vector $\mathbf{r} \in \mathbb{R}^n$. This is the case for example of a vectorial continuous random variable $\mathbf{X} = (X_1, X_2, ..., X_n)$.

Definition 2.4 (Probability density function)**.** *A probability density function is a distribution $f : \mathbb{R}^n \mapsto \mathbb{R}$ which satisfies the normalization and non-negativity conditions:*

$$\int_{\mathbb{R}^n} f(\mathbf{r})\mathrm{d}\mathbf{r} = 1, \qquad \textit{and} \qquad f(\mathbf{r}) \geq 0 \qquad \forall \mathbf{r} \in \mathbb{R}^n. \tag{2.5}$$

Note that to be even more general we can consider a function f defined over a subset $\mathbb{A} \subseteq \mathbb{R}^n$, in which case the integration is performed over $\mathbb{A}$ and the non-negativity condition is checked for all $\mathbf{r} \in \mathbb{A}$. In practice we will often simply refer to probability density distributions as probability distributions.

2.2 Shannon differential entropy and its properties

The standard information-theoretic tool to evaluate the information content of a probability distribution is its Shannon entropy. It benefits indeed from interesting properties that we will explain in this section. What is remarkable about Shannon entropy is that it only depends on the values taken by the probability distributions. As such, it is invariant under a relabeling of the outcomes and only depends on the probability vector or probability distributions associated to the random variable.

As we have seen, a random variable can behave very differently whether it is discrete or continuous. For that reason, the definition of entropy that we use is different in each two cases, depending on the mathematical tools that we are using.

Definition 2.5 (Shannon entropy of a probability vector)**.** *The Shannon (discrete) entropy of a probability vector* $\mathbf{p}$ *is the quantity* $H(\mathbf{p})$ *defined as:*

$$H(\mathbf{p}) = -\sum_{i=1}^{N} p_i \ln p_i \tag{2.6}$$

The Shannon discrete entropy of a discrete random variable X is $H(X) = H(\mathbf{p}_X)$ where $\mathbf{p}_X$ is the probability vector associated to X. Sometimes the notation $H(p_X)$ is used, when p_X is a probability mass function that takes non-zero values only over a discrete domain. In that case we define $H(p_X)$ to be equal to $H(\mathbf{p}_X)$. The capital letter H is used when the discrete entropy is computed.

Definition 2.6 (Shannon entropy of a probability density distribution)**.** *The Shannon (differential) entropy of a probability density distribution* f *is the quantity* $h(f)$ *defined as:*

$$h(f) = -\int f(\mathbf{r}) \ln f(\mathbf{r}) \mathrm{d}\mathbf{r} \tag{2.7}$$

The Shannon differential entropy of a continuous variable X is the entropy of its probability density function, so that we have $h(X) = h(f_X)$.

Note that from an abstract mathematical point of view, it is possible to compute the continuous entropy of a discrete variable, or conversely the discrete entropy of a continuous random variable. From Eq. (2.3) and the two definitions of Shannon entropy that we have introduced, the following relation can be derived:

$$h(X) = H(X) + \ln(\Delta) \tag{2.8}$$

where $\Delta \to 0^+$ is an infinitely small vanishing quantity which describes the discrete sampling of a continuous function [16].

In this thesis, we are interested in systems that are described by continuous probability distributions. For that reason, we give a particular interest to Shannon differential entropy. In the following subsections, we are going to highlight several if its useful properties.

2.2.1 Concavity of entropy

Definition 2.7 (Concave function)**.** *A function* $\varphi : \mathbb{A} \mapsto \mathbb{R}$ *is concave over* $\mathbb{A} \subseteq \mathbb{R}$ *if and only if:*

$$\varphi(p_1 x_1 + p_2 x_2) \geq p_1 \varphi(x_1) + p_2 \varphi(x_2) \qquad \forall x_1, x_2 \in \mathbb{A} \tag{2.9}$$

where (p_1, p_2) *is a probability vector. Conversely,* φ *is convex over* $\mathbb{A}$ *if and only if* $-\varphi$ *is concave over* $\mathbb{A}$*.*

A sufficient condition to "φ is concave over $\mathbb{A}$" is that the second derivative of φ is negative over $\mathbb{A}$. Convex and concave functions have a particular interest when deriving inequalities. Indeed, they are subject to the Jensen's inequality [18]. Jensen's inequality can be stated in various ways, such as $\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)]$. In a measure-theoretic from we can write $\varphi\left(\int g \mathrm{d}\mu\right) \leq \int \varphi(g) \mathrm{d}\mu$. When associating $\mathrm{d}\mu = f(\mathbf{r})\mathrm{d}\mathbf{r}$, we find the version that we will use in this thesis.

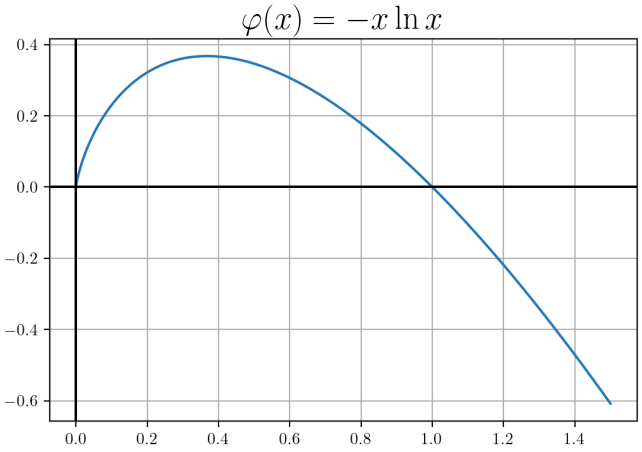

Figure 2.1: Graph of the function $\varphi(x) = -x \ln x$ for $x \geq 0$. When $x \to 0^+$, $\varphi(x) \to 0^+$. When $x \to \infty$, $\varphi(x) \to -\infty$. It clearly appears that φ is concave on $\mathbb{R}_+$.

Definition 2.8 (Jensen's inequality)**.** *Let $\varphi : \mathbb{A} \mapsto \mathbb{R}$ be concave over $\mathbb{A} \subseteq \mathbb{R}$. Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a proability density distribution. Let $g : \mathbb{R}^n \mapsto \mathbb{A}$ be any real-valued distribution. Then, the following inequality holds:*

$$\varphi\left(\int g(\mathbf{r})\ f(\mathbf{r})\ \mathrm{d}\mathbf{r}\right)\ \geq\ \int \varphi\Big(g(\mathbf{r})\Big)\ f(\mathbf{r})\ \mathrm{d}\mathbf{r} \tag{2.10}$$

The inequality is reversed when φ is convex.

Let us define the function $\varphi(x) = -x \ln x$, which is the function that appears in the integrand of Eq. (2.7). A simple graph shows us that φ is concave over $\mathbb{R}_+$, see Fig. 2.1. Let $f_1(\mathbf{r})$ and $f_2(\mathbf{r})$ be two probability distributions, and let $f(\mathbf{r}) = p_1 f_1(\mathbf{r}) + p_2 f_2(\mathbf{r})$, where (p_1, p_2) is a probability vector.

$$\begin{aligned}
h(f) &= \int \varphi\left(f(\mathbf{r})\right) \mathrm{d}\mathbf{r} && (2.11)\\
&= \int \varphi\left(p_1 f_1(\mathbf{r}) + p_2 f_2(\mathbf{r})\right) \mathrm{d}\mathbf{r} && (2.12)\\
&\geq \int \left(p_1 \varphi\left(f_1(\mathbf{r})\right) + p_2 \varphi\left(f_2(\mathbf{r})\right)\right) \mathrm{d}\mathbf{r} && (2.13)\\
&= p_1 \int \varphi(f_1(\mathbf{r}))\mathrm{d}\mathbf{r} + p_2 \int \varphi(f_2(\mathbf{r}))\mathrm{d}\mathbf{r} && (2.14)\\
&= p_1 h(f_1) + p_2 h(f_2) && (2.15)
\end{aligned}$$

This shows that the Shannon entropy of the function f is lower-bounded by a convex combination of the entropies of f_1 and f_2 as $h(f) \geq p_1 h(f_1) + p_2 h(f_2)$. Note that this property is shared by any functional of the form $\Phi(f) = \int \varphi(f(\mathbf{r})))\mathrm{d}\mathbf{r}$ where φ is a concave function. The property of concavity can be extended to mixtures between probability distributions $\{f_i\}$ according to a probability vector $\mathbf{p}$ as:

$$h\left(\sum_i p_i f_i\right) \geq \sum_i p_i h(f_i). \tag{2.16}$$

2.2.2 Entropy of a convolution

The convolution of probability distributions is an operation that arises naturally as the sum of random independent variables leads to a convolution of their probability density functions. As an example, if

we define the continuous random variable $\mathbf{Z} = \mathbf{X}+\mathbf{Y}$, the probability density function of $\mathbf{Z}$ is obtained as $f_{\mathbf{Z}} = f_{\mathbf{X}} * f_{\mathbf{Y}}$ (provided that $\mathbf{X}$ and $\mathbf{Y}$ are independent).

Let the probability distribution f be defined as the convolution of f_1 and f_2, so that we have:

$$f(\mathbf{r}) = (f_1 * f_2)(\mathbf{r}) \tag{2.17}$$
$$= \int f_1(\mathbf{r}')f_2(\mathbf{r}-\mathbf{r}')\mathrm{d}\mathbf{r}' \tag{2.18}$$
$$= \int f_1(\mathbf{r}-\mathbf{r}')f_2(\mathbf{r}')\mathrm{d}\mathbf{r}' \tag{2.19}$$

Using the notation $\varphi(x) = -x\ln x$ and remembering that it is concave over $\mathbb{R}_+$, we can apply Jensen's inequality to lower-bound the entropy of f:

$$h(f) = \int \varphi\left(f(\mathbf{r})\right)\mathrm{d}\mathbf{r} \tag{2.20}$$
$$= \int \varphi\left(\int f_1(\mathbf{r}')f_2(\mathbf{r}-\mathbf{r}')\mathrm{d}\mathbf{r}'\right)\mathrm{d}\mathbf{r} \tag{2.21}$$
$$\geq \iint \varphi\left(f_1(\mathbf{r}')\right)f_2(\mathbf{r}-\mathbf{r}')\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r} \tag{2.22}$$
$$= \int \varphi(f_1(\mathbf{s}))\mathrm{d}\mathbf{s} \tag{2.23}$$
$$= h(f_1) \tag{2.24}$$

We have used the fact that $\tilde{f}_2(\mathbf{r}') = f_2(\mathbf{r}-\mathbf{r}')$ is a probability density distribution. A similar development exchanging the place of f_1 and f_2 gives the inequality $h(f) \geq h(f_2)$. We can thus write the inequality $h(f) \geq \max\left(h(f_1), h(f_2)\right)$. In a random variables writing we can write $h(\mathbf{X}+\mathbf{Y}) \geq \max\left(h(\mathbf{X}), h(\mathbf{Y})\right)$.

2.2.3 Entropy of a rescaling

A rescaling transformation is a common operation for continous variables. It corresponds to changing the precision with which we measure the outcome of the variable. We understand naturally that increasing the precision of the measure will increase the entropy associated to the variable. In a random variable notation, a rescaling transformation corresponds simply to multiplying the random variable with a constant. For example, we can define the continuous random variable $\mathbf{Z}$ from the continuous random variable $\mathbf{X}$ as $\mathbf{Z} = s\mathbf{X}$. Their probability density functions will then be related as $f_{\mathbf{Z}} = \mathcal{L}_s\left[f_{\mathbf{X}}\right]$, where $\mathcal{L}_s$ is the rescaling operator defined as:

$$\mathcal{L}_s\left[f\right](\mathbf{r}) = \frac{1}{|s|^n} f\left(\frac{\mathbf{r}}{s}\right) \tag{2.25}$$

where the notation $\mathbf{r}/s$ means that each component r_i is divided by s. Let us now interest ourselves to the entropy of a rescaled probability distribution:

$$h\left(\mathcal{L}_s\left[f\right]\right) = -\int \frac{1}{|s|^n} f\left(\frac{\mathbf{r}}{s}\right)\ln\left(\frac{1}{|s|^n} f\left(\frac{\mathbf{r}}{s}\right)\right)\mathrm{d}\mathbf{r} \tag{2.26}$$
$$= -\int f\left(\mathbf{r}'\right)\ln\left(\frac{1}{|s|^n} f\left(\mathbf{r}'\right)\right)\mathrm{d}\mathbf{r}' \tag{2.27}$$
$$= \int f\left(\mathbf{r}'\right)\ln\left(|s|^n\right)\mathrm{d}\mathbf{r}' - \int f\left(\mathbf{r}'\right)\ln\left(f(\mathbf{r}')\right)\mathrm{d}\mathbf{r}' \tag{2.28}$$
$$= h(f) + n\ln|s| \tag{2.29}$$

where we have used $\mathbf{r}' = \mathbf{r}/|s|$, so that $\mathrm{d}\mathbf{r}' = \mathrm{d}\mathbf{r}/|s|^n$. We have thus the equality $h\left(\mathcal{L}_s\left[f\right]\right) = h\left(f\right) + n\ln|s|$. In a random variables writing we can write $h(s\mathbf{X}) = h(\mathbf{X}) + n\ln|s|$. A major consequence of that rescaling property is that Shannon differential entropy can become negative.

Note that here we have considered a uniform rescaling of all the components of $\mathbf{X}$. In general, when the random variable $\mathbf{X}$ obeys the linear transformation $\mathbf{X} \mapsto \mathbf{A}\mathbf{X}$, its entropy becomes $h(\mathbf{X}) \mapsto h(\mathbf{X}) + \ln|\det\mathbf{A}|$. Synthetically we can write $h(\mathbf{A}\mathbf{X}) = h(\mathbf{X}) + \ln|\det\mathbf{A}|$.

2.2.4 Additivity of entropy

Entropy is said to be an additive quantity because when two isolated system are brought together, their total entropy is the sum of the entropy of both subsystems. When it comes to probability distributions, the action of bringing two independent system together comes down to a tensoring of their probability distributions. Let f_1 and f_2 describe two separate systems, then the global system is described by f which is defined as follows:

$$f(\mathbf{r}_1, \mathbf{r}_2) = f_1(\mathbf{r}_1) f_2(\mathbf{r}_2) \tag{2.30}$$

The entropy of the probability distribution f is then computed as:

$$\begin{aligned}
h(f) &= -\iint f_1(\mathbf{r}_1) f_2(\mathbf{r}_2) \ln\left(f_1(\mathbf{r}_1) f_2(\mathbf{r}_2)\right) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 && (2.31)\\
&= -\iint f_1(\mathbf{r}_1) f_2(\mathbf{r}_2) \left(\ln f_1(\mathbf{r}_1) + \ln f_2(\mathbf{r}_2)\right) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 && (2.32)\\
&= -\iint f_1(\mathbf{r}_1) f_2(\mathbf{r}_2) \ln f(\mathbf{r}_1) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 - \iint f_1(\mathbf{r}_1) f_2(\mathbf{r}_2) \ln f_2(\mathbf{r}_2) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 && (2.33)\\
&= -\int f_1(\mathbf{r}_1) \ln f_1(\mathbf{r}_1) \mathrm{d}\mathbf{r}_1 - \int f_2(\mathbf{r}_2) \ln f_2(\mathbf{r}_2) \mathrm{d}\mathbf{r}_2 && (2.34)\\
&= h(f_1) + h(f_2) && (2.35)
\end{aligned}$$

We have the equality $h(f) = h(f_1) + h(f_2)$. In a random variables writing we write $h(\mathbf{X} \otimes \mathbf{Y}) = h(\mathbf{X}) + h(\mathbf{Y})$.

2.2.5 Subadditivity of entropy

We have seen that when bringing two system together, their entropies sum up. Now, we are going to look at the case of a system which is split into two subsystem. Let $f(\mathbf{r})$ be a probability density function. The parameter $\mathbf{r}$ is the concatenation of two vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, so that $\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2)$. We define $f_1(\mathbf{r}_1)$ as the marginal probability density function of $f(\mathbf{r}_1, \mathbf{r}_2)$ along $\mathbf{r}_1$, and similarly $f_2(\mathbf{r}_2)$ as the marginal probability density function of $f(\mathbf{r}_1, \mathbf{r}_2)$ along $\mathbf{r}_2$.

$$f_1(\mathbf{r}_1) = \int f(\mathbf{r}_1, \mathbf{r}_2) \mathrm{d}\mathbf{r}_2, \qquad f_2(\mathbf{r}_2) = \int f(\mathbf{r}_1, \mathbf{r}_2) \mathrm{d}\mathbf{r}_1. \tag{2.36}$$

We are going to compare the entropy of f with the sum of the entropies of f_1 and f_2. We are going to show that $h(f) \leq h(f_1) + h(f_2)$, or equivalently that $h(f_1) + h(f_2) - h(f) \geq 0$ [16]. Let us first develop $h(f_1) + h(f_2)$ as follows:

$$\begin{aligned}
h(f_1) + h(f_2) &= -\int f_1(\mathbf{r}_1) \ln f_1(\mathbf{r}_1) \mathrm{d}\mathbf{r}_1 - \int f_2(\mathbf{r}_2) \ln f_2(\mathbf{r}_2) \mathrm{d}\mathbf{r}_2 && (2.37)\\
&= -\iint f(\mathbf{r_1}, \mathbf{r}_2) \ln f_1(\mathbf{r}_1) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 - \iint f(\mathbf{r_1}, \mathbf{r}_2) \ln f_2(\mathbf{r}_2) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 && (2.38)\\
&= -\iint f(\mathbf{r}_1, \mathbf{r}_2) \ln\left(f_1(\mathbf{r}_1) f_2(\mathbf{r}_2)\right) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 && (2.39)
\end{aligned}$$

Using this, we can now write the difference $h(f_1) + h(f_2) - h(f)$ as:

$$\begin{aligned}
h(f_1) + h(f_2) - h(f) &= -\iint f(\mathbf{r}_1, \mathbf{r}_2) \ln\left(f_1(\mathbf{r}_1) f_2(\mathbf{r}_2)\right) \mathrm{d}\mathbf{r} + \iint f(\mathbf{r}_1, \mathbf{r}_2) \ln\left(f(\mathbf{r}_1, \mathbf{r}_2)\right) \mathrm{d}\mathbf{r} && (2.40)\\
&= -\iint f(\mathbf{r}_1, \mathbf{r}_2) \ln\left(\frac{f_1(\mathbf{r}_1) f_2(\mathbf{r}_2)}{f(\mathbf{r}_1, \mathbf{r}_2)}\right) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 && (2.41)\\
&\geq -\ln\left(\iint f(\mathbf{r}_1, \mathbf{r}_2) \frac{f_1(\mathbf{r}_1) f_2(\mathbf{r}_2)}{f(\mathbf{r}_1, \mathbf{r}_2)} \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2\right) && (2.42)\\
&= -\ln\left(\int f_1(\mathbf{r}_1) f_2(\mathbf{r}_2) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2\right) = 0 && (2.43)
\end{aligned}$$

where the inequality comes from Jensen's inequality on the convex function $\varphi(x) = -\ln x$. We have thus the inequality $h(f) \leq h(f_1) + h(f_2)$. The non-negative quantity $f(f_1) + h(f_2) - h(f)$ is usually called the mutual information between f_1 and f_2. It quantifies the shared information between the two distributions. Note that when the distribution f is a tensor product of f_1 and f_2, the mutual information is zero.

In a random variables writing, we usually write $h(\mathbf{X}, \mathbf{Y}) \leq h(\mathbf{X}) + h(\mathbf{Y})$, and $h(\mathbf{X}, \mathbf{Y})$ denotes the entropy of the joint distribution between $\mathbf{X}$ and $\mathbf{Y}$. Note also that the property of subadditivity can be used in chain, so that we find the inequality:

$$h(X_1, X_2, ..., X_k) \leq \sum_{i=1}^{k} h(X_i) \tag{2.44}$$

2.3 From Gaussian distributions to the entropy power inequality

Gaussian distributions are ubiquitous in physics. They can serve as a first approximation of any distribution. Also, they are very robust as many transformations preserve the Gaussian character of a distribution. We devote a section of this chapter to Gaussian distributions because of their special relation with entropy. As we are going to see, it is possible to exploit some of their properties to derive stronger entropic inequalities.

2.3.1 First statistical moments

Before presenting Gaussian distribution, let us introduce first the tools required to characterize them. The general shape of a distribution can be described in a first approximation by its statistical moments. The first statistical moment corresponds to the mean or expectation value of the distribution. For multi-dimensional random variables, it is a vector $\bar{\mathbf{r}}$ of components $\bar{r}_i$ such that $\bar{\mathbf{r}} = (\bar{r}_1, \bar{r}_2, ..., \bar{r}_n)$. It is defined as:

Definition 2.9 (Mean value)**.** *The mean value $\bar{\mathbf{r}} \in \mathbb{R}^n$ of a probability distribution $f : \mathbb{R}^n \mapsto \mathbb{R}$ is its first statistical moment and is computed as:*

$$\bar{\mathbf{r}} = \int \mathbf{r}\ f(\mathbf{r})\ \mathrm{d}\mathbf{r}. \tag{2.45}$$

The second statistical moment of a probability distribution is a matrix $\mathbf{V} \in \mathbb{R}^{n \times n}$ The covariance matrix $\mathbf{V}$ is a $n \times n$ matrix.

Definition 2.10 (Covariance matrix)**.** *The covariance matrix $\mathbf{V} \in \mathbb{R}^{n \times n}$ of a probability distribution $f : \mathbb{R}^n \mapsto \mathbb{R}$ is its second statistical moment and is computed as:*

$$(\mathbf{V})_{ij} = \int (r_i - \bar{r}_i)\ (r_j - \bar{r}_j)\ f(\mathbf{r})\ \mathrm{d}\mathbf{r}. \tag{2.46}$$

$$= \int r_i\ r_j\ f(\mathbf{r})\ \mathrm{d}\mathbf{r} - \ \bar{r}_i\ \bar{r}_j. \tag{2.47}$$

In general, any real symmetric positive-definite matrix is an acceptable covariance matrix: $\mathbf{V} \in \mathbb{R}^{n \times n}$, $\mathbf{V}^\intercal = \mathbf{V}$ and $\mathbf{V} > 0$.

The mean $\bar{\mathbf{r}}$ and covariance matrix $\mathbf{V}$ are the two first statistical moment of a continuous distribution. In practice, they described roughly the shape of the distribution. Any probability distribution can be associated to a mean value and a covariance matrix. Gaussian distribution are remarkable in the sense that they are unequivocally defined only by these two first moments.

2.3.2 Gaussian distributions and their properties

Definition 2.11 (Gaussian distribution)**.** *The n-dimensional Gaussian distribution of mean value $\bar{\mathbf{r}}$ and covariance matrix $\mathbf{V}$ is the probability density distribution $G_{\bar{\mathbf{r}},\mathbf{V}} : \mathbb{R}^n \mapsto \mathbb{R}$ defined as follows:*

$$G_{\bar{\mathbf{r}},\mathbf{V}}(\mathbf{r}) = \frac{1}{\sqrt{(2\pi)^n \det \mathbf{V}}} \exp\left(-\frac{1}{2}(\mathbf{r} - \bar{\mathbf{r}})^\intercal \mathbf{V}^{-1} (\mathbf{r} - \bar{\mathbf{r}})\right) \tag{2.48}$$

In practice, we will sometimes simply refer to a Gaussian distribution as a Gaussian.

- Partial tracing a Gaussian gives a Gaussian

  First of all, let us mention that partial tracing a Gaussian by integrating it over one of its variable yields another Gaussian. Indeed, let us say we have a Gaussian defined by its first two statistical moments $\bar{\mathbf{r}}, \mathbf{V}$. If we partial trace that Gaussian over its i^{th} variable, we will obtain another Gaussian with first statistical moments $\bar{\mathbf{r}}', \mathbf{V}'$, so that $\bar{\mathbf{r}}'$ is $\bar{\mathbf{r}}$ with its i^{th} component removed, and $\mathbf{V}'$ is $\mathbf{V}$ with its i^{th} row and i^{th} column removed.

$$\int G_{\bar{\mathbf{r}},\mathbf{V}}\left(r_1, ..., r_i, ..., r_n\right) \mathrm{d}r_i = G_{\bar{\mathbf{r}}',\mathbf{V}'}\left(r_1, ..., r_{i-1}, r_{i+1}, ..., r_n\right) \tag{2.49}$$

- Tensoring two Gaussians gives a Gaussian

  Similarly, if we tensor two Gaussian distribution of respective first statistical moments $\bar{\mathbf{r}}_1, \mathbf{V}_1$ and $\bar{\mathbf{r}}_2, \mathbf{V}_2$, we will obtain a new Gaussian with first statistical moments $\bar{\mathbf{r}}, \mathbf{V}$ defined as $\bar{\mathbf{r}} = (\bar{\mathbf{r}}_1, \bar{\mathbf{r}}_2)$ and $\mathbf{V} = \mathbf{V}_1 \oplus \mathbf{V}_2$.

$$G_{\bar{\mathbf{r}}_1,\mathbf{V}_1}\left(r_1, ..., r_{n_1}\right) \cdot G_{\bar{\mathbf{r}}_2,\mathbf{V}_2}\left(r_{n_1+1}, ..., r_{n_1+n_2}\right) = G_{\bar{\mathbf{r}},\mathbf{V}}\left(r_1, ..., r_{n_1+n_2}\right) \tag{2.50}$$

- Convolving two Gaussians gives a Gaussian

  The Gaussian distribution has the particularity that a convolution of Gaussian distributions is also a Gaussian distribution. In general, convolving two probability distributions of respective first statistical moments $\bar{\mathbf{r}}_1, \mathbf{V}_1$ and $\bar{\mathbf{r}}_2, \mathbf{V}_2$ results in a new probability distribution of first statistical moments $\bar{\mathbf{r}}, \mathbf{V}$ such that $\bar{\mathbf{r}} = \bar{\mathbf{r}}_1 + \bar{\mathbf{r}}_2$ and $\mathbf{V} = \mathbf{V}_1 + \mathbf{V}_2$. In the case of Gaussian distributions we have thus the relation:

$$G_{\bar{\mathbf{r}}_1,\mathbf{V}_1} * G_{\bar{\mathbf{r}}_2,\mathbf{V}_2} = G_{\bar{\mathbf{r}}_1+\bar{\mathbf{r}}_2,\mathbf{V}_1+\mathbf{V}_2} \tag{2.51}$$

- Gaussian maximize entropy for a given variance

  The Gaussian distribution is the distribution that maximizes the entropy under a constrained mean and constrained covariance matrix. Among the set of distributions with a given variance, the Gaussian distribution is the distribution with the highest entropy. Equivalent saying : among the set of distribution with a given entropy, the Gaussian is the distribution with the lowest variance. Let $f_{\bar{\mathbf{r}},\mathbf{V}}$ be a distribution with mean value $\bar{\mathbf{r}}$ and covariance matrix $\mathbf{V}$, and let $G_{\bar{\mathbf{r}},\mathbf{V}}$ be a Gaussian with same mean value and covariance matrix. Then the following holds:

$$h\left(f_{\bar{\mathbf{r}},\mathbf{V}}\right) \leq h\left(G_{\bar{\mathbf{r}},\mathbf{V}}\right) = \frac{1}{2} \ln\left((2\pi e)^n \det \mathbf{V}\right) \tag{2.52}$$

- Gaussians minimize entropy production

  For fixed entropy, Gaussian distributions minimize the entropy of the convolution [55]. Let the distribution f be the result of a convolution of two distributions, such that $f = f_1 * f_2$. Let G_1 and G_2 be two Gaussian distributions with diagonal covariance matrices such that $h(G_1) = h(f_1)$ and $h(G_2) = h(f_2)$. Let G be the Gaussian distribution resulting of the convolution of G_1 and G_2 such that $G = G_1 * G_2$. Then, we have the inequality $h(G) \leq h(f)$. We can also write that as $h(f_1 * f_2) \geq h(G_1 * G_2)$. This can also be formulated in a random variable writing. Let $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$ and $\mathbf{Z}_G = \mathbf{X}_G + \mathbf{Y}_G$, such that $h(X) = h(X_G)$ and $h(Y) = h(Y_G)$. Then, we have the inequality $h(\mathbf{Z}) \geq h(\mathbf{Z}_G)$.

2.3.3 Entropy power inequality

In the previous section, we have shown that the entropy of a convolution of probability distributions can be lower-bounded, see development (2.24). We are going to show that it is possible to take advantage of the properties of Gaussian distribution to derive a stronger inequality. As an illustration, let us consider a Gaussian distribution centered around the origin, such that $\bar{\mathbf{r}} = \mathbf{0}$ and $\mathbf{V} = V\mathbf{I}$. This

implies that $\mathbf{V}^{-1} = \mathbf{I}/V$ and that $\det \mathbf{V} = V^n$. We can compute the expression of G_V as well as its entropy:

$$G_V(\mathbf{r}) = \frac{1}{(\sqrt{2\pi V})^n} \exp\left(-\frac{1}{2}\frac{\mathbf{r}^2}{V}\right), \qquad h(G_V) = \frac{n}{2}\ln(2\pi e V). \tag{2.53}$$

The entropy power $N(f)$ of a distribution is the variance σ^2 of the n-dimensional Gaussian with the same entropy. It is a lowerbound on the variance of the distribution, since the Gaussian minimizes the variance for a given entropy.

Definition 2.12 (Entropy power)**.** *The entropy power of a probability distribution $f : \mathbb{R}^n \mapsto \mathbb{R}$ is the variance of a Gaussian with same entropy. It is equal to the quantity $N(f)$:*

$$N(f) = \frac{1}{2\pi e}\exp\left(\frac{2}{n}h(f)\right) \tag{2.54}$$

The entropy power of a probability distribution can be used to formulate a lower-bound on the determinant of the covariance matrix $\mathbf{V}$ of f as follows:

$$\det \mathbf{V}(f) \geq (N(f))^n. \tag{2.55}$$

Let us consider two distributions f, g. To these two distributions we associate two Gaussian distributions f_G, g_G with same respective entropy. We have $h(f) = h(f_G)$ and $h(g) = h(g_G)$. From these relations it directly follows that $N(f) = N(f_G)$ and $N(g) = N(g_G)$.

$$\begin{aligned} N(f) + N(g) &= N(f_G) + N(g_G) && (2.56)\\ &= V(f_G) + V(g_G) && (2.57)\\ &= V(f_G * g_G) && (2.58)\\ &= N(f_G * g_G) && (2.59)\\ &\leq N(f * g) && (2.60) \end{aligned}$$

The inequality directly follows from the property that Gaussian distributions minimize entropy production under convolution. The entropy power inequality is in fact nothing but a restatement of that property. The entropy power inequality is a non-trivial lower-bound on the entropy of a convolution. Let us now rework that inequation. We propose the following development:

$$\begin{aligned} & N(f*g) \geq N(f) + N(g) && (2.61)\\ \Leftrightarrow \quad & \exp\left(\frac{2}{n}h(f*g)\right) \geq \exp\left(\frac{2}{n}h(f)\right) + \exp\left(\frac{2}{n}h(g)\right) && (2.62)\\ \Leftrightarrow \quad & \frac{2}{n}h(f*g) \geq \ln\left(\exp\left(\frac{2}{n}h(f)\right) + \exp\left(\frac{2}{n}h(g)\right)\right) && (2.63)\\ \Leftrightarrow \quad & \frac{2}{n}h(f*g) \geq \ln\left(p_1\left(\frac{1}{p_1}\exp\left(\frac{2}{n}h(f)\right)\right) + p_2\left(\frac{1}{p_2}\exp\left(\frac{2}{n}h(g)\right)\right)\right) && (2.64)\\ \Rightarrow \quad & \frac{2}{n}h(f*g) \geq p_1\ln\left(\frac{1}{p_1}\exp\left(\frac{2}{n}h(f)\right)\right) + p_2\ln\left(\frac{1}{p_2}\exp\left(\frac{2}{n}h(g)\right)\right) && (2.65)\\ \Leftrightarrow \quad & \frac{2}{n}h(f*g) \geq p_1\frac{2}{n}h(f) - p_1\ln p_1 + p_2\frac{2}{n}h(g) - p_2\ln p_2 && (2.66)\\ \Leftrightarrow \quad & h(f*g) \geq p_1 h(f) + p_2 h(g) + \frac{n}{2}(-p_1\ln p_1 - p_2\ln p_2) && (2.67)\\ \Leftrightarrow \quad & h(f*g) \geq p_1 h(f) + p_2 h(g) + \frac{n}{2}H(\mathbf{p}) && (2.68) \end{aligned}$$

where the parameters p_1, p_2 are such that (p_1, p_2) is a probability vector. It is interesting to see that the entropy power inequality can be restated directly as a lower-bound on $h(f * g)$ which depends on

$h(f), h(g)$ and an arbitrarily chosen probability vector. From that observation, we see that the entropy power inequality is a refinement of the property that entropy increases under convolution. Indeed, if we choose (p_1, p_2) to be $(1, 0)$ or $(0, 1)$, we find the inequality $h(f * g) \geq h(f)$ or $h(f * g) \geq h(g)$, respectively.

To conclude this section, let us present the straightforward extension of these relations to several convolutions:

$$N\left(\overset{k}{\underset{i=1}{\ast}} f_i\right) \geq \sum_{i=1}^{k} N(f_i) \tag{2.69}$$

$$h\left(\overset{k}{\underset{i=1}{\ast}} f_i\right) \geq \sum_{i=1}^{k} p_i h(f_i) + \frac{n}{2} H(\mathbf{p}) \tag{2.70}$$

Finally, let us note how the entropy power of a distribution evolves under a rescaling transformation:

$$N\left(\mathcal{L}_s\left[f\right]\right) = \frac{1}{2\pi e} \exp\left(\frac{2}{n} h\left(\mathcal{L}_s\left[f\right]\right)\right) \tag{2.71}$$

$$= \frac{1}{2\pi e} \exp\left(\frac{2}{n}\left(h(f) + n \ln|s|\right)\right) \tag{2.72}$$

$$= \frac{1}{2\pi e} \exp\left(\frac{2}{n} h(f)\right) \exp\left(2 \ln|s|\right) \tag{2.73}$$

$$= s^2 N(f) \tag{2.74}$$

2.4 Other information-theoretical measures

In the previous sections, we have focused our interest on the central measure of information that is Shannon entropy. Let us now presents some other measures.

2.4.1 p-norms

Let us introduce the p-norms.

Definition 2.13 (p-norm)**.** *The p-norm of a distribution $f : \mathbb{R}^n \mapsto \mathbb{R}$ is defined as:*

$$\|f\|_p = \left(\int |f(\mathbf{r})|^p \mathrm{d}\mathbf{r}\right)^{\frac{1}{p}} \tag{2.75}$$

Similarly, the p-norm of a vector $\mathbf{x} \in \mathbb{R}^N$ is defined as:

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{N} |x_i|^p\right)^{\frac{1}{p}} \tag{2.76}$$

The summation is carried until infinity if $\mathbf{x} \in \mathbb{R}^{\mathbb{N}}$.

The case $p = 0$ and $p = \infty$ require some clarification. In the limit $p \to 0$, the p-norm $\|f\|_p$ tends towards the size of the domain which has a non-zero value, so that we identify $\|f\|_0 = \nu\left(\{\mathbf{r} : f(\mathbf{r}) \neq 0\}\right)$, where ν is the Lebesgue measure. In the limit $p \to \infty$, the p-norm $\|f\|_p$ tends towards the maximum absolute value of f, so that we identify $\|f\|_\infty = \max(|f(\mathbf{r})|)$.

One of the main interest of p-norms is that they enable to formulate many general inequalities. Let us present some of them.

Hölder's inequality

Hölder's inequality establish a relation between the p-norms of a product of function and the p-norms of the respective functions [31]. It reads as follows:

$$\|fg\|_1 \leq \|f\|_p \|g\|_q \tag{2.77}$$

where $p, q \geq 1$ and $1/p + 1/q = 1$. The inequality becomes an equality if and only if $|f|^p$ and $|g|^q$ are linearly dependent. Note that in the case $p = q = 2$, Hölder's inequality implies the famous Cauchy-Schwarz inequality. Indeed, in that case we obtain $\|fg\|_1 \leq \|f\|_2 \|g\|_2$, which is a stronger version of Cauchy-Schwarz inequality because $\int |f(\mathbf{r})||g(\mathbf{r})| \mathrm{d}\mathbf{r} \geq \int f(\mathbf{r}) g(\mathbf{r}) \mathrm{d}\mathbf{r}$.

Young's convolution inequality

Hölder's inequality can be used to derive an ineqality for the convolution of two functions. This leads to Young's convolution inequality [68], which reads as follows:

$$\|f * g\|_r \leq \|f\|_p \|g\|_q \tag{2.78}$$

where $p, q, r \geq 1$ and $1/p + 1/q = 1/r + 1$.

Babenko-Beckner inequality

Babenko-Beckner inequality relates the p-norm of a function to the p-norm its Fourier transform [3, 5]. Let $\psi : \mathbb{R}^n \mapsto \mathbb{C}$ and $\phi : \mathbb{R}^n \mapsto \mathbb{C}$ be related by a Fourier transformation, so that $\phi = \mathcal{F}[\psi]$.

$$\|\phi\|_p \leq \left(\frac{2\pi}{q}\right)^{\frac{n}{2q}} \left(\frac{2\pi}{p}\right)^{\frac{n}{2q}} \|\psi\|_q \tag{2.79}$$

where $1 < p \leq 2$, $q \geq 2$ and $1/p + 1/q = 1$. Interestingly, in the case $p = q = 2$ the inequality becomes an equality, which is Parseval's identity $\|\psi\|_2 = \|\phi\|_2$.

2.4.2 Rényi entropy

Let us now introduce a family of entropies that is built upon p-norms.

Definition 2.14 (Rényi entropy)**.** *The Rényi entropy of parameter α of a distribution $f : \mathbb{R}^n \mapsto \mathbb{R}$ is defined as follows:*

$$h_\alpha(f) = \frac{\alpha}{1-\alpha} \ln\left(\|f\|_\alpha\right) \tag{2.80}$$

$$= \frac{1}{1-\alpha} \ln\left(\int |f(\mathbf{r})|^\alpha \mathrm{d}\mathbf{r}\right) \tag{2.81}$$

Similarly, the Rényi entropy of parameter α of a vector $\mathbf{x} \in \mathbb{R}^N$ is defined as follows:

$$H_\alpha(\mathbf{x}) = \frac{\alpha}{1-\alpha} \ln\left(\|\mathbf{x}\|_\alpha\right) \tag{2.82}$$

$$= \frac{1}{1-\alpha} \ln\left(\sum_{i=1}^{N} |x_i|^p\right) \tag{2.83}$$

The summation is carried until infinity if $\mathbf{x} \in \mathbb{R}^{\mathbb{N}}$.

First, let us notice that Shannon entropy can be obtain as a limit case of Rényi entropy. In the limit $\alpha \to 1$, the definition of the Rényi entropy tends towards an indetermination. However, a simple

development shows that it tends towards the Shannon entropy:

$$\lim_{\alpha\to 1} \quad h_\alpha(f) = \lim_{\alpha\to 1} \quad \frac{1}{1-\alpha}\ln\left(\int (f(\mathbf{r}))^\alpha \,\mathrm{d}\mathbf{r}\right) \tag{2.84}$$

$$= \lim_{\alpha\to 1} \quad \frac{1}{\frac{\mathrm{d}}{\mathrm{d}\alpha}(1-\alpha)} \quad \frac{\mathrm{d}}{\mathrm{d}\alpha}\ln\left(\int (f(\mathbf{r}))^\alpha \,\mathrm{d}\mathbf{r}\right) \tag{2.85}$$

$$= \lim_{\alpha\to 1} \quad -1 \quad \frac{1}{\int (f(\mathbf{r}))^\alpha \,\mathrm{d}\mathbf{r}} \quad \frac{\mathrm{d}}{\mathrm{d}\alpha}\int (f(\mathbf{r}))^\alpha \,\mathrm{d}\mathbf{r} \tag{2.86}$$

$$= \lim_{\alpha\to 1} \quad -\frac{1}{(\|f\|_\alpha)^\alpha} \quad \int \frac{\mathrm{d}}{\mathrm{d}\alpha}(f(\mathbf{r}))^\alpha \,\mathrm{d}\mathbf{r} \tag{2.87}$$

$$= -\frac{1}{\|f\|_1} \quad \lim_{\alpha\to 1} \quad \int (f(\mathbf{r}))^\alpha \ln(f(\mathbf{r}))\,\mathrm{d}\mathbf{r} \tag{2.88}$$

$$= -\frac{1}{\|f\|_1} \quad \int f(\mathbf{r})\ln(f(\mathbf{r}))\,\mathrm{d}\mathbf{r} \tag{2.89}$$

$$= h(f) \tag{2.90}$$

Let us now observe how the rescaling operator $\mathcal{L}_s$ affects Rényi entropies. First, let us compute its effect on p-norms.

$$\|\mathcal{L}_s[f]\|_p = \left(\int \left|\frac{1}{|s|^n} f\left(\frac{\mathbf{r}}{s}\right)\right|^p \mathrm{d}\mathbf{r}\right)^{\frac{1}{p}} \tag{2.91}$$

$$= \left(\int |s|^{-np}\left|f\left(\frac{\mathbf{r}}{s}\right)\right|^p \mathrm{d}\mathbf{r}\right)^{\frac{1}{p}} \tag{2.92}$$

$$= \left(\int |s|^{n(1-p)}\left|f\left(\mathbf{r}'\right)\right|^p \mathrm{d}\mathbf{r}'\right)^{\frac{1}{p}} \tag{2.93}$$

$$= |s|^{\frac{n(1-p)}{p}}\left(\int \left|f\left(\mathbf{r}'\right)\right|^p \mathrm{d}\mathbf{r}'\right)^{\frac{1}{p}} \tag{2.94}$$

$$= |s|^{\frac{n(1-p)}{p}}\|f\|_p \tag{2.95}$$

From the above expression, we can readily find the effect of $\mathcal{L}_s$ on Rényi entropies as:

$$h_\alpha(\mathcal{L}_s[f]) = h_\alpha(f) + n\ln|s| \tag{2.96}$$

Interestingly, the rescaling operator has the same effect regardless of the parameter α. It is also similar to the scaling on Shannon entropy, see (2.29).

2.5 Quantum applications

The different tools we have introduced in this chapter find application in the field of quantum information. Indeed, the physicality conditions that apply on a density operator $\hat{\rho}$ correspond to the fact that its eigenvalues should correspond to a probability distribution. In the following, let us write as $\boldsymbol{\lambda}$ the vector of eigenvalues of $\hat{\rho}$, so that $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ...)$. As we said, $\boldsymbol{\lambda}$ is a probability vector.

Definition 2.15 (Purity)**.** *The purity of a quantum state is a measure of its statistical uncertainty.*

$$\mu(\hat{\rho}) = \sum_{i=1}^{N} \lambda_i^2 \tag{2.97}$$

We can also write $\mu(\hat{\rho}) = \mathrm{Tr}\left[\hat{\rho}^2\right]$. In general, the purity of a state is below 1, which corresponds to a pure state, as $0 < \mu \leq 1$. We have $\mu = \|\boldsymbol{\lambda}\|_2^2$.

Definition 2.16 (Von Neumann entropy)**.** *The von Neumann entropy of a quantum state described by a density operator $\hat{\rho}$ is the Shannon discrete entropy of the probability vector $\boldsymbol{\lambda}$ made from the eigenvalues of $\hat{\rho}$.*

$$S(\hat{\rho}) = H(\boldsymbol{\lambda}) \tag{2.98}$$

We can also write $S(\hat{\rho}) = -\mathrm{Tr}\left[\hat{\rho}\ln\hat{\rho}\right]$. The von Neumann entropy of a quantum state is always non-negative, and reaches zero for pure states.

Let us now consider two canonically conjugated observable with continuous spectrum, $\hat{x}$ and $\hat{p}$. A quantum state is associated to two probability distributions $\rho_x(x)$ and $\rho_p(p)$. Bialynicki-Birula and Mycielski derived the following equation in [6]. The result is a consequence of Babenko-Beckner inequality.

$$h_\alpha\left(\rho_x\right) + h_\beta\left(\rho_p\right) \geq \ln\pi + \frac{1}{2}\left(\frac{\ln\alpha}{\alpha-1} + \frac{\ln\beta}{\beta-1}\right) \tag{2.99}$$

where α and β are related through $1/\alpha + 1/\beta = 2$. In the special of $\alpha = \beta = 1$, the inequation becomes particularly interesting in the sense that the same functional is evaluated on ρ_x and ρ_p. In that case, we find the entropic uncertainty relation:

$$h(\rho_x) + h(\rho_p) \geq \ln\pi + 1 \tag{2.100}$$

It is instructive to note that this relation directly implies Heisenberg uncertainty relation. Indeed, an equivalent formulation of (2.100) is the relation $N(\rho_p)N(\rho_p) \geq 1/4$ [29]. Then, using the fact that $V(\rho_x) \geq N(\rho_x)$ and $V(\rho_p) \geq V(\rho_p)$, we find the relation $V(\rho_x)V(\rho_p) \geq 1/4$. Defining the standard deviation of ρ_x and ρ_p as respectively σ_x and σ_p, we find:

$$\sigma_x \cdot \sigma_p \;\geq\; \frac{1}{2} \tag{2.101}$$

Chapter 3

Theory of majorization

In this chapter, we set up the tools of the main mathematical theory that we plan to use in order to study the disorder of Wigner functions. The theory of majorization is a very powerful formalism that allows to compare distributions in terms of disorder. This has been studied in detail in [48] which will be our reference throughout this chapter. A main interest of majorization for our purposes is that it extends to distributions that do not necessarily correspond to probability distributions. Indeed, we can use thoery of majorization to compare in terms of disorder distributions taking both positive and negative values, and which are not normalized to the same value. The case of finite-domain distributions, discrete and continuous, taking both positive and negative values has been covered in reference [48], as well as the case of infinite-domain distributions, discrete and continuous, taking only non-negative values. The case of infinite-domain distributions taking both positive and negative values, which is precisely the case which we are interested in, has not been addressed yet.

The objective of this chapter is therefore to propose a formulation of majorization theory that is compatible with Wigner functions, i.e. infinite-domain continuous distributions taking both positive and negative values. To do this, we will proceed in stages. We will start by looking at distributions defined on a finite size domain, first in the discrete case (Section 3.1), then in the continuous case (Section 3.2). Note that throughout this chapter, we will always consider distributions that can take both positive and negative values. As we announced, these two cases have already been covered in [48], but we will approach them in order to extend the formalism to distributions defined over an infinite domain.

Then, in Section 3.3, we will address the case of infinite dimension vectors taking both positive and negative values. This will allow us to pinpoint discontinuities appearing as a result of their infinite size domain. However, by defining the appropriate objects, we will manage to extend the formalism developed in the previous sections to the infinite dimensional case. Finally, in Section 3.3, we will extend our formulation to the case that interests us here, which is the case of infinite-domain continuous distributions.

Note that this chapter dealing with a very mathematical subject is approached with the perspective of a physics thesis. A thorough study of majorization is not our primary motivation in the present work. The transition from the finite-domain case to the infinite-domain case is supported by arguments of continuity and consistency between the two cases. We leave to a future work a fully rigorous mathematical demonstration of the relations we formulate in the infinite-domain case. Let us note that the relations we propose reduce, as it should, to the known case when we consider distributions taking non-negative values.

3.1 Discrete majorization on a finite-size domain

In this first section, we limit our study to the case of finite-dimensional vectors which can have positive and negative components. We don't consider vectors that are normalized, but we consider vectors

such that their absolute norm is finite:

$$\mathbf{x} \in \mathbb{R}^N, \qquad \sum_{i=1}^{N} |x_i| < \infty. \tag{3.1}$$

Note that condition $\sum_{i=1}^{N} |x_i| < \infty$ is obviously implied by condition $\mathbf{x} \in \mathbb{R}^N$. We are going to construct step by step the different tools that we will use in order to set up a discrete majorization relation.

3.1.1 Level-equivalence

A central key in majorization is the notion of level-equivalence. Indeed, from the point of view of majorization, two level-equivalent vectors are completely indistinguishable, as it will appear clearly later. In the present case, it is easy to define what are two level-equivalent finite-dimensional vectors. We say that $\mathbf{x}$ and $\mathbf{y}$ are level-equivalent if they have the same set of components:

$$\mathbf{x} \equiv \mathbf{y} \qquad \Leftrightarrow \qquad \{x_i\} = \{y_i\} \tag{3.2}$$

Two level-equivalent vectors have the same components, but possibly in a different order. For that reason, they are related by some permutation matrix $\mathbf{P}_i$. Remember that a permutation matrix is a matrix made only of 0 and 1, such that there is exactly one 1 by row and by column. The condition $\mathbf{x} \equiv \mathbf{y}$ is equivalent to the statement $\mathbf{x} = \mathbf{P}_i\mathbf{y}$. Applying a permutation matrix on any vector creates another level-equivalent vector.

$$\mathbf{P}_i\mathbf{x} = \mathbf{x}^{(i)} \equiv \mathbf{x} \tag{3.3}$$

We use the notation $\mathbf{x}^{(i)}$ to express that the state is obtained from $\mathbf{x}$ through the permutation matrix $\mathbf{P}_i$. So that in our writing, $\mathbf{x}^{(i)}$ is always level-equivalent to $\mathbf{x}$.

Now that we have defined the notion of level-equivalence for finite-dimensional vectors, let us introduce a tool that is specifically designed to enclose all the information about level-equivalence : the level-function. The level-function of a distribution is a function which characterizes the values taken by the distribution, irrespectively of their order. We define two different versions of the level-function.

Definition 3.1 (Level-functions of a vector)**.** *The upper level-function $m_{\mathbf{x}}^{+}(t)$ of a vector $\mathbf{x}$ associates to any real value t the number of components of $\mathbf{x}$ that are greater or equal to t.*

$$m_{\mathbf{x}}^{+}(t) = \#\left(\{i : x_i \geq t\}\right). \tag{3.4}$$

Conversely, the lower level-function $m_{\mathbf{x}}^{-}(t)$ of a vector $\mathbf{x}$ associates to any real value t the number of components of $\mathbf{x}$ that are lower or equal to t.

$$m_{\mathbf{x}}^{-}(t) = \#\left(\{i : x_i \leq t\}\right). \tag{3.5}$$

From their definitions, we can express the level-functions as:

$$m_{\mathbf{x}}^{+}(t) = \sum_{i=1}^{N} \Theta(x_i - t), \qquad m_{\mathbf{x}}^{-}(t) = \sum_{i=1}^{N} \Theta(t - x_i). \tag{3.6}$$

where Θ is the Heaviside step function, such that $\Theta(z) = 1$ if $z \geq 0$ and is zero otherwise. The level-function is a precious tool as it enclose all the information about the set of components of a vector. It memorizes what are the values taken, but doesn't keep the information about the order of the values.

As both level-functions contain the information about the number of components with value higher or lower than t, we understand that the derivative of the level-functions is related to the number of

components having a value equal to t. We define the upper level-density function $\mu^+_{\mathbf{x}}$ and lower level-density function $\mu^-_{\mathbf{x}}$ of a vector as the respective derivative of the upper and lower level-functions:

$$\mu^+_{\mathbf{x}}(t) = -\frac{\mathrm{d}}{\mathrm{d}t} m^+_{\mathbf{x}}(t), \qquad \mu^-_{\mathbf{x}}(t) = \frac{\mathrm{d}}{\mathrm{d}t} m^-_{\mathbf{x}}(t). \tag{3.7}$$

As it appears, in the case of finite-dimensional vectors that we are considering here, the upper and lower level-density functions are equal, so that we define the level-density function $\mu_{\mathbf{x}}$ as follows:

$$\mu_{\mathbf{x}}(t) = \mu^+_{\mathbf{x}}(t) = \mu^-_{\mathbf{x}}(t) = \sum_{i=1}^{N} \delta(x_i - t). \tag{3.8}$$

From the construction of the level-functions and level-density function, it appears that two vectors are level-equivalent if and only if they have the same level-functions (or level-density functions). This leads us to the following definition of level-equivalence for finite-dimensional vectors.

Definition 3.2 (Level-equivalence for finite-dimensional vectors)**.** *Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$. We say that $\mathbf{x}$ is level-equivalent to $\mathbf{y}$, written $\mathbf{x} \equiv \mathbf{y}$ if and only if their upper level-functions are equal.*

$$\mathbf{x} \equiv \mathbf{y} \qquad \Leftrightarrow \qquad m^+_{\mathbf{x}}(t) = m^+_{\mathbf{y}}(t) \qquad \forall t \tag{3.9}$$

Obviously, two finite-dimensional vectors which have the same lower level-functions are also level-equivalent, since the upper and lower level-functions are related to each other in the finite-dimensional case. Later in this section, we will present other conditions which are equivalent to the condition of level-equivalence.

3.1.2 Symmetric functions

As we may guess, level-equivalent vectors share many similar properties. We are going to introduce this with the use of symmetric functions. Let us define the function $\Phi : \mathbb{R}^N \mapsto \mathbb{R}$ that takes a vector as argument and that is constructed as follows:

$$\Phi(\mathbf{x}) = \sum_{i=1}^{N} \varphi(x_i) \tag{3.10}$$

where $\varphi : \mathbb{R} \mapsto \mathbb{R}$ is some function with no particular constraint. From its construction, it appears that Φ takes the same value for vectors that are related through a permutation, since the summation is performed indistinguishably over all indices. Using the level-density function, we can rewrite Φ as follows:

$$\Phi(\mathbf{x}) = \int \varphi(t)\mu_{\mathbf{x}}(t)\mathrm{d}t \tag{3.11}$$

which appears clearly from Equation (3.8). In the above equation, $\mu_{\mathbf{x}}(t)\mathrm{d}t$ can be understood as the number of components taking the value t, and $\varphi(t)$ is the value taken by φ for these components. We say that Φ is symmetric, because it is invariant under permutation of its argument. A symmetric function takes the same value for all level-equivalent vectors.

Definition 3.3 (Symmetric function)**.** *The function $\Phi : \mathbb{R}^N \mapsto \mathbb{R}$ is said to be symmetric if and only if it takes the same value for level-equivalent vectors.*

$$\mathbf{x} \equiv \mathbf{y} \qquad \Rightarrow \qquad \Phi(\mathbf{x}) = \Phi(\mathbf{y}) \tag{3.12}$$

Let us now go a step further and define two families of functions $\mathbb{R} \mapsto \mathbb{R}$ depending on a real parameter t. We define γ^+_t and γ^-_t as follows:

$$\gamma^+_t(z) = [z - t]^+, \qquad \gamma^-_t(z) = [z - t]^-. \tag{3.13}$$

where we use the notation $[z]^+ = \max(z, 0)$ and $[z]^- = \min(z, 0)$. In that notation, $[z]^+ = z$ when $z \geq 0$ and is zero otherwise, and $[z]^- = z$ when $z \leq 0$ and is zero otherwise. From the two functions γ_t^+ and γ_t^- we construct the two symmetric functions Φ_t^+ and Φ_t^- as follows:

$$\Phi_t^+(\mathbf{x}) = \sum_{i=1}^{N} \gamma_t^+(x_i), \qquad \Phi_t^-(\mathbf{x}) = \sum_{i=1}^{N} \gamma_t^-(x_i). \tag{3.14}$$

These two functions are related to each other, as we can see from the following development:

$$\begin{aligned}
\Phi_t^+(\mathbf{x}) + \Phi_t^-(\mathbf{x}) &= \sum_{i=1}^{N} [x_i - t]^+ + \sum_{i=1}^{N} [x_i - t]^- && (3.15)\\
&= \sum_{x_i \geq t} (x_i - t) + \sum_{x_i \leq t} (x_i - t) && (3.16)\\
&= \sum_{i=1}^{N} (x_i - t) && (3.17)\\
&= \sum_{i=1}^{N} x_i - Nt && (3.18)
\end{aligned}$$

These two set of symmetric functions have a particular interest regarding level-functions. Indeed, a short development shows us that we can relate them to the level-functions:

$$\begin{aligned}
\Phi_t^+(\mathbf{x}) &= \sum_{i=1}^{N} \gamma_t^+(x_i) && (3.19)\\
&= \int \gamma_t^+(u) \mu_{\mathbf{x}}(u) \mathrm{d}u && (3.20)\\
&= \int_t^{\infty} (u - t) \mu_{\mathbf{x}}(u) \mathrm{d}u && (3.21)\\
&= -\Big[(u - t) m_{\mathbf{x}}^+(u)\Big]_t^{\infty} + \int_t^{\infty} m_{\mathbf{x}}^+(u) \mathrm{d}u && (3.22)\\
&= \int_t^{\infty} m_{\mathbf{x}}^+(u) \mathrm{d}u && (3.23)
\end{aligned}$$

where we have use integration by parts and the fact that $m_{\mathbf{x}}^+(t)$ is zero when $t > \max(\mathbf{x})$. A similar development yields:

$$\begin{aligned}
\Phi_t^-(\mathbf{x}) &= \sum_{i=1}^{N} \gamma_t^-(x_i) && (3.24)\\
&= \int \gamma_t^-(u) \mu_{\mathbf{x}}(u) \mathrm{d}u && (3.25)\\
&= \int_{-\infty}^{t} (u - t) \mu_{\mathbf{x}}(u) \mathrm{d}u && (3.26)\\
&= \Big[(u - t) m_{\mathbf{x}}^-(u)\Big]_{-\infty}^{t} - \int_{-\infty}^{t} m_{\mathbf{x}}^-(u) \mathrm{d}u && (3.27)\\
&= -\int_{-\infty}^{t} m_{\mathbf{x}}^-(u) \mathrm{d}u && (3.28)
\end{aligned}$$

As a consequence, this implies that if $\Phi_t^+(\mathbf{x}) = \Phi_t^+(\mathbf{y})$ for all t, then their level-functions are equal for all t. The same observation applies to Φ_t^-. The condition of level-equivalence that we have introduced at the beginning of this subsection can now be reformulated in other equivalent conditions. For any two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$, the following statements are equivalent:

$$\mathbf{x} \equiv \mathbf{y} \quad \Leftrightarrow \quad m_{\mathbf{x}}^+(t) = m_{\mathbf{y}}^+(t) \qquad \forall t \tag{3.29}$$

$$\Leftrightarrow \quad m_{\mathbf{x}}^-(t) = m_{\mathbf{y}}^-(t) \qquad \forall t \tag{3.30}$$

$$\Leftrightarrow \quad \mu_{\mathbf{x}}(t) = \mu_{\mathbf{y}}(t) \qquad \forall t \tag{3.31}$$

$$\Leftrightarrow \quad \sum_{i=1}^{N} \varphi(x_i) = \sum_{i=1}^{N} \varphi(y_i) \qquad \forall \varphi \tag{3.32}$$

$$\Leftrightarrow \quad \Phi(\mathbf{x}) = \Phi(\mathbf{y}) \qquad \forall \Phi \text{ symmetric} \tag{3.33}$$

It is important to note that the implication $\mathbf{x} \equiv \mathbf{y} \Rightarrow \Phi(\mathbf{x}) = \Phi(\mathbf{y})\ \forall \Phi$ symmetric is a direct consequence of the definition of a symmetric function. The converse implication, namely $\Phi(\mathbf{x}) = \Phi(\mathbf{y})\ \forall \Phi$ symmetric $\Rightarrow \mathbf{x} \equiv \mathbf{y}$, follows from the existence of the symmetric functions Φ_t^+ and Φ_t^-, and from their relation to the level-function.

3.1.3 Rearrangements and cumulative sums

Now that we have properly defined what are level-equivalent vectors, let us introduce rearrangements. Rearrangements are very diverse. Essentially, the rearrangement of a vector is a vector with same components, but ordered in a special order. By construction, a rearranged vector and the original vector are level-equivalent, when dealing with finite-dimensional vectors.

In majorization, two rearrangements plays a prominent role : the decreasing rearrangement and increasing rearrangement. As their names indicate, they correspond respectively to a vector whose components are sorted by decreasing order, and increasing order. Let us consider a vector $\mathbf{x} \in \mathbb{R}^N$. Its decreasing rearrangement is written as $\mathbf{x}^\downarrow$ and its increasing rearrangement is written as $\mathbf{x}^\uparrow$. These are defined as follows:

$$\left(\mathbf{x}^\downarrow\right)_i = x_i^\downarrow, \qquad \left(\mathbf{x}^\uparrow\right)_i = x_i^\uparrow \tag{3.34}$$

where $x_i^\downarrow$ (resp. $x_i^\uparrow$) is the i^{th} highest (resp. lowest) component of $\mathbf{x}$. Notice that the i^{th} highest component is also the $(N+1-i)^{\text{th}}$ lowest component, so that we have $x_i^\downarrow = x_{N+1-i}^\uparrow$. By construction, any finite-dimensional vector $\mathbf{x}$ is level-equivalent to its decreasing rearrangement $\mathbf{x}^\downarrow$ and its increasing rearrangement $\mathbf{x}^\uparrow$, so that we have $\mathbf{x}^\downarrow \equiv \mathbf{x}^\uparrow \equiv \mathbf{x}$.

Let us now highlight how the decreasing and increasing rearrangements are related to the level-function. To that purpose, we make the assumption in what follows that $\mathbf{x}$ is a non-degenerate vector, which means that there are no components that take the same value. It is then obvious that the number of components having a value greater or equal to the k^{th} highest component is precisely k. Conversely, the number of components having a value lower or equal to the k^{th} lowest component is k. We can then write:

$$m_{\mathbf{x}}^+(x_k^\downarrow) = \#\left(\{i : x_i \geq x_k^\downarrow\}\right) = k, \qquad m_{\mathbf{x}}^-(x_k^\uparrow) = \#\left(\{i : x_i \leq x_k^\uparrow\}\right) = k. \tag{3.35}$$

Taking advantage of that observation, we introduce the inverse function of $m_{\mathbf{x}}^+$ and $m_{\mathbf{x}}^-$ as respectively $M_{\mathbf{x}}^+$ and $M_{\mathbf{x}}^-$, so that we have the following relations:

$$\left(m_{\mathbf{x}}^+\right)^{-1}(k) = M_{\mathbf{x}}^+(k) = x_k^\downarrow, \qquad \left(m_{\mathbf{x}}^-\right)^{-1}(k) = M_{\mathbf{x}}^-(k) = x_k^\uparrow. \tag{3.36}$$

Note that more generally, the functions $M_{\mathbf{x}}^+$ and $M_{\mathbf{x}}^-$ can be defined as follows:

$$M_{\mathbf{x}}^+(u) = \max\left(\{t : m_{\mathbf{x}}^+(t) \geq u\}\right), \qquad M_{\mathbf{x}}^-(u) = \min\left(\{t : m_{\mathbf{x}}^-(t) \geq u\}\right). \tag{3.37}$$

Let us now introduce the cumulative sums of a vector. As we are going to see, these function play a central role when establishing a majorization relation.

Definition 3.4 (Cumulative sums)**.** *The cumulative sum* $S_k : \mathbb{R}^N \mapsto \mathbb{R}$ *of a vector* $\mathbf{x} \in \mathbb{R}^N$ *is the sum of the first* k *components of* $\mathbf{x}$*:*

$$S_k(\mathbf{x}) = \sum_{i=1}^{k} x_i. \tag{3.38}$$

The decreasing cumulative sum $S_k^{\downarrow} : \mathbb{R}^N \mapsto \mathbb{R}$ *of a vector* $\mathbf{x} \in \mathbb{R}^N$ *is the sum of the* k *highest components of* $\mathbf{x}$*:*

$$S_k^{\downarrow}(\mathbf{x}) = \sum_{i=1}^{k} x_i^{\downarrow}. \tag{3.39}$$

The increasing cumulative sum $S_k^{\uparrow} : \mathbb{R}^N \mapsto \mathbb{R}$ *of a vector* $\mathbf{x} \in \mathbb{R}^N$ *is the sum of the* k *lowest components of* $\mathbf{x}$*:*

$$S_k^{\uparrow}(\mathbf{x}) = \sum_{i=1}^{k} x_i^{\uparrow}. \tag{3.40}$$

As it appears, the cumulative sum S_k is not a symmetric function, since the components that are included vary under a permutation. On the contrary, the decreasing cumulative sum $S_k^{\downarrow}$ and increasing cumulative sum $S_k^{\uparrow}$ are symmetric by construction. As it appears, $S_k^{\downarrow}(\mathbf{x}) = S_k\left(\mathbf{x}^{\downarrow}\right)$ and $S_k^{\uparrow}(\mathbf{x}) = S_k\left(\mathbf{x}^{\uparrow}\right)$. Summing the k highest components of $\mathbf{x}$ with its $N-k$ lowest components comes to summing all the components of $\mathbf{x}$, so that we have the following relation: $S_k^{\downarrow}(\mathbf{x}) + S_{N-k}^{\uparrow}(\mathbf{x}) = S_N(\mathbf{x})$. Since the decreasing cumulative sum $S_k^{\downarrow}$ is summing the k highest components of $\mathbf{x}$ it will always be greater or equal to the sum of any k components of $\mathbf{x}$. Conversely, the increasing cumulative sum $S_k^{\uparrow}$ is always lower or equal to the sum of any k components of $\mathbf{x}$. We have the relation $S_k^{\uparrow}(\mathbf{x}) \leq S_k(\mathbf{x}) \leq S_k^{\downarrow}(\mathbf{x})$. Also, note that $S_N(\mathbf{x}) = S_N^{\downarrow}(\mathbf{x}) = S_N^{\uparrow}(\mathbf{x}) = \sum_{i=1}^{N} x_i$.

3.1.4 Relations of majorization

A relation of majorization between two vectors is a strong statement that one of the two vectors is more disordered than the other according to a large variety of disorder measures. Before introducing the standard relation of majorization, we are going to introduce weak-majorization, which compares vectors of possibly different norms. Then, we will come to the standard relation of majorization. The reader should know that different relations of majorization we present here have been extensively studied in the case of finite-dimensional vectors, possibly with negative values, in [48]. We choose in this subsection to present a brief and straight to the point overview of the different definitions of majorization.

Weak-majorization from below (sub-majorization)

A relation of weak-majorization can be established between two vectors of different norms. The relation "$\mathbf{x}$ sub-majorizes $\mathbf{y}$" can be understood as the fact that $\mathbf{x}$ is more ordered than $\mathbf{y}$, and that $\mathbf{x}$ has a larger norm than $\mathbf{y}$. It is defined as follows:

Definition 3.5 (Sub-majorization for finite-dimensional vectors)**.** *The vector* $\mathbf{x} \in \mathbb{R}^N$ *weak-majorizes from below (sub-majorizes) the vector* $\mathbf{y} \in \mathbb{R}^N$*, written* $\mathbf{x} \succ_{\mathrm{w}} \mathbf{y}$*, is equivalent to any of these three*

conditions

- $$\sum_{i=1}^{N} \varphi(x_i) \geq \sum_{i=1}^{N} \varphi(y_i) \qquad \forall \varphi \quad \textit{convex increasing} \tag{3.41}$$

- $$\Phi_t^+(\mathbf{x}) \geq \Phi_t^+(\mathbf{y}) \qquad \forall t \in \mathbb{R} \tag{3.42}$$

- $$S_k^{\downarrow}(\mathbf{x}) \geq S_k^{\downarrow}(\mathbf{y}) \qquad \forall k \in \{1, ..., N\} \tag{3.43}$$

And these three conditions are equivalent to each other.

Remember that $\Phi_t^+(\mathbf{x}) + \Phi_t^-(\mathbf{x}) = S_N(\mathbf{x}) - Nt$ and $S_k^{\downarrow}(\mathbf{x}) + S_{N-k}^{\uparrow}(\mathbf{x}) = S_N(\mathbf{x})$. Using this, Conditions (3.42) and (3.43) can be written equivalently as:

$$\Phi_t^-(\mathbf{x}) - S_N(\mathbf{x}) \leq \Phi_t^-(\mathbf{y}) - S_N(\mathbf{y}) \qquad \forall t \in \mathbb{R} \tag{3.44}$$

$$S_k^{\uparrow}(\mathbf{x}) - S_N(\mathbf{x}) \leq S_k^{\uparrow}(\mathbf{y}) - S_N(\mathbf{y}) \qquad \forall k \in \{1, ..., N\} \tag{3.45}$$

It is also possible to formulate an hybrid equivalent condition to (3.42) as:

$$\begin{cases} \Phi_t^+(\mathbf{x}) \geq \Phi_t^+(\mathbf{y}) & \forall t \geq t_0 \\ \Phi_t^-(\mathbf{x}) - S_N(\mathbf{x}) \leq \Phi_t^-(\mathbf{y}) - S_N(\mathbf{y}) & \forall t \leq t_0 \end{cases} \tag{3.46}$$

where $t_0 \in \mathbb{R}$. And similarly, we can formulate an hybrid equivalent condition to (3.43) as:

$$\begin{cases} S_k^{\downarrow}(\mathbf{x}) \geq S_k^{\downarrow}(\mathbf{y}) & \forall k \in \{1, ..., k_1\} \\ S_k^{\uparrow}(\mathbf{x}) - S_N(\mathbf{x}) \leq S_k^{\uparrow}(\mathbf{y}) - S_N(\mathbf{y}) & \forall k \in \{1, ..., k_2\} \end{cases} \tag{3.47}$$

such that $k_1, k_2 \in \mathbb{N}_0$ and $k_1 + k_2 = N$.

Weak-majorization from above (super-majorization)

The relation "$\mathbf{x}$ super-majorizes $\mathbf{y}$" can be understood as the fact that $\mathbf{x}$ is more ordered than $\mathbf{y}$, and that $\mathbf{x}$ has a lower norm than $\mathbf{y}$. It is defined as follows:

Definition 3.6 (Super-majorization for finite-dimensional vectors)**.** *The vector $\mathbf{x} \in \mathbb{R}^N$ weak-majorizes from above (super-majorizes) the vector $\mathbf{y} \in \mathbb{R}^N$, written $\mathbf{x} \succ^{\mathrm{w}} \mathbf{y}$, is equivalent to any of these three conditions*

- $$\sum_{i=1}^{N} \varphi(x_i) \geq \sum_{i=1}^{N} \varphi(y_i) \qquad \forall \varphi \quad \textit{convex decreasing} \tag{3.48}$$

- $$\Phi_t^-(\mathbf{x}) \leq \Phi_t^-(\mathbf{y}) \qquad \forall t \in \mathbb{R} \tag{3.49}$$

- $$S_k^{\uparrow}(\mathbf{x}) \leq S_k^{\uparrow}(\mathbf{y}) \qquad \forall k \in \{1, ..., N\} \tag{3.50}$$

And these three conditions are equivalent to each other.

Similarly to the case of sub-majorization, we can rewrite conditions (3.49) and (3.50) as:

$$\Phi_t^+(\mathbf{x}) - S_N(\mathbf{x}) \geq \Phi_t^+(\mathbf{y}) - S_N(\mathbf{y}) \qquad \forall t \in \mathbb{R} \tag{3.51}$$

$$S_k^{\downarrow}(\mathbf{x}) - S_N(\mathbf{x}) \geq S_k^{\downarrow}(\mathbf{y}) - S_N(\mathbf{y}) \qquad \forall k \in \{1, ..., N\} \tag{3.52}$$

We can also formlulate an hybrid equivalent condition to (3.49) as:

$$\begin{cases} \Phi_t^+(\mathbf{x}) - S_N(\mathbf{x}) \geq \Phi_t^+(\mathbf{y}) - S_N(\mathbf{y}) & \forall t \geq t_0 \\ \Phi_t^-(\mathbf{x}) \leq \Phi_t^-(\mathbf{y}) & \forall t \leq t_0 \end{cases} \tag{3.53}$$

where $t_0 \in \mathbb{R}$. We also give an hybrid equivalent condition to (3.50) as:

$$\begin{cases} S_k^{\downarrow}(\mathbf{x}) - S_N(\mathbf{x}) \geq S_k^{\downarrow}(\mathbf{y}) - S_N(\mathbf{y}) & \forall k \in \{1, ..., k_1\} \\ S_k^{\uparrow}(\mathbf{x}) \leq S_k^{\uparrow}(\mathbf{y}) & \forall k \in \{1, ..., k_2\} \end{cases} \tag{3.54}$$

such that $k_1, k_2 \in \mathbb{N}_0$ and $k_1 + k_2 = N$.

Majorization

We now come to standard majorization relations. The relation "$\mathbf{x}$ majorizes $\mathbf{y}$" can be understood as the fact that $\mathbf{x}$ is more ordered than $\mathbf{y}$, and their norms are equal. It is defined as follows:

Definition 3.7 (Majorization for finite-dimensional vectors)**.** *The vector $\mathbf{x} \in \mathbb{R}^N$ majorizes the vector $\mathbf{y} \in \mathbb{R}^N$, written $\mathbf{x} \succ \mathbf{y}$, is equivalent to any of these conditions:*

- $$\sum_{i=1}^{N} \varphi(x_i) \geq \sum_{i=1}^{N} \varphi(y_i) \qquad \forall \varphi \quad \textit{convex} \tag{3.55}$$
- $$\mathbf{x} \succ_{\mathrm{w}} \mathbf{y} \qquad \textit{and} \qquad \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} y_i \tag{3.56}$$
- $$\mathbf{x} \succ^{\mathrm{w}} \mathbf{y} \qquad \textit{and} \qquad \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} y_i \tag{3.57}$$

And these three conditions are equivalent to each other.

As a consequence, a relation of majorization implies all the properties implied by a relation of sub-majorization and a relation of super-majorization. Note also, that if $\mathbf{x} \succ_{\mathrm{w}} \mathbf{y}$ and $\mathbf{x} \succ^{\mathrm{w}} \mathbf{y}$, then $\mathbf{x}$ and $\mathbf{y}$ have the same normalization so that $\mathbf{x} \succ \mathbf{y}$. Let us conclude this section by formulating two equivalent conditions for a majorization relation between two vectors. In the same manner as we did for weak-majorization, we derive the following equivalent condition for majorization:

$$\mathbf{x} \succ \mathbf{y} \qquad \Leftrightarrow \qquad \begin{cases} S_N(\mathbf{x}) = S_N(\mathbf{y}) & \\ \Phi_t^+(\mathbf{x}) \geq \Phi_t^+(\mathbf{y}) & \forall t \geq t_0 \\ \Phi_t^-(\mathbf{x}) \leq \Phi_t^-(\mathbf{y}) & \forall t \leq t_0 \end{cases} \tag{3.58}$$

where $t_0 \in \mathbb{R}$. Then, we proceed similarly and fives another equivalent condition as follows:

$$\mathbf{x} \succ \mathbf{y} \quad \Leftrightarrow \quad \begin{cases} S_N(\mathbf{x}) = S_N(\mathbf{y}) & \\ S_k^{\downarrow}(\mathbf{x}) \geq S_k^{\downarrow}(\mathbf{y}) & \forall k \in \{1, ..., k_1\} \\ S_k^{\uparrow}(\mathbf{x}) \leq S_k^{\uparrow}(\mathbf{y}) & \forall k \in \{1, ..., k_2\} \end{cases} \tag{3.59}$$

where $k_1, k_2 \in \mathbb{N}_0$ and $k_1 + k_2 = N$.

On the contrary to weak-majorization, a relation of majorization imposes that the two vectors are normalized to the same value. Note that when $\mathbf{x}$ and $\mathbf{y}$ are normalized to the same value, the three majorization are equivalent.

Now that we have properly defined a majorization relation, we can introduce a type of function which play a particular role with respect to majorization. We define Schur-convex and Schur-concave functions as follows:

Definition 3.8 (Schur-convex/concave function)**.** *The function $\Phi : \mathbb{R}^n \mapsto \mathbb{R}$ is said to be Schur-convex if and only if $\mathbf{x} \succ \mathbf{y}$ implies $\Phi(\mathbf{x}) \geq \Phi(\mathbf{y})$:*

$$\mathbf{x} \succ \mathbf{y} \quad \Rightarrow \quad \Phi(\mathbf{x}) \geq \Phi(\mathbf{y}) \tag{3.60}$$

Conversely, Φ is said to be Schur-concave if $-\Phi$ is Schur-convex.

As it appears, Schur-convex and Scur-concave functions are intrinsically linked to majorization. Let us note that the functions $S_k^{\downarrow}$ and Φ_t^+ are Schur-convex. Conversely, the functions $S_k^{\uparrow}$ and Φ_t^- are Schur-concave. As a consequence, we can write the following equivalence:

$$\mathbf{x} \succ \mathbf{y} \quad \Leftrightarrow \quad \Phi(\mathbf{x}) \geq \Phi(\mathbf{y}) \quad \forall \Phi \text{ Schur-convex} \tag{3.61}$$

Finally, let us conclude this section with some general considerations regarding majorization. Majorization is said to be a pre-order, because it lacks several properties to make it an order. First, a relation of majorization is not anti-symmetric. Indeed, when both $\mathbf{x} \succ \mathbf{y}$ and $\mathbf{x} \prec \mathbf{y}$ hold, this does not imply that $\mathbf{x} = \mathbf{y}$, but simply implies that $\mathbf{x} \equiv \mathbf{y}$. Level-equivalent vectors do not need to be equal. Secondly, it can happen that neither $\mathbf{x} \succ \mathbf{y}$ or $\mathbf{x} \prec \mathbf{y}$ hold. In that case, $\mathbf{x}$ and $\mathbf{y}$ are said to be incomparable, and we will write it as $\mathbf{x} \not\gtrless \mathbf{y}$. The considerations we have just made are not limited to the present section, and concern in general discrete and continuous majorization, for finite and infinite-size domain.

3.2 Continuous majorization on a finite-size domain

We now turn our attention to continuous distributions. In a first time, we will consider distribution that are defined over a finite domain. It should be noted that the definitions of majorization for continuous distribution defined over a finite domain have been studied in detail in [48]. Mostly, each property is the continuous analog of discrete majorization for finite-dimensional vectors. We consider a domain $\mathbb{A}$ that is a subset of $\mathbb{R}^n$ with finite size.

$$f : \mathbb{A} \mapsto \mathbb{R}, \qquad \mathbb{A} \subset \mathbb{R}^n, \qquad \nu(\mathbb{A}) < \infty \tag{3.62}$$

where ν is the Lebesgue measure which measures the size of a set. For example, $\mathbb{A}$ could be the interval $[0, 1]$. We consider continuous distributions that can take positive and negative values, and which are not normalized to a precise value. We only impose that the integral of the distribution converges absolutely over $\mathbb{A}$, and that the distribution does not have infinite discontinuities:

$$\int_{\mathbb{A}} |f(\mathbf{r})| \mathrm{d}\mathbf{r} < \infty, \qquad |f(\mathbf{r})| < \infty \quad \forall \mathbf{r} \tag{3.63}$$

Note that sometimes we will omit to write $\mathbb{A}$ under the integral. When no boundaries are specified for the integration variable $\mathbf{r}$, it means that the integral is carried over the whole domain $\mathbb{A}$.

3.2.1 Level-equivalence

The idea of level-equivalent functions follows the idea of level-equivalent vectors. In simple terms, we say that two distributions are level-equivalent if they are related by a bijection over their domain. The notion of level-equivalence finds an appropriate formulation through level-functions. Similarly to the discrete case, we define two versions.

Definition 3.9 (Level-functions of a distribution)**.** *The upper level-function $m_f^+(t)$ of a distribution f associate to any real value t the size of the domain of f that has an image greater or equal to t:*

$$m_f^+(t) = \nu\left(\{\mathbf{r} : f(\mathbf{r}) \geq t\}\right) \tag{3.64}$$

Conversely, the lower level-function $m_f^-(t)$ of a distribution f assocaite to any real value t the size of the domain of f that has an image lower or equal to t:

$$m_f^-(t) = \nu\left(\{\mathbf{r} : f(\mathbf{r}) \leq t\}\right) \tag{3.65}$$

From their definition, we understand that these level-functions are formally equal to the following expression:

$$m_f^+(t) = \int \Theta\left(f(\mathbf{r}) - t\right) \mathrm{d}\mathbf{r}, \qquad m_f^-(t) = \int \Theta\left(t - f(\mathbf{r})\right) \mathrm{d}\mathbf{r}. \tag{3.66}$$

We now introduce the upper and lower density-level function as the respective derivative of the upper and lower level-functions.

$$\mu_f^+(t) = -\frac{\mathrm{d}}{\mathrm{d}t} m_f^+(t), \qquad \mu_f^-(t) = \frac{\mathrm{d}}{\mathrm{d}t} m_f^-(t) \tag{3.67}$$

In the present case of distribution defined over a finite support, these two level-density functions are equal, and we will simply call them as $\mu_f(t)$. Indeed, it can be seen that

$$\mu_f(t) = \mu_f^+(t) = \mu_f^-(t) = \int \delta\left(f(\mathbf{r}) - t\right) \mathrm{d}\mathbf{r} \tag{3.68}$$

We can now define the notion of level-equivalence for functions defined over a finite-size domain.

Definition 3.10 (Level-equivalence for finite-domain functions)**.** *Let $f, g \in \mathbb{R}^{\mathbb{A}}$, where $\mathbb{A}$ is such that $\nu(\mathbb{A}) < \infty$. We say that f is level-equivalent to g, written $f \equiv g$ if and only if their level-functions are equal.*

$$f \equiv g \qquad \Leftrightarrow \qquad m_f^+(t) = m_g^+(t) \qquad \forall t \tag{3.69}$$

3.2.2 Symmetric functionals

A symmetric functional follows the same idea as what we explained for symmetric function in the case of vectors. For example, let us build functional $\Phi : \mathbb{R}^{\mathbb{A}} \mapsto \mathbb{R}$ in that particular manner:

$$\Phi(f) = \int \varphi\left(f(\mathbf{r})\right) \mathrm{d}\mathbf{r} \tag{3.70}$$

$$= \int \varphi(t) \mu_f(t) \mathrm{d}t \tag{3.71}$$

As it appears, that functional is invariant for level-equivalent distributions since it only depends on the level-density function of f. More generally, we define a symmetric functional as follows:

Definition 3.11 (Symmetric functional)**.** *The functional $\Phi : \mathbb{R}^{\mathbb{A}} \mapsto \mathbb{R}$ is said to be symmetric if and only if it takes the same value for level-equivalent distributions.*

$$f \equiv g \qquad \Rightarrow \qquad \Phi(f) = \Phi(g) \tag{3.72}$$

.

Similarly as we did in the disrete case, we define Φ_t^+ and Φ_t^-:

$$\Phi_t^+(f) = \int_{\mathbb{A}} \gamma_t^+\left(f(\mathbf{r})\right) \mathrm{d}\mathbf{r}, \qquad \Phi_t^-(f) = \int_{\mathbb{A}} \gamma_t^-\left(f(\mathbf{r})\right) \mathrm{d}\mathbf{r} \tag{3.73}$$

Notice that we use the same notation as in Equation 3.14, but no confusion is possible since they don't take the same type of arguments (vectors and functions). A similar derivation as (3.23) and (3.28) yields:

$$\Phi_t^+(f) = \int_t^{\infty} m_f^+(u)\mathrm{d}u, \qquad \Phi_t^-(f) = \int_{-\infty}^{t} m_f^-(u)\mathrm{d}u \tag{3.74}$$

Note that similarly to the discrete case, Φ_t^+ and Φ_t^- are related as:

$$\Phi_t^+(f) + \Phi_t^-(f) = \int_{\mathbb{A}} f(\mathbf{r})\mathrm{d}\mathbf{r} - \nu\left(\mathbb{A}\right) t \tag{3.75}$$

Let $f : \mathbb{A} \mapsto \mathbb{R}$ and $g : \mathbb{A} \mapsto \mathbb{R}$ be two distributions defined over a finite domain $\mathbb{A}$. Then, the statement f is level-equivalent to g, written $f \equiv g$, is equivalent to any of these properties:

$$f \equiv g \quad \Leftrightarrow \quad m_f^+(t) = m_g^+(t) \qquad \forall t \in \mathbb{R} \tag{3.76}$$

$$\Leftrightarrow \quad m_f^-(t) = m_g^-(t) \qquad \forall t \in \mathbb{R} \tag{3.77}$$

$$\Leftrightarrow \quad \mu_f(t) = \mu_g(t) \qquad \forall t \in \mathbb{R} \tag{3.78}$$

$$\Leftrightarrow \quad \int \varphi(f(\mathbf{r}))\mathrm{d}\mathbf{r} = \int \varphi(g(\mathbf{r}))\mathrm{d}\mathbf{r} \qquad \forall \varphi \tag{3.79}$$

$$\Leftrightarrow \quad \Phi(f) = \Phi(g) \qquad \forall \Phi \text{ symmetric} \tag{3.80}$$

3.2.3 Rearrangements and cumulative integrals

The decreasing rearrangement of a continuous function is easy to understand. It is more difficult to describe in mathematical terms, but informally one can easily picture a function sorted from the maximum to the minimum. To define it properly, we need to introduce a notion of distance, which will be embodied by the norm of a vector. We will consider that a point is further than another one if its distance from the origin is greater. Then, we define $V_{\mathbb{A}}(s)$ as the volume of points of $\mathbb{A}$ that are at a distance lower or equal to s from the origin.

$$V_{\mathbb{A}}(s) = \nu\left(\{\mathbf{r} \in \mathbb{A} : \|\mathbf{r}\| \leq s\}\right) \tag{3.81}$$

$V_{\mathbb{A}}(s)$ is the volume of points in $\mathbb{A}$ that have a norm lower or equal to s.

Also, as previously, we define the inverse function of the level-function as:

$$\left(m_f^+\right)^{-1} = M_f^+, \qquad \left(m_f^-\right)^{-1} = M_f^-, \tag{3.82}$$

which can also be computed as

$$M_f^+(u) = \max\left(\{t : m_f^+(t) \geq u\}\right), \qquad M_f^-(u) = \min\left(\{t : m_f^-(t) \geq u\}\right). \tag{3.83}$$

The decreasing rearrangement of a function f is obtained when sorting f by decreasing order, so that highest values are the closest to the origin, and lowest values the furthest. Conversely, the decreasing rearrangement of a function f is obtained when sorting f by increasing order, so that the lowest values are the closest to the origin, and highest values the furthest. We define them as follows:

Definition 3.12 (Rearrangements of a function)**.** *Let f be a continuous function defined over the domain $\mathbb{A}$. The decreasing rearrangement of f, written $f^{\downarrow}$, is a radial function which is monotonically decreasing from the origin, which has the same upper level-function as the original function. It is defined as:*

$$f^{\downarrow}(\mathbf{r}) = M_f^+ \left(V_{\mathbb{A}}(\|\mathbf{r}\|)\right), \tag{3.84}$$

The increasing rearrangement of f, written $f^{\uparrow}$, is a radial function which is monotonically increasing from the origin, which has the same lower level-function as the original function. It is defined as:

$$f^{\uparrow}(\mathbf{r}) = M_f^- \left(V_{\mathbb{A}}(\|\mathbf{r}\|)\right), \tag{3.85}$$

It sould be noted that $f^{\downarrow}$ and $f^{\uparrow}$ are defined on the same domain as f, namely $\mathbb{A}$. We say that a function is radial when the value it takes at some value $\mathbf{r}$ only depends on the norm $\|\mathbf{r}\|$. We say that a function is radial decreasing if it is radial and decreases when moving away from the origin. Conversely we say that a function is radial increasing if it is radial and increases when moving away from the origin.

As it appears, for functions defined over a finite-size domain, the decreasing and increasing rearrangements are level-equivalent to the original distribution, and we write $f \equiv f^{\downarrow} \equiv f^{\uparrow}$.

We define the cumulative integral S_s of a function f as the integration of f over a subset of $\mathbb{A}$ determined continuously by the real parameter s: The cumulative integral plays the role of the cumulative sum in a continuous setting. In a similar fashion, we define the decreasing cumulative integral and increasing integral as the cumulative integral of respectively the increasing and decreasing rearrangement of f.

Definition 3.13 (Cumulative integrals of a function)**.** *Let $f : \mathbb{A} \mapsto \mathbb{R}$ be a distribution. The cumulative integral $S_s : \mathbb{R}^{\mathbb{A}} \mapsto \mathbb{R}$ of f is the integral of f over a subset of $\mathbb{A}$ centered around the origin with measure s:*

$$S_s(f) = \int\limits_{\substack{\mathbf{r}\in\mathbb{A} \\ V_{\mathbb{A}}(\|\mathbf{r}\|)\leq s}} f(\mathbf{r})\mathrm{d}\mathbf{r}, \tag{3.86}$$

The decreasing cumulative integral $S_s^{\downarrow} : \mathbb{R}^{\mathbb{A}} \mapsto \mathbb{R}$ of f is the cumulative integral of its decreasing rearrangement $f^{\downarrow}$.

$$S_s^{\downarrow}(f) = \int\limits_{\substack{\mathbf{r}\in\mathbb{A} \\ V_{\mathbb{A}}(\|\mathbf{r}\|)\leq s}} f^{\downarrow}(\mathbf{r})\mathrm{d}\mathbf{r} \tag{3.87}$$

The increasing cumulative integral $S_s^{\uparrow} : \mathbb{R}^{\mathbb{A}} \mapsto \mathbb{R}$ of f is the cumulative integral of its increasing rearrangement $f^{\uparrow}$.

$$S_s^{\uparrow}(f) = \int\limits_{\substack{\mathbf{r}\in\mathbb{A} \\ V_{\mathbb{A}}(\|\mathbf{r}\|)\leq s}} f^{\uparrow}(\mathbf{r})\mathrm{d}\mathbf{r} \tag{3.88}$$

Obviously, we have that $S_s^{\downarrow}(f) = S_s(f^{\downarrow})$ and that $S_s^{\uparrow}(f) = S_s(f^{\uparrow})$. The increasing and decreasing cumulative integrals are symmetric functionals. Note that we have the framing $S_s^{\uparrow}(f) \leq S_s(f) \leq S_s^{\downarrow}(f)$. Similarly to the discrete case, we can establish a relation between the decreasing and increasing cumulative integrals, which reads as follows:

$$S_s^{\downarrow}(f) + S_{\nu(\mathbb{A})-s}^{\uparrow}(f) = \int_{\mathbb{A}} f(\mathbf{r})\mathrm{d}\mathbf{r}. \tag{3.89}$$

When the parameter of the cumulative integral is equal to $\nu(\mathbb{A})$, the integration is carried over the whole domain $\mathbb{A}$, so that we have the equality $S_{\nu(\mathbb{A})} = S_{\nu(\mathbb{A})}^{\downarrow} = S_{\nu(\mathbb{A})}^{\uparrow} = \int_{\mathbb{A}} f(\mathbf{r})\mathrm{d}\mathbf{r}$.

3.2.4 Relations of majorization

With the formalism that we have introduced, the definitions of continuous majorization follow exactly the same pattern as for discrete majorization.

Weak-majorization from below (sub-majorization)

A relation of weak-majorization can be established between two functions of different normalization. The relation "f sub-majorizes g" can be understood as the fact that f is more ordered than g, and that f has a larger normalization than g. It is defined as follows:

Definition 3.14 (Sub-majorization for finite-domain distributions)**.** *The distribution $f : \mathbb{A} \mapsto \mathbb{R}$ weak-majorizes from below (sub-majorizes) the distribution $g : \mathbb{A} \mapsto \mathbb{R}$, written $f \succ_{\mathrm{w}} g$, is equivalent to any of these three conditions*

- $$\int_{\mathbb{A}} \varphi(f(\mathbf{r}))\mathrm{d}\mathbf{r} \geq \int_{\mathbb{A}} \varphi(g(\mathbf{r}))\mathrm{d}\mathbf{r} \qquad \forall \varphi \quad \textit{convex increasing} \tag{3.90}$$

- $$\Phi_t^+(f) \geq \Phi_t^+(g) \qquad \forall t \in \mathbb{R} \tag{3.91}$$

- $$S_s^{\downarrow}(f) \geq S_s^{\downarrow}(g) \qquad \forall s : \quad 0 \leq s \leq \nu(\mathbb{A}) \tag{3.92}$$

And these three conditions are equivalent to each other. Remember that $\mathbb{A}$ is finite, so that $\nu(\mathbb{A}) < \infty$.

Remember that $\Phi_t^+(f) + \Phi_t^-(f) = S_{\nu(\mathbb{A})}(f) + \nu(\mathbb{A})t$ and that $S_s^{\downarrow}(f) + S^{\uparrow}_{\nu(\mathbb{A})-s}(f) = S_{\nu(\mathbb{A})}(f)$. It is thus possible to formulate Conditions (3.91) and (3.91) as follows:

$$\Phi_t^-(f) - S_{\nu(\mathbb{A})}(f) \leq \Phi_t^-(g) - S_{\nu(\mathbb{A})}(g) \qquad \forall t \in \mathbb{R} \tag{3.93}$$

$$S_s^{\uparrow}(f) - S_{\nu(\mathbb{A})}(f) \leq S_s^{\uparrow}(g) - S_{\nu(\mathbb{A})}(g) \qquad \forall s : 0 \leq s \leq \nu(\mathbb{A}) \tag{3.94}$$

Similarly to what we did for the discrete case, we present a hybrid equivalent condition to (3.91):

$$\begin{cases} \Phi_t^+(f) \geq \Phi_t^+(g) & \forall t \geq t_0 \\ \Phi_t^-(f) - S_{\nu(\mathbb{A})}(f) \geq \Phi_t^-(g) - S_{\nu(\mathbb{A})}(g) & \forall t \leq t_0 \end{cases} \tag{3.95}$$

where $t_0 \in \mathbb{R}$. We also give a hybrid equivalent condition to (3.91):

$$\begin{cases} S_s^{\downarrow}(f) \geq S_s^{\downarrow}(g) & \forall s : \quad 0 \leq s \leq s_1 \\ S_s^{\uparrow}(f) - S_{\nu(\mathbb{A})}(f) \leq S_s^{\uparrow}(g) - S_{\nu(\mathbb{A})}(g) & \forall s : \quad 0 \leq s \leq s_2 \end{cases} \tag{3.96}$$

such that $s_1, s_2 \geq 0$ and $s_1 + s_2 = \nu(\mathbb{A})$.

Weak-majorization from above (super-majorization)

The relation "f super-majorizes g" can be understood as the fact that f is more ordered than g, and that f has a lower normalization than g. It is defined as follows:

Definition 3.15 (Super-majorization for finite-domain distributions)**.** *The distribution $f : \mathbb{A} \mapsto \mathbb{R}$ weak-majorizes from above (super-majorizes) the distribution $g : \mathbb{A} \mapsto \mathbb{R}$, written $f \succ^{\mathrm{w}} g$, is equivalent to any of these three conditions:*

- $$\int_{\mathbb{A}} \varphi(f(\mathbf{r}))\mathrm{d}\mathbf{r} \geq \int_{\mathbb{A}} \varphi(g(\mathbf{r}))\mathrm{d}\mathbf{r} \qquad \forall \varphi \quad \textit{convex decreasing} \tag{3.97}$$

- $$\Phi_t^-(f) \leq \Phi_t^-(g) \qquad \forall t \in \mathbb{R} \tag{3.98}$$

- $$S_s^{\uparrow}(f) \leq S_s^{\uparrow}(g) \qquad \forall s : \quad 0 \leq s \leq \nu(\mathbb{A}) \tag{3.99}$$

And these three conditions are equivalent to each other. Remember that $\mathbb{A}$ is finite, so that $\nu(\mathbb{A}) < \infty$.

Conditions (3.98) and (3.99) can be equivalently formulated as:

$$\Phi_t^+(f) - S_{\nu(\mathbb{A})}(f) \geq \Phi_t^+(g) - S_{\nu(\mathbb{A})}(g) \qquad \forall t \tag{3.100}$$

$$S_s^{\downarrow}(f) - S_{\nu(\mathbb{A})}(f) \geq S_s^{\downarrow}(g) - S_{\nu(\mathbb{A})}(g) \qquad \forall s : 0 \leq s \leq \nu(\mathbb{A}) \tag{3.101}$$

And similarly, we can obtain a hybrid condition equivalent to (3.98) as follows:

$$\begin{cases} \Phi_t^+(f) - S_{\nu(\mathbb{A})}(f) \geq \Phi_t^+(g) - S_{\nu(\mathbb{A})}(g) & \forall t \geq t_0 \\ \Phi_t^-(f) \geq \Phi_t^-(g) & \forall t \leq t_0 \end{cases} \tag{3.102}$$

where $t_0 \in \mathbb{R}$. And we find a hybrid equivalent condition to (3.99) as:

$$\begin{cases} S_s^{\downarrow}(f) \geq S_s^{\downarrow}(g) & \forall s : 0 \leq s \leq s_1 \\ S_s^{\uparrow}(f) - S_{\nu(\mathbb{A})}(f) \leq S_s^{\uparrow}(g) - S_{\nu(\mathbb{A})}(g) & \forall s : 0 \leq s \leq s_2 \end{cases} \tag{3.103}$$

where $s_1, s_2 \geq 0$ and $s_1 + s_2 = \nu(\mathbb{A})$.

Majorization

We now come to standard majorization relations. The relation "f majorizes g" can be understood as the fact that f is more ordered than g, and their norms are equal. It is defined as follows:

Definition 3.16 (Majorization for finite-domain distribution)**.** *The distribution $f : \mathbb{A} \mapsto \mathbb{R}$ majorizes the distribution $g : \mathbb{A} \mapsto \mathbb{R}$, written $f \succ g$, is equivalent to any of these conditions:*

- $$\int_{\mathbb{A}} \varphi(f(\mathbf{r}))\mathrm{d}\mathbf{r} \geq \int_{\mathbb{A}} \varphi(g(\mathbf{r}))\mathrm{d}\mathbf{r} \qquad \forall \varphi \quad \textit{convex} \tag{3.104}$$

- $$f \succ_{\mathrm{w}} g \qquad \textit{and} \qquad \int_{\mathbb{A}} f(\mathbf{r})\mathrm{d}\mathbf{r} = \int_{\mathbb{A}} g(\mathbf{r})\mathrm{d}\mathbf{r} \tag{3.105}$$

- $$f \succ^{\mathrm{w}} g \qquad \textit{and} \qquad \int_{\mathbb{A}} f(\mathbf{r})\mathrm{d}\mathbf{r} = \int_{\mathbb{A}} g(\mathbf{r})\mathrm{d}\mathbf{r} \tag{3.106}$$

And these three conditions are equivalent to each other. Remember that $\mathbb{A}$ is finite, so that $\nu(\mathbb{A}) < \infty$.

As a consequence, a relation of majorization implies all the properties implied by a relation of sub-majorization and a relation of super-majorization. Note also, that if $f \succ_{\mathrm{w}} g$ and $f \succ^{\mathrm{w}} g$, then f and g have the same normalization so that $f \succ g$.

As we have done for weak-majorization, it is possible to formulate hybrid conditions that use both families of functionals Φ_t^+ and Φ_t^-:

$$f \succ g \qquad \Leftrightarrow \qquad \begin{cases} S_{\nu(\mathbb{A})}(f) = S_{\nu(\mathbb{A})}(g) & \\ \Phi_t^+(f) \geq \Phi_t^+(g) & \forall t \geq t_0 \\ \Phi_t^-(f) \leq \Phi_t^-(g) & \forall t \leq t_0 \end{cases} \tag{3.107}$$

where $t_0 \in \mathbb{R}$. It is also possible to formulate hybrid conditions that use both the decreasing cumulative integral $S_s^{\downarrow}$ and $S_s^{\uparrow}$:

$$f \succ g \qquad \Leftrightarrow \qquad \begin{cases} S_{\nu(\mathbb{A})}(f) = S_{\nu(\mathbb{A})}(g) & \\ S_s^{\downarrow}(f) \geq S_s^{\downarrow}(g) & \forall s : 0 \leq s \leq s_1 \\ S_s^{\uparrow}(f) \leq S_s^{\uparrow}(g) & \forall s : 0 \leq s \leq s_2 \end{cases} \tag{3.108}$$

such that $s_1, s_2 \geq 0$ and $s_1 + s_2 = \nu(\mathbb{A})$.

Similarly to the discrete case, we define Schur-convex and Schur-concave functionals as follows.

Definition 3.17 (Schur-convex/concave functional)**.** *The functional $\Phi : \mathbb{R}^{\mathbb{A}} \mapsto \mathbb{R}$ is said to be Schur-convex if and only if $f \succ g$ implies $\Phi(f) \geq \Phi(g)$:*

$$f \succ g \qquad \Rightarrow \qquad \Phi(f) \geq \Phi(g) \tag{3.109}$$

Conversely, Φ is said to be Schur-concave if $-\Phi$ is Schur-convex.

The functionals $S_s^{\downarrow}$ and Φ_t^+ are Schur-convex, while the functionals $S_s^{\uparrow}$ and Φ_t^- are Schur-concave. As a consequence we have the following equivalence:

$$f \succ g \qquad \Leftrightarrow \qquad \Phi(f) \geq \Phi(g) \qquad \forall \Phi \text{ Schur-convex.} \tag{3.110}$$

Finally, let us recall that the considerations we laid at the end of the previous section still hold in the present case. Continuous majorization is a pre-order. As a consequence, if both $f \succ g$ and $f \prec g$ hold, we have that $f \equiv g$ (but not necessarily that $f = g$). When neither $f \succ g$ or $f \prec g$ hold, we say that f and g are incomparable, and we write it as $f \nsim g$.

3.3 Discrete majorization on an infinite-size domain

We are now considering discrete majorization for vector defined over an infinite support. In general, we consider in the present section vectors of the form:

$$\mathbf{x} \in \mathbb{R}^{\mathbb{N}}, \qquad \sum_{i=1}^{\infty} |x_i| < \infty. \tag{3.111}$$

We consider vectors that can take negative values and that are not necessarily normalized to the same value. However, we consider vectors whose absolute norm is finite. Because of the sum of the absolute values of the components is finite, we need for convergence that:

$$\lim_{n \to \infty} x_n = 0. \tag{3.112}$$

3.3.1 Level-equivalence

Another consequence of this is that the vectors of our concerns have an infinite number of components with value arbitrarily close to zero. This means that for any $t < 0$ there is an infinite number of components with value greater than t, and that for any $t > 0$ there is an infinite number of components with value smaller than t. This can be translated in terms of level-function. We define the level-functions of an infinite-dimensional vector in the same manner as we did for finite-dimensional vector, so that we also refer to Definition 3.1. As a concequence of the infinite number of components with value arbitrarily close to zero, we observe the following:

$$m_{\mathbf{x}}^+(t) = \infty \qquad \forall t < 0, \qquad\qquad m_{\mathbf{x}}^-(t) = \infty \qquad \forall t > 0. \tag{3.113}$$

We see that the upper level-function $m_{\mathbf{x}}^+(t)$ is then only properly defined for $t > 0$ and the lower level-function $m_{\mathbf{x}}^-(t)$ for $t < 0$. This means that the upper level-function only contains the information about the positive parts of $\mathbf{x}$, and that the lower level-function only contains the information about its negative parts. A consequence of that observation is that the upper level-function and the lower level-function are not related to each other anymore. The infinite discontinuity in $t = 0$ makes them decorrelated from each other. $m_{\mathbf{x}}^+$ contains all the information about the positive components of $\mathbf{x}$, whereas $m_{\mathbf{x}}^-$ contains all the information about the negative components of $\mathbf{x}$.

When it comes to the upper and lower level-density functions, $\mu_{\mathbf{x}}^+(t)$ is only defined for $t > 0$ and $\mu_{\mathbf{x}}^-(t)$ is only defined for $t < 0$. We define the level-density function $\mu_{\mathbf{x}}$ of an infinite-dimensional vector $\mathbf{x} \in \mathbb{R}^{\mathbb{N}}$ as follows:

$$\mu_{\mathbf{x}}(t) = \begin{cases} \mu_{\mathbf{x}}^+(t) & \forall t > 0 \\ \\ \mu_{\mathbf{x}}^-(t) & \forall t < 0 \end{cases} \tag{3.114}$$

The level-density function encloses all the information about the values taken by the vector $\mathbf{x}$. One can see that $\mu_{\mathbf{x}} = \sum \delta(x_i - t)$. Note that $\mu_{\mathbf{x}}(t)$ has an infinite singularity in $t = 0$, as a consequence of what we have explained.

Checking whether two infinite-dimensional vectors are level-equivalent has te be done by comparing both the upper level-function and lower level-function. It is not sufficient to compare only one of them, as they separately incorporate the information about the positive or negative parts.

Definition 3.18 (Level-equivalence for infinite-dimensional vectors)**.** *Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{\mathbb{N}}$. We say that $\mathbf{x}$ is level-equivalent to $\mathbf{y}$, written $\mathbf{x} \equiv \mathbf{y}$ if and only if their upper level-functions are equal and their lower level-functions are equal.*

$$\mathbf{x} \equiv \mathbf{y} \quad \Leftrightarrow \quad \begin{cases} m^+_{\mathbf{x}}(t) = m^+_{\mathbf{y}}(t) & \forall t > 0 \\ m^-_{\mathbf{x}}(t) = m^-_{\mathbf{y}}(t) & \forall t < 0 \end{cases} \tag{3.115}$$

An interesting observation should be made here. Two vectors can be level-equivalent even if they do not exactly have the same components. To understand this, let us consider the following example:

$$\mathbf{x} = \left(2^{-1}, 2^{-2}, ..., 2^{-n}, ...\right) \qquad \mathbf{y} = \left(0, 2^{-1}, 2^{-2}, ..., 2^{-n}, ...\right) \tag{3.116}$$

One easily understand that $\mathbf{x} \equiv \mathbf{y}$, even if $\mathbf{x}$ has no component with zero value, while $\mathbf{y}$ has one. This is possible because of the fact that $\mathbf{x}$ has a infinite number of components arbitrarily close to zero. For that reason, $\mathbf{x}$ has virtually an infinite number of zero-valued components. Any infinite-dimensional vector satisfying the condition of finite absolute norm can be considered to have an infinite batch of components with value zero.

3.3.2 Symmetric functions

In the case of infinite-dimensional vectors, symmetric functions are defined similarly to Definition 3.3. A function Φ is symmetric if and only if $\Phi(\mathbf{x}) = \Phi(\mathbf{y})$ as soon as $\mathbf{x} \equiv \mathbf{y}$. Let us now consider the following example of symmetric function:

$$\Phi(\mathbf{x}) = \sum_{i=1}^{\infty} \varphi(x_i) = \int \varphi(t)\mu_{\mathbf{x}}(t)\mathrm{d}t. \tag{3.117}$$

Note that because of the infinite discontinuity of $\mu_{\mathbf{x}}(t)$ in $t = 0$, the function $\Phi(\mathbf{x})$ does not always converge towards a finite value. That observation has direct consequences on the functions Φ^+_t and Φ^-_t that we have previously defined. Indeed, it follows that they converge only over a subset of values of t:

$$\Phi^+_t(\mathbf{x}) = \infty \qquad \forall t < 0, \qquad\qquad \Phi^-_t(\mathbf{x}) = \infty \qquad \forall t > 0. \tag{3.118}$$

Note that the upper and lower level-function can be retrieved from Φ^+_t and Φ^-_t using relations (3.23) and (3.23). For two infinite-dimensional vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{\mathbb{N}}$, we can write the following equivalent statements to a relation of level-equivalence:

$$\mathbf{x} \equiv \mathbf{y} \quad \Leftrightarrow \quad \begin{cases} m^+_{\mathbf{x}}(t) = m^+_{\mathbf{y}}(t) & \forall t > 0 \\ m^-_{\mathbf{x}}(t) = m^-_{\mathbf{y}}(t) & \forall t < 0 \end{cases} \tag{3.119}$$

$$\Leftrightarrow \quad \mu_{\mathbf{x}}(t) = \mu_{\mathbf{y}}(t) \qquad \forall t \in \mathbb{R} \tag{3.120}$$

$$\Leftrightarrow \quad \sum_{i=1}^{\infty} \varphi(x_i) = \sum_{i=1}^{\infty} \varphi(y_i) \qquad \forall \varphi \tag{3.121}$$

$$\Leftrightarrow \quad \Phi(\mathbf{x}) = \Phi(\mathbf{y}) \qquad \forall \Phi \text{ symmetric} \tag{3.122}$$

3.3.3 Rearrangements and cumulative sums

The rearrangement of an infinite-dimensional vector presents subtleties that we do not face in the case of a finite-dimensional vector. We say that the vector $\mathbf{y}$ is a rearrangement of the vector $\mathbf{x}$ if each components of $\mathbf{y}$ comes from a component of $\mathbf{x}$, and that each component of $\mathbf{x}$ is used at most one time. Let $\mathbf{y}$ be a rearrangement of $\mathbf{x}$. We define the function $f_{\mathbf{y}\to\mathbf{x}}$ as the function which gets as input an index of $\mathbf{y}$, and gives as output the index of $\mathbf{x}$ from which the component is coming.

$$f_{\mathbf{y}\to\mathbf{x}} : \mathbb{N} \mapsto \mathbb{N} : i_{\mathbf{y}} \mapsto i_{\mathbf{x}} \tag{3.123}$$

That function is such that $x_{f_{\mathbf{y}\to\mathbf{x}}(i)} = y_i$. The function $f_{\mathbf{y}\to\mathbf{x}}$ is invertible and its inverse is $f^{-1}_{\mathbf{y}\to\mathbf{x}} = f_{\mathbf{x}\to\mathbf{y}}$. In the case of finite-dimensional vectors, the function $f_{\mathbf{y}\to\mathbf{x}}$ bijective. Indeed, each component of $\mathbf{y}$ comes from one single component of $\mathbf{x}$, and each component of $\mathbf{x}$ yields one single component of $\mathbf{y}$.

However, in the case of infinite-dimensional vector, the function $f_{\mathbf{y}\to\mathbf{x}}$ is only injective. Indeed, it is possible to construct a vector $\mathbf{y}$ such that each component of $\mathbf{y}$ comes from one single component of $\mathbf{x}$, but such that not all the components of $\mathbf{x}$ are used in $\mathbf{y}$. As an example, let us imagine that $\mathbf{y}$ is the vector made of all the even components of $\mathbf{x}$. In that case, $\mathbf{y}$ is completely defined and contains only components from $\mathbf{x}$, however not all the components of $\mathbf{x}$ are included in $\mathbf{y}$. This observation has the surprising implication that a rearranged vector is not necessarily level-equivalent to the original vector when we are considering infinite-dimensional vectors, as we will show in Figure 3.1.

The decreasing and increasing rearrangements are defined as previously. The vector $\mathbf{x}^{\downarrow}$ (resp. $\mathbf{x}^{\uparrow}$) is a vector made of the components of $\mathbf{x}$ sorted by decreasing order (resp. increasing order). Remember that $\mathbf{x}$ has an infinite number of components arbitrarily close to zero. That means that when rearranged by decreasing order, the negative values of $\mathbf{x}$ will be arranged beyond infinity, and thus be lost. The same happens with the negative values of $\mathbf{x}$ in its increasing rearrangement $\mathbf{x}^{\uparrow}$. In general, we have:

$$\mathbf{x}^{\downarrow} \equiv [\mathbf{x}]^{+} \qquad\qquad \mathbf{x}^{\uparrow} \equiv [\mathbf{x}]^{-} \tag{3.124}$$

where we denote the vector $[\mathbf{x}]^{+}$ as the vector $\mathbf{x}$ where the negative components have been replaced by zero, and the vector $[\mathbf{x}]^{-}$ as the vector $\mathbf{x}$ where the positive components have been replaced by zero.

Let us illustrate this by taking a particular example of vector $\mathbf{x}$. Let us define the vector $\mathbf{x}$ as a vector made of components $x_n = -(-a)^{-n}$, where $a > 1$ is some real parameter. When $n \to \infty$, $x_n \to 0$ and the absolute sum of $\mathbf{x}$ converges towards a real value. Hereafter we build its decreasing and increasing rearrangements, as well as its positive and negative parts:

$$\begin{aligned}
\mathbf{x} &= \left(a^{-1}, -a^{-2}, a^{-3}, -a^{-4}, ..., -(-a)^{-n}, ...\right) \\
\mathbf{x}^{\downarrow} &= \left(a^{-1}, a^{-3}, a^{-5}, a^{-7}, a^{-9}, ..., a^{-2n+1}, ...\right) \\
\mathbf{x}^{\uparrow} &= \left(-a^{-2}, -a^{-4}, -a^{-6}, -a^{-8}, ..., -a^{-2n}, ...\right) \\
[\mathbf{x}]^{+} &= \left(a^{-1}, 0, a^{-3}, 0, a^{-5}, 0, ..., a^{-2n+1}, 0, ...\right) \\
[\mathbf{x}]^{-} &= \left(0, -a^{-2}, 0, -a^{-4}, 0, -a^{-6}, ..., 0, -a^{-2n}, ...\right)
\end{aligned} \tag{3.125}$$

Figure 3.1 illustrates that example for some particular value of a. The example illustrates the fact that in general for infinite-dimensional vectors, $\mathbf{x}^{\downarrow} \not\equiv \mathbf{x}$ neither $\mathbf{x}^{\uparrow} \not\equiv \mathbf{x}$. However, in general we have the relation $\mathbf{x}^{\downarrow} \equiv [\mathbf{x}]^{+}$ and $\mathbf{x}^{\uparrow} \equiv [\mathbf{x}]^{-}$. These observations are consequences of the fact that $\mathbf{x}$ as an infinite number of components arbitrarily close to zero.

We define the cumulative sums identically to the case of finite-dimensional vectors (see Definition 3.4), and we allow the parameter k to become infinity. It should be noted that they depends respectively only on the positive components and negative components of the vector. For finite-dimensional vector, we remember that $S^{\downarrow}_k(\mathbf{x}) + S^{\uparrow}_{N-k}(\mathbf{x}) = S_N(\mathbf{x})$, where $S_N(\mathbf{x})$ is simply the normalization of

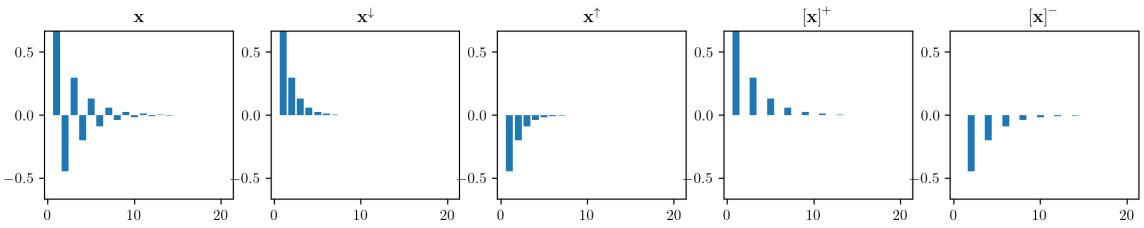

Figure 3.1: Example of infinite-dimensional vector with both positive and negative components built in the manner of (3.125) with $a = 3/2$. We compute from left to right: the original vector $\mathbf{x}$, its decreasing rearrangement $\mathbf{x}^{\downarrow}$, its increasing rearrangement $\mathbf{x}^{\uparrow}$, its positive parts $[\mathbf{x}]^{+}$ and its negative parts $[\mathbf{x}]^{-}$. Only the 20 first components of each vector are plotted, but each vector is infinite-dimensional. The components quickly tends towards zero. From the illustration, it clearly appears that $\mathbf{x} \not\equiv \mathbf{x}^{\downarrow}$ and $\mathbf{x} \not\equiv \mathbf{x}^{\uparrow}$. However, we see that $\mathbf{x}^{\downarrow} \equiv [\mathbf{x}]^{+}$ and $\mathbf{x}^{\uparrow} \equiv [\mathbf{x}]^{-}$.

$\mathbf{x}$. For infinite-dimensional vectors, the decreasing and increasing cumulative sums are not related anymore since N has to be taken to infinity. The information about the decreasing cumulative sum does not give information about the increasing cumulative sum and *vice versa*. In light of these observations, we define the infinite cumulative sum S_{∞}, the infinite decreasing cumulative sum $S_{\infty}^{\downarrow}$ and the infinite increasing cumulative sum $S_{\infty}^{\uparrow}$ as follows:

$$S_{\infty}(\mathbf{x}) = \sum_{i=1}^{\infty} x_i \qquad S_{\infty}^{\downarrow}(\mathbf{x}) = \sum_{x_i>0} x_i \qquad S_{\infty}^{\uparrow}(\mathbf{x}) = \sum_{x_i<0} x_i \tag{3.126}$$

Intuitively, we see that $S_{\infty}^{\downarrow}(\mathbf{x}) + S_{\infty}^{\uparrow}(\mathbf{x}) = S_{\infty}(\mathbf{x})$. That relation is to be compared with the finite-dimensional case in which we have $S_{N/2}^{\downarrow}(\mathbf{x}) + S_{N/2}^{\uparrow}(\mathbf{x}) = S_N(\mathbf{x})$ (supposing N is even). Note also that $S_{\infty}^{\downarrow}(\mathbf{x}) = S_{\infty}\left([\mathbf{x}]^{+}\right)$ and $S_{\infty}^{\uparrow}(\mathbf{x}) = S_{\infty}\left([\mathbf{x}]^{-}\right)$.

We see that, in the case of infinite-dimensional vectors, the positive and negative parts interacts separately. The different objects that we have defined are linked to each other as represented hereafter:

$$\begin{array}{ccccccccc} \mathbf{x}^{\downarrow} & \leftrightarrow & [\mathbf{x}]^{+} & \leftrightarrow & m_{\mathbf{x}}^{+}(t) & \leftrightarrow & \Phi_t^{+}(\mathbf{x}) & \leftrightarrow & S_k^{\downarrow}(\mathbf{x}) \\ \mathbf{x}^{\uparrow} & \leftrightarrow & [\mathbf{x}]^{-} & \leftrightarrow & m_{\mathbf{x}}^{-}(t) & \leftrightarrow & \Phi_t^{-}(\mathbf{x}) & \leftrightarrow & S_k^{\uparrow}(\mathbf{x}) \end{array}$$

3.3.4 Relations of majorization

Let us now build a version of discrete majorization which is compatible with infinite-dimensional vectors taking both positive and negative values. In order to do so, we start from the hybrid formulations we introduced in the finite-dimensional case, and we take the limite of N going towards infinity. This will leads us to the following extended version of discrete majorization.

Weak-majorization from below (sub-majorization)

From the hybrid conditions (3.46) and (3.47) that we have derived earlier, we define sub-majorization for infinite-dimensional vectors as follows:

Definition 3.19 (Sub-majorization for infinite-dimensional vectors)**.** *The vector $\mathbf{x} \in \mathbb{R}^{\mathbb{N}}$ weak-majorizes from below (sub-majorizes) the vector $\mathbf{y} \in \mathbb{R}^{\mathbb{N}}$, written $\mathbf{x} \succ_{\mathrm{w}} \mathbf{y}$, is equivalent to any of these three con-*

ditions:

- $$\sum_{i=1}^{\infty} \varphi(x_i) \geq \sum_{i=1}^{\infty} \varphi(y_i) \qquad \forall \varphi \quad \textit{convex increasing} \tag{3.127}$$

- $$\begin{cases} \Phi_t^+(\mathbf{x}) \geq \Phi_t^+(\mathbf{y}) & \forall t \geq 0 \\ \Phi_t^-(\mathbf{x}) - S_\infty(\mathbf{x}) \leq \Phi_t^-(\mathbf{y}) - S_\infty(\mathbf{y}) & \forall t \leq 0 \end{cases} \tag{3.128}$$

- $$\begin{cases} S_k^\downarrow(\mathbf{x}) \geq S_k^\downarrow(\mathbf{y}) & \forall k \in \mathbb{N} \\ S_k^\uparrow(\mathbf{x}) - S_\infty(\mathbf{x}) \leq S_k^\uparrow(\mathbf{y}) - S_\infty(\mathbf{y}) & \forall k \in \mathbb{N} \end{cases} \tag{3.129}$$

And these three conditions are equivalent to each other.

Weak-majorization from above (super-majorization)

From the hybrid conditions (3.53) and (3.54) that we have derived earlier, we define super-majorization for infinite-dimensional vectors as follows:

Definition 3.20 (Super-majorization for infinite-dimensional vectors)**.** *The vector* $\mathbf{x} \in \mathbb{R}^{\mathbb{N}}$ *weak-majorizes from above (super-majorizes) the vector* $\mathbf{y} \in \mathbb{R}^{\mathbb{N}}$*, written* $\mathbf{x} \succ^{\mathrm{w}} \mathbf{y}$*, is equivalent to any of these three conditions:*

- $$\sum_{i=1}^{\infty} \varphi(x_i) \geq \sum_{i=1}^{\infty} \varphi(y_i) \qquad \forall \varphi \quad \textit{convex decreasing} \tag{3.130}$$

- $$\begin{cases} \Phi_t^+(\mathbf{x}) - S_\infty(\mathbf{x}) \geq \Phi_t^+(\mathbf{y}) - S_\infty(\mathbf{y}) & \forall t \geq 0 \\ \Phi_t^-(\mathbf{x}) \leq \Phi_t^-(\mathbf{y}) & \forall t \leq 0 \end{cases} \tag{3.131}$$

- $$\begin{cases} S_k^\downarrow(\mathbf{x}) - S_\infty(\mathbf{x}) \geq S_k^\downarrow(\mathbf{y}) - S_\infty(\mathbf{y}) & \forall k \in \mathbb{N} \\ S_k^\uparrow(\mathbf{x}) \leq S_k^\uparrow(\mathbf{y}) & \forall k \in \mathbb{N} \end{cases} \tag{3.132}$$

And these three conditions are equivalent to each other.

Majorization

We then come to regular majorization, and define it as previously.

Definition 3.21 (Majorization for infinite-dimensional vectors)**.** *The vector* $\mathbf{x} \in \mathbb{R}^{\mathbb{N}}$ *majorizes the*

vector $\mathbf{y} \in \mathbb{R}^{\mathbb{N}}$*, written* $\mathbf{x} \succ \mathbf{y}$*, is equivalent to any of these three conditions*

- $$\sum_{i=1}^{\infty} \varphi(x_i) \geq \sum_{i=1}^{\infty} \varphi(y_i) \qquad \forall \varphi \quad \textit{convex} \tag{3.133}$$

- $$\mathbf{x} \succ_{\mathrm{w}} \mathbf{y} \qquad \textit{and} \qquad \sum_{i=1}^{\infty} x_i = \sum_{i=1}^{\infty} y_i \tag{3.134}$$

- $$\mathbf{x} \succ^{\mathrm{w}} \mathbf{y} \qquad \textit{and} \qquad \sum_{i=1}^{\infty} x_i = \sum_{i=1}^{\infty} y_i \tag{3.135}$$

And these three conditions are equivalent to each other.

In practice, we will often use the following conditions to check whether two infinite-dimensional vectors majorize each other. The following condition comes from (3.58):

$$\mathbf{x} \succ \mathbf{y} \qquad \Leftrightarrow \qquad \begin{cases} S_\infty(\mathbf{x}) = S_\infty(\mathbf{y}) & \\ \Phi_t^+(\mathbf{x}) \geq \Phi_t^+(\mathbf{y}) & \forall t \geq 0 \\ \Phi_t^-(\mathbf{x}) \leq \Phi_t^-(\mathbf{y}) & \forall t \leq 0 \end{cases} \tag{3.136}$$

The following condition comes from (3.59):

$$\mathbf{x} \succ \mathbf{y} \qquad \Leftrightarrow \qquad \begin{cases} S_\infty(\mathbf{x}) = S_\infty(\mathbf{y}) & \\ S_k^{\downarrow}(\mathbf{x}) \geq S_k^{\downarrow}(\mathbf{y}) & \forall k \in \mathbb{N} \\ S_k^{\uparrow}(\mathbf{x}) \leq S_k^{\uparrow}(\mathbf{y}) & \forall k \in \mathbb{N} \end{cases} \tag{3.137}$$

This concludes this section. Let us mention that the formulation of majorization that we have developed in the present section applies to infinite-dimensional vectors, but it is also valid for finite-dimensional vectors. As such, it constitutes a generalization of Section 3.1.

3.4 Continuous majorization on an infinite-size domain

This section is the continuous counterpart of the previous section. For this reason, we follow the same reasoning as explained above. In this section, we consider continuous distribution defined over a domain $\mathbb{A}$ of infinite measure.

$$f : \mathbb{A} \mapsto \mathbb{R} \qquad \mathbb{A} \subseteq \mathbb{R}^n, \qquad \nu(\mathbb{A}) = \infty \tag{3.138}$$

In practice, $\mathbb{A}$ could be for example $\mathbb{R}^n$ or $\mathbb{R}_+$. We will only consider functions that takes finite values and such that the integration of their absolute value over $\mathbb{A}$ is finite.

$$|f(\mathbf{r})| < \infty \quad \forall \mathbf{r}, \qquad \int_{\mathbb{A}} |f(\mathbf{r})| \mathrm{d}\mathbf{r} < \infty \tag{3.139}$$

A direct consequence of these constraints is that the distribution needs to vanish at infinity.

$$\lim_{\|\mathbf{r}\| \to \infty} f(\mathbf{r}) = 0 \tag{3.140}$$

3.4.1 Level-equivalence

We understand that such a distribution has an infinite portion of its domain arbitrarily close to zero, so that the upper and lower level-functions are only defined over one side of $t = 0$.

$$m_f^+(t) = \infty \qquad \forall t < 0, \qquad\qquad m_f^-(t) = \infty \qquad \forall t > 0. \tag{3.141}$$

We define the level-density function of f as follows:

$$\mu_f(t) = \begin{cases} \mu_f^+(t) & \forall t \geq 0 \\ \mu_f^-(t) & \forall t < 0 \end{cases} \tag{3.142}$$

This way, the level-density function incorporates the information about the positive and negative values taken by f. Note that there is always a discontinuity in $t = 0$, and we have $\mu_f(0) = \infty$.

Definition 3.22 (Level-equivalence for infinite-domain functions)**.** *Let $f, g \in \mathbb{R}^{\mathbb{A}}$, where $\mathbb{A}$ is such that $\nu(\mathbb{A}) = \infty$. We say that f is level-equivalent to g, written $f \equiv g$ if and only if their upper level-functions are equal and their lower level-functions are equal.*

$$f \equiv g \qquad \Leftrightarrow \qquad \begin{cases} m_f^+(t) = m_g^+(t) & \forall t > 0 \\ m_f^-(t) = m_g^-(t) & \forall t < 0 \end{cases} \tag{3.143}$$

3.4.2 Symmetric functionals

In the case of infinite-domain functions, symmetric functionals are defined similarly to Definition 3.11. A functional Φ is symmetric if and only if $\Phi(f) = \Phi(g)$ as soon as $f \equiv g$. Let us now consider the following example of symmetric function:

$$\Phi(f) = \int_{\mathbb{A}} \varphi(f(\mathbf{r}))\mathrm{d}\mathbf{r} = \int \varphi(t)\mu_f(t)\mathrm{d}t. \tag{3.144}$$

Note that because of the infinite discontinuity of $\mu_f(t)$ in $t = 0$, the function $\Phi(\mathbf{x})$ does not always converge towards a finite value. That observation has direct consequences on the functions Φ_t^+ and Φ_t^- that we have previously defined. Indeed, it follows that they converge only over a subset of values of t:

$$\Phi_t^+(f) = \infty \qquad \forall t < 0, \qquad\qquad \Phi_t^-(f) = \infty \qquad \forall t > 0. \tag{3.145}$$

Note that the upper and lower level-function can be retrieved from Φ_t^+ and Φ_t^- using relations (3.23) and (3.23). For two infinite-domain functions $f, g \in \mathbb{R}^{\mathbb{A}}$, we can write the following equivalent statements to a relation of level-equivalence:

$$f \equiv g \qquad \Leftrightarrow \qquad \begin{cases} m_f^+(t) = m_g^+(t) & \forall t > 0 \\ m_f^-(t) = m_g^-(t) & \forall t < 0 \end{cases} \tag{3.146}$$

$$\Leftrightarrow \qquad \mu_f(t) = \mu_g(t) \qquad \forall t \in \mathbb{R} \tag{3.147}$$

$$\Leftrightarrow \qquad \int \varphi(f(\mathbf{r}))\mathrm{d}\mathbf{r} = \int \varphi(g(\mathbf{r}))\mathrm{d}\mathbf{r} \qquad \forall \varphi \tag{3.148}$$

$$\Leftrightarrow \qquad \Phi(f) = \Phi(g) \qquad \forall \Phi \text{ symmetric} \tag{3.149}$$

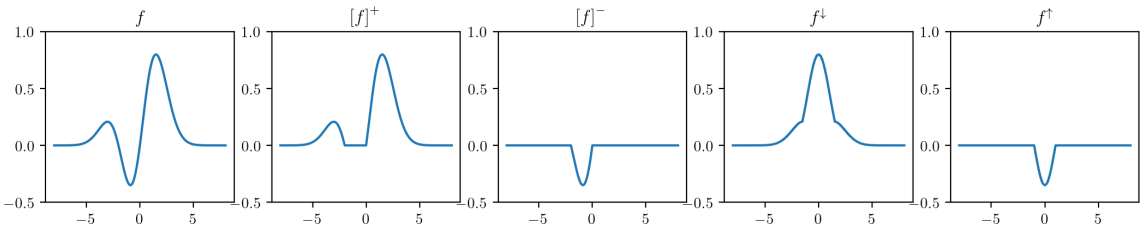

Figure 3.2: Example of infinite-domain distribution $f : \mathbb{R} \mapsto \mathbb{R}$. The distribution f is defined over $\mathbb{R}$ and takes both positive and negative values. We display from left to right: the original distribution f, its positive parts $[f]^+$ and its negative parts $[f]^-$., its decreasing rearrangement $f^\downarrow$, its increasing rearrangement $f^\uparrow$. From the illustration, it clearly appears that $f \not\equiv f^\downarrow$ and $f \not\equiv f^\uparrow$. However, we see that $f^\downarrow \equiv [f]^+$ and $f^\uparrow \equiv [f]^-$.

3.4.3 Rearrangements and cumulative integrals

Let us mention here some useful volume function of infinite sets. As they will be used in the present thesis, let us mention the volume function of $\mathbb{R}^n$ and $\mathbb{R}_+$.

$$V_{\mathbb{R}^n}(s) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} s^n \tag{3.150}$$

It can be seen that the volume function of $\mathbb{R}^n_+$ is simply obtained from the volume function of $\mathbb{R}^n$ as $V_{\mathbb{R}^n_+}(s) = V_{\mathbb{R}^n}(s)/2^n$. The volume function of $\mathbb{R}_+$ is simply $V_{\mathbb{R}_+}(s) = s$.

The cumulative integrals of a continuous distribution defined over an infinite domain are defined identically to the finite-domain case (see Definition 3.13). When the parameter s is chosen to be infinite, the cumulative integrals $S_s(f)$, $S^\downarrow_s(f)$ and $S^\uparrow_s(f)$ become respectively:

$$S_\infty(f) = \int f(\mathbf{r})\mathrm{d}\mathbf{r}, \qquad S^\downarrow_\infty(f) = \int [f(\mathbf{r})]^+ \,\mathrm{d}\mathbf{r}, \qquad S^\uparrow_\infty(f) = \int [f(\mathbf{r})]^- \,\mathrm{d}\mathbf{r}. \tag{3.151}$$

3.4.4 Relations of majorization

In a similar fashion as we did for infinite-dimensional vectors, we are going to build a version of continuous majorization which is compatible with infinite-domain functions taking both positive and negative values. In order to do so, we start from the hybrid formulations we introduced in the finite-domain case (see Section 3.2). This will leads us to the following extended version of continuous majorization.

Weak-majorization from below (sub-majorization)

From the hybrid conditions (3.95) and (3.96) that we have derived earlier, we define sub-majorization for infinite-domain functions as follows:

Definition 3.23 (Sub-majorization for infinite-domain distributions)**.** *The distribution $f : \mathbb{A} \mapsto \mathbb{R}$ weak-majorizes from below (sub-majorizes) the distribution $g : \mathbb{A} \mapsto \mathbb{R}$, written $f \succ_\mathrm{w} g$, is equivalent*

to any of these three conditions:

- $$\int_{\mathbb{A}} \varphi(f(\mathbf{r}))\mathrm{d}\mathbf{r} \geq \int_{\mathbb{A}} \varphi(f(\mathbf{r}))\mathrm{d}\mathbf{r} \qquad \forall\varphi \quad \textit{convex increasing} \tag{3.152}$$

- $$\begin{cases} \Phi_t^+(f) \geq \Phi_t^+(g) & \forall t \geq 0 \\ \Phi_t^-(f) - S_\infty(f) \leq \Phi_t^-(g) - S_\infty(g) & \forall t \leq 0 \end{cases} \tag{3.153}$$

- $$\begin{cases} S_s^\downarrow(f) \geq S_s^\downarrow(f) & \forall s \geq 0 \\ S_s^\uparrow(f) - S_\infty(f) \leq S_s^\uparrow(g) - S_\infty(g) & \forall s \geq 0 \end{cases} \tag{3.154}$$

And these three conditions are equivalent to each other.

Weak-majorization from above (super-majorization)

From the hybrid conditions (3.102) and (3.103) that we have derived earlier, we define super-majorization for infinite-domain functions as follows:

Definition 3.24 (Super-majorization for infinite-domain distributions)**.** *The distribution $f : \mathbb{A} \mapsto \mathbb{R}$ weak-majorizes from above (super-majorizes) the distribution $g : \mathbb{A} \mapsto \mathbb{R}$, written $f \succ^{\mathrm{w}} g$, is equivalent to any of these three conditions:*

- $$\int_{\mathbb{A}} \varphi(f(\mathbf{r}))\mathrm{d}\mathbf{r} \geq \int_{\mathbb{A}} \varphi(f(\mathbf{r}))\mathrm{d}\mathbf{r} \qquad \forall\varphi \quad \textit{convex decreasing} \tag{3.155}$$

- $$\begin{cases} \Phi_t^+(f) - S_\infty(f) \geq \Phi_t^+(g) - S_\infty(g) & \forall t \geq 0 \\ \Phi_t^-(f) \leq \Phi_t^-(g) & \forall t \leq 0 \end{cases} \tag{3.156}$$

- $$\begin{cases} S_s^\downarrow(f) - S_\infty(f) \geq S_s^\downarrow(f) - S_\infty(g) & \forall s \geq 0 \\ S_s^\uparrow(f) \leq S_s^\uparrow(g) & \forall s \geq 0 \end{cases} \tag{3.157}$$

And these three conditions are equivalent to each other.

Majorization

Definition 3.25 (Majorization for infinite-domain distributions)**.** *The distribution $f : \mathbb{A} \mapsto \mathbb{R}$ majorizes the distribution $g : \mathbb{A} \mapsto \mathbb{R}$, written $f \succ g$, is equivalent to any of these three conditions*

- $$\int_{\mathbb{A}} \varphi(f(\mathbf{r}))\mathrm{d}\mathbf{r} \geq \int_{\mathbb{A}} \varphi(g(\mathbf{r}))\mathrm{d}\mathbf{r} \qquad \forall\varphi \quad \textit{convex} \tag{3.158}$$

- $$f \succ_{\mathrm{w}} g \qquad \textit{and} \qquad \int_{\mathbb{A}} f(\mathbf{r})\mathrm{d}\mathbf{r} = \int_{\mathbb{A}} g(\mathbf{r})\mathrm{d}\mathbf{r} \tag{3.159}$$

- $$f \succ^{\mathrm{w}} g \qquad \textit{and} \qquad \int_{\mathbb{A}} f(\mathbf{r})\mathrm{d}\mathbf{r} = \int_{\mathbb{A}} g(\mathbf{r})\mathrm{d}\mathbf{r} \tag{3.160}$$

And these three conditions are equivalent to each other.

In practice, we will often use the following conditions to check whether two infinite-dimensional vectors majorize each other. The following condition comes from (3.107):

$$f \succ g \qquad \Leftrightarrow \qquad \begin{cases} S_\infty(f) = S_\infty(g) & \\ \Phi_t^+(f) \geq \Phi_t^+(g) & \forall t \geq 0 \\ \Phi_t^-(f) \leq \Phi_t^-(g) & \forall t \leq 0 \end{cases} \tag{3.161}$$

The following condition comes from (3.108):

$$f \succ g \qquad \Leftrightarrow \qquad \begin{cases} S_\infty(f) = S_\infty(g) & \\ S_s^\downarrow(f) \geq S_s^\downarrow(g) & \forall s \geq 0 \\ S_s^\uparrow(f) \leq S_s^\uparrow(g) & \forall s \geq 0 \end{cases} \tag{3.162}$$

Let us mention that the formulation of majorization that we have developed in the present section applies to infinite-domain functions, but it is also valid for finite-domain functions. As such, it constitutes a generalization of Section 3.2.

Chapter 4

Quantum optics in Wigner space

In Chapter 1 we have introduced the basics of quantum phase space in a very general setting. In that context, we have introduced the Wigner function, which is a central tool in our thesis. In this chapter, we will give a more precise physical framework to our considerations. In this sense, we are going to look in detail at quantum optics with continuous variables, and focus on one or several modes of the electromagnetic field. In a first step, we will define precisely the description of the physical background on which we work. Then, we will introduce several quantum state bases. Following this, we will present a number of optical transformations acting on quantum states, and we will explain precisely how these act on the Wigner functions of the states. Finally, we will conclude this chapter by introducing the broad family of Gaussian states.

4.1 Physical background

Here we follow mostly reference [45]. Quantum electrodynamics promote classical fields to quantum operators. The usual potential vector becomes then a quantum vectorial operator $\hat{\mathbf{A}} = (\hat{A}_x, \hat{A}_y, \hat{A}_z)$ according to the correspondence rule. The electric field vectorial operator $\hat{\mathbf{E}}$ and the magnetic field vectorial operator $\hat{\mathbf{B}}$ can be derived from the potential vector as follows:

$$\hat{\mathbf{E}} = -\frac{\partial \hat{\mathbf{A}}}{\partial t}, \qquad \hat{\mathbf{B}} = \boldsymbol{\nabla} \times \hat{\mathbf{A}}. \tag{4.1}$$

From Maxwell's equations, it appears that the potential vector is solution of the wave equation:

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right)\hat{\mathbf{A}} = 0, \tag{4.2}$$

where c denotes the speed of light in vacuum. It is then natural to develop the potential vector $\hat{\mathbf{A}}$ onto the set of plane waves. We define the propagation vector $\mathbf{k} \in \mathbb{R}^3$ which is related to an angular pulsation $\omega_k = \|\mathbf{k}\|$ and two orthogonal directions of polarization $\mathbf{e}_{\mathbf{k}\lambda}$ labeled by $\lambda \in \{1, 2\}$, such that the three vectors $\left\{\vec{\mathbf{1}}_{\mathbf{k}}, \vec{\mathbf{e}}_{\mathbf{k}1}, \vec{\mathbf{e}}_{\mathbf{k}2}\right\}$ form an oriented orthogonal basis. We consider a large box of volume V. The potential vector operator at some point $\mathbf{r}$ and time t is then given as follows:

$$\hat{\mathbf{A}}(\mathbf{r}, t) = \sum_{\mathbf{k}\lambda} \sqrt{\frac{\hbar}{2\varepsilon_0 V \omega_k}} \left(\hat{a}_{\mathbf{k}\lambda} \exp\left(-i\omega_k t + i\mathbf{kr}\right) + \hat{a}^\dagger_{\mathbf{k}\lambda} \exp\left(i\omega_k t - i\mathbf{kr}\right)\right) \vec{\mathbf{e}}_{\mathbf{k}\lambda}, \tag{4.3}$$

$$= \sum_{\mathbf{k}\lambda} \sqrt{\frac{1}{\varepsilon_0 V \omega_k}} \left(\hat{x}_{\mathbf{k}\lambda} \cos\left(\omega_k t - \mathbf{kr}\right) + \hat{p}_{\mathbf{k}\lambda} \sin\left(\omega_k t - \mathbf{kr}\right)\right) \vec{\mathbf{e}}_{\mathbf{k}\lambda}, \tag{4.4}$$

where ε_0 is the vacuum permittivity constant, related to the vacuum permeability μ_0 and speed of light c as $\varepsilon_0 \mu_0 c^2 = 1$. The operator $\hat{a}_{\mathbf{k}\lambda}$ is the quantum mode operator associated to the mode of propagation $(\mathbf{k}, \lambda)$. It is the quantum analog to the complex amplitude and is dimensionless. The operator $\hat{a}_{\mathbf{k}\lambda}$ has complex eigenvalues and is thus not an observable. On the contrary, the quadratures

operators $\hat{x}_{\mathbf{k}\lambda}$ and $\hat{p}_{\mathbf{k}\lambda}$ have a real spectrum and corresponds to observables. The mode operator and quadrature operators are related to each other, as we are going to show. In what follows, we will label the modes by a single index k, so that each value of k accounts for a different value of $(\mathbf{k}, \lambda)$.

$$\hat{x}_k = \sqrt{\frac{\hbar}{2}}\left(\hat{a}_k + \hat{a}_k^\dagger\right) \qquad \hat{p}_k = -i\sqrt{\frac{\hbar}{2}}\left(\hat{a}_k - \hat{a}_k^\dagger\right) \tag{4.5}$$

$$\hat{a}_k = \sqrt{\frac{1}{2\hbar}}\left(\hat{x}_k + i\hat{p}_k\right) \qquad \hat{a}_k^\dagger = \sqrt{\frac{1}{2\hbar}}\left(\hat{x}_k - i\hat{p}_k\right) \tag{4.6}$$

When considering the commutation inside a single mode of the elctromagnetic field (for k fixed), the commutation relation are the following: The mode operator of one mode obey the bosonic commutation relation $\left[\hat{a}, \hat{a}^\dagger\right] = 1$, and the quadrature operators obey the canonical commutation relation $[\hat{x}, \hat{p}] = i\hbar$. Two mode operators or quadratures operators associated to orthogonal modes commute together. When considering a set of N orthogonal modes of electromagnetic waves, it is possible to express the commutation relations in an simple formulation. To that purpose, we define the mode vector $\hat{\mathbf{b}}$ and the quadrature vector $\hat{\mathbf{q}}$ as vectors of $2N$ operators defined as:

$$\hat{\mathbf{b}} = \left(\hat{a}_1, \hat{a}_1^\dagger, \hat{a}_2, \hat{a}_2^\dagger ..., \hat{a}_N, \hat{a}_N^\dagger\right)^\intercal, \qquad \hat{\mathbf{q}} = \left(\hat{x}_1, \hat{p}_1, \hat{x}_2, \hat{p}_2, ..., \hat{x}_N, \hat{p}_N\right)^\intercal. \tag{4.7}$$

Note that we also define the reduced mode operator $\hat{\mathbf{a}}$ with N components as $\hat{\mathbf{a}} = (\hat{a}_1, \hat{a}_2, ..., \hat{a}_N)^\intercal$. Note that in alternative convention, the quadrature vector $\hat{\mathbf{q}}$ is ordered in a different way. We now introduce the symplectic form $\mathbf{\Omega}$ as the $2N \times 2N$ matrix defined as follows:

$$\mathbf{\Omega} = \bigoplus_{n=1}^{N} \boldsymbol{\omega} = \begin{pmatrix} \boldsymbol{\omega} & & \\ & \ddots & \\ & & \boldsymbol{\omega} \end{pmatrix}, \qquad \boldsymbol{\omega} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{4.8}$$

Note that the inverse of the symplectic form is $\mathbf{\Omega}^{-1} = \mathbf{\Omega}^\intercal = -\mathbf{\Omega}$. Then, the multimode bosonic commutation relations read as $\left[\hat{b}_i, \hat{b}_j\right] = \Omega_{ij}$ and the multimode canonical commutation relations read $[\hat{q}_i, \hat{q}_j] = i\hbar\Omega_{ij}$.

The Hamiltonian of a quantum electromagnetic wave corresponds to the sum of the energy stored in the electric field and the magnetic field. It can also be formulated from the mode operators and the quadrature operators, as follows:

$$\hat{H} = \frac{1}{2}\int_V \left(\varepsilon_0 \hat{\mathbf{E}}^2 + \frac{1}{\mu_0}\hat{\mathbf{B}}^2\right) \mathrm{d}\mathbf{r} \tag{4.9}$$

$$= \frac{1}{2}\sum_k \omega_k \left(\hat{x}_k^2 + \hat{p}_k^2\right) \tag{4.10}$$

$$= \sum_k \hbar\omega_k \left(\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2}\right) \tag{4.11}$$

We will refer to the above Hamiltonian as the bosonic Hamiltonian. In the following of this thesis we will in general omit the units. This corresponds to considering that we work in a system of units such that $\hbar = 1$. Also, we when dealing with multimode systems, we will only consider systems that have the same pulsation ω, and we will choose units such that the value of $\omega = 1$.

4.2 Essential optical pure states

Now that we have introduced the physical background, let us define some pure that are ubiquitous in quantum optics. The states we present here constitute different bases of quantum states upon which any state can be decomposed. We present single-mode states, but multi-mode states can be built by tensoring single-mode state. To that purpose, we consider a single mode of harmonic oscillator associated with a mode operator $\hat{a}$ and two quadrature operators $\hat{x}$ and $\hat{p}$.

4.2.1 Quadrature basis

A quadrature state is an eigenstate of one of the two quadratures $\hat{x}$ or $\hat{p}$.

$$\hat{x}\,|x\rangle = x\,|x\rangle \qquad x \in \mathbb{R}, \qquad\qquad \hat{p}\,|p\rangle = p\,|p\rangle \qquad p \in \mathbb{R}. \tag{4.12}$$

Quadrature states are not physical, as their wave-functions cannot be normalized. However, they are interesting as they form the most natural continuous basis to work in phase space. Indeed, they form a complete basis and are orthonormal:

$$\int |x\rangle\langle x|\,\mathrm{d}x = \hat{\mathbb{1}} \qquad\qquad \langle x_2|x_1\rangle = \delta(x_1 - x_2) \tag{4.13}$$

$$\int |p\rangle\langle p|\,\mathrm{d}p = \hat{\mathbb{1}} \qquad\qquad \langle p_2|p_1\rangle = \delta(p_1 - p_2) \tag{4.14}$$

It is possible to go from one to another as they are related by a Fourier transform. Indeed, we have the relation $\langle x|p\rangle = \exp(ixp/\hbar)/\sqrt{2\pi\hbar}$. As any single-mode basis, the quadrature basis can be tensored over several mode to construct a mulitmode basis. This leads us to define the vectorial quadrature state as follows:

$$|\mathbf{x}\rangle = \bigotimes_{i=1}^{N} |x_i\rangle = |x_1\rangle \otimes |x_2\rangle \otimes ... \otimes |x_N\rangle \tag{4.15}$$

where $\mathbf{x} \in \mathbb{R}^N$ is such that $\mathbf{x} = (x_1, ..., x_N)$.

4.2.2 Fock basis

The Fock basis plays a central role in quantum optics. Indeed, they are eigenstates of the Hamiltonian of the electromagnetic field. We introduce the photon-number operator as $\hat{n} = \hat{a}^\dagger\hat{a}$. From its construction, one can see that $\hat{n}$ is Hermitian. Moreover, it can be shown that its spectrum is discrete, integer and non-negative. The eigenstates of the photon-number operator $\hat{n}$ are called Fock states and are associated to the letter n, represented as $|n\rangle$:

$$\hat{n}\,|n\rangle = n\,|n\rangle \qquad\qquad n \in \mathbb{N}_0 \tag{4.16}$$

A Fock state expressed in the quadrature basis $\hat{x}$ admits the wave-function $\psi_n(x) = \langle x|n\rangle$ and Wigner function $W_n(x,p)$:

$$\psi_n(x) = \left(\sqrt{\pi}2^n n!\right)^{-\frac{1}{2}} H_n(x) \exp\left(-\frac{x^2}{2}\right) \tag{4.17}$$

$$W_n(x,p) = \frac{(-1)^n}{\pi} L_n\left(2x^2 + 2p^2\right) \exp\left(-x^2 - p^2\right) \tag{4.18}$$

Notice that the Wigner function of Fock state is radial, so that we will sometimes refer to it as $W_n(r)$. Throughout this thesis, we will refer to these Wigner function as W_n. So that W_0 is the Wigner function of the Fock state with eigenvalue 0, W_1 the Wigner function of the Fock state with eigenvalue 1, and so on. The Wigner function of the Fock state with eigenvalue 0 plays a prominent role in quantum optics, and in quantum physics in general. Indeed, it is the ground state of the bosonic Hamiltonian, which is usually called *vacuum*. The Wigner function of vacuum is W_0 and reads as:

$$W_0(x,p) = \frac{1}{\pi} \exp\left(-x^2 - p^2\right) \tag{4.19}$$

We plot in Figure 4.1 the Wigner function of the Fock states assocaited to $n = 0, 1, 2, 3$. In Figure 4.2 we plot the Wigner function of the Fock state associated to $n = 10$ together with its marginal distributions.

The Fock states form a complete orthonormal basis, meaning that the following relation holds

$$\sum_{n=0}^{\infty} |n\rangle\langle n| = \hat{\mathbb{1}} \qquad\qquad \langle i|j\rangle = \delta_{ij}, \tag{4.20}$$

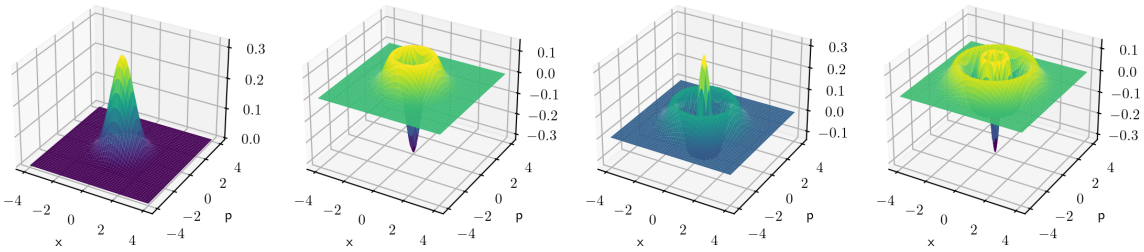

Figure 4.1: Wigner functions of the Fock states assocociated to $n = 0$, 1, 2 and 3, from left to right. The Wigner function of a Fock state is radial and only depends on the parameter $r = \sqrt{x^2 + p^2}$. Observe that it is in general partly negative, except in the case $n = 0$. At $(x, p) = (0, 0)$, the value of the Wigner function of n^{th} Fock state is equal to $(-1)^n/\pi$.

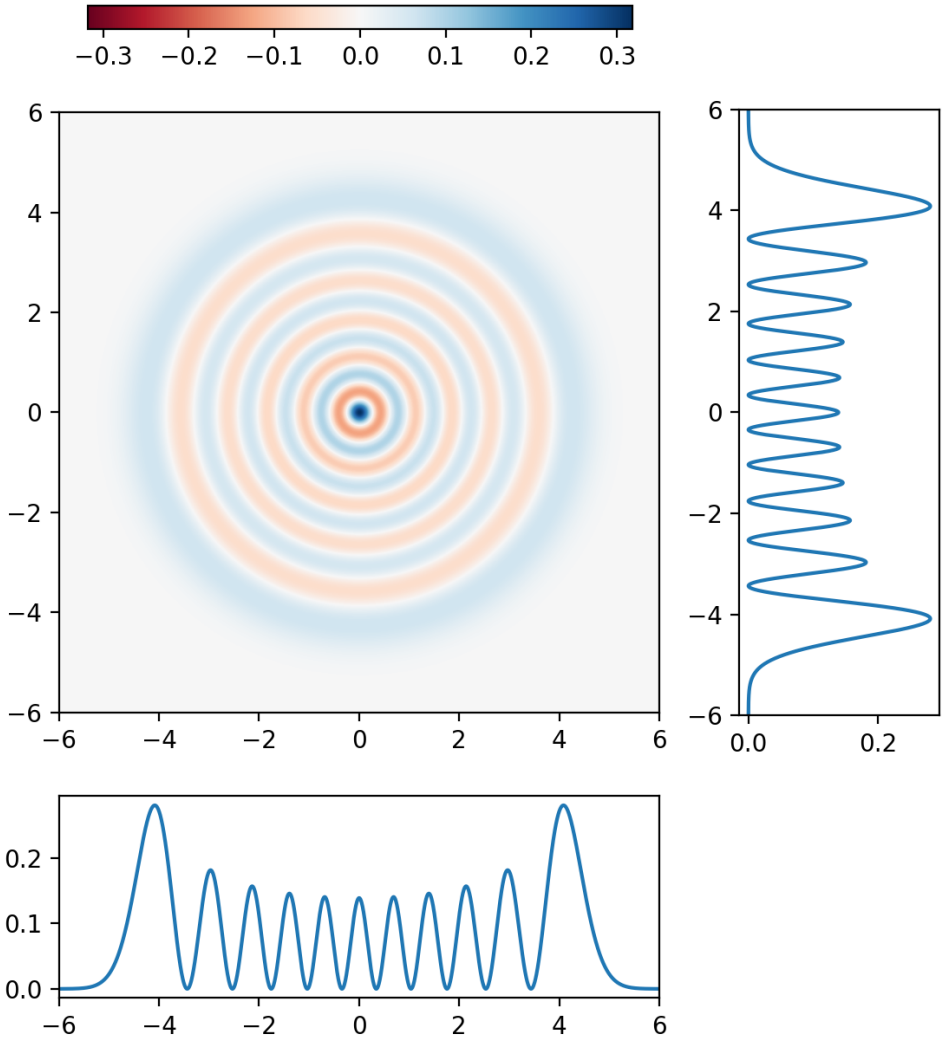

Figure 4.2: Wigner function of the Fock state $n = 10$, together with its marginal distributions $\rho_x(x) = \int W_n(x, p)\mathrm{d}p$ and $\rho_p(p) = \int W_n(x, p)\mathrm{d}x$. Since Fock states are invariant by rotation, their marginal distributions are identical for each quadrature. Observe that despite the fact that the Wigner function is partly negative, the marginal distributions are non-negative, as expected.

where δ_{ij} is the Kronecker delta such that $\delta_{ij} = 1$ if $i = j$ and is zero otherwise. The Fock basis is the most natural discrete basis to work in state space. The fact that it is a discrete basis is very useful in order to express a quantum state as a matrix. Indeed, it is then possible to write any single-mode state $\hat{\rho}$ as a matrix ρ such that

$$(\rho)_{ij} = \rho_{ij} = \langle i| \hat{\rho} |j\rangle , \qquad \hat{\rho} = \sum_{ij} \rho_{ij} |i\rangle \langle j| . \tag{4.21}$$

where the summation is performed over both indices i and j from 0 to ∞. We use a shorter notation so that $\sum_{ij} = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty}$. A direct consequence of that observation is that the Wigner function of any quantum state can be expressed from the Wigner-Weyl transform of the element $|i\rangle \langle j|$ in the Wigner basis. Indeed, we have the following relation:

$$W_{\hat{\rho}} = \sum_{ij} \rho_{ij} W_{ij}, \qquad W_{ij} = \mathcal{T}_{\mathrm{w}} \left[|i\rangle \langle j| \right] . \tag{4.22}$$

We understand then that the expression of the elements W_{ij}, corresponding to the operators $|i\rangle \langle j|$ have a particular importance in order to switch from a matrix representation to a phase-space representation. When $i \geq j$, the expression of W_{ij} can be found in [38] and is the following:

$$W_{ij}(x,p) = \frac{(-1)^i}{\pi} \left(\sqrt{2}(x+ip) \right)^{i-j} L_j^{(i-j)} \left(2x^2 + 2p^2 \right) \exp\left(-x^2 - p^2 \right) \qquad (i \geq j) \tag{4.23}$$

where we use the notation $L_n^{(\alpha)}$ to denote the generalized Laguerre polynomial. When $i < j$, the expression of W_{ij} can be found easily using the identity $W_{ji} = W_{ij}^*$. It is then expressed as follows:

$$W_{ij}(x,p) = \frac{(-1)^j}{\pi} \left(\sqrt{2}(x-ip) \right)^{j-i} L_i^{(j-i)} \left(2x^2 + 2p^2 \right) \exp\left(-x^2 - p^2 \right) \qquad (i \leq j) \tag{4.24}$$

Notice that the case $i = j$ resumes to (4.18), as we expect.

We define a multimode Fock state from a vector $\mathbf{n} \in \mathbb{N}_0^N$ as follows:

$$|\mathbf{n}\rangle = \bigotimes_{i=1}^{N} |n_i\rangle \tag{4.25}$$

where $\mathbf{n} = (n_1, ..., n_N)$.

4.2.3 Coherent basis

Coherent states are eigenstates of the mode operator $\hat{a}$.

$$\hat{a} |\alpha\rangle = \alpha |\alpha\rangle \qquad \alpha \in \mathbb{C} \tag{4.26}$$

In practice, we will associates the complex parameter α to two real parameters x_α and p_α which are such that $\alpha = (x_\alpha + ip_\alpha)/\sqrt{2}$. These two parameters will be useful to express several relations in a simpler and more intuitive form. Note the importance of the factor $\sqrt{2}$. The parameter α is the eigenvalue of the coherent state $|\alpha\rangle$ with respect to the operator $\hat{a}$. The parameters x_α and p_α are related to the phase space localization of the coherent state, as we are going to see. The wave-function of a coherent state in the quadrature basis is $\psi_\alpha(x) = \langle x|\alpha\rangle$, defined as follows:

$$\psi_\alpha(x) = \pi^{-\frac{1}{4}} \exp\left(\frac{1}{2} \left(\alpha^2 - |\alpha|^2 \right) \right) \exp\left(-\frac{1}{2} \left(x - \sqrt{2}\alpha \right)^2 \right) \tag{4.27}$$

$$= \pi^{-\frac{1}{4}} \exp\left(-\frac{1}{2}(x - x_\alpha)^2 \right) \exp\left(ip_\alpha \Big(x - x_\alpha/2 \Big) \right) \tag{4.28}$$

The Wigner function of a coherent state is analog to the one of vacuum, as it is simply the Wigner function W_0 shifted in phase space according to the parameters (x_α, p_α). Indeed, the Wigner function of a coherent state is $W_\alpha(x,p) = W_0(x - x_\alpha, p - p_\alpha)$.

Coherent states form an over-complete basis, which means that any state can be decomposed onto the set of coherent state, even if two coherent states are never orthogonal. We have the following relations:

$$\frac{1}{\pi}\iint |\alpha\rangle\langle\alpha| \,\mathrm{d}^2\alpha = \hat{\mathbb{1}}, \qquad\qquad \langle\beta|\alpha\rangle = \exp\left(-\frac{1}{2}\left(|\alpha|^2 + |\beta|^2 - 2\alpha\beta^*\right)\right). \tag{4.29}$$

Since coherent states form a complete basis, any quantum operator $\hat{A}$ can be decomposed unequivocally onto the elements $A(\alpha,\beta) = \langle\alpha|\,\hat{A}\,|\beta\rangle$. Then, the Weyl transform of the operator $\hat{A}$ can be reconstructed by using the Weyl transform of $|\alpha\rangle\langle\beta|$, which we denote as $W_{\alpha\beta} = \mathcal{T}_{\mathrm{w}}\left[|\alpha\rangle\langle\beta|\right]$. The expression of $W_{\alpha\beta}$ can be computed as the following:

$$W_{\alpha\beta}(x,p) = \frac{1}{\pi}\exp\left(-\frac{1}{2}\left(|\alpha|^2 + |\beta|^2 - 2\alpha\beta^*\right) - \left(x - \frac{\alpha+\beta^*}{\sqrt{2}}\right)^2 - \left(p - \frac{\alpha-\beta^*}{\sqrt{2}i}\right)^2\right). \tag{4.30}$$

Then, it is possible to express $A(x,p)$ from the elements $A(\alpha,\beta)$ as follows:

$$A(x,p) = \frac{1}{\pi^2}\iint A(\alpha,\beta)W_{\alpha\beta}(x,p)\mathrm{d}^2\alpha\mathrm{d}^2\beta. \tag{4.31}$$

A multimode coherent state is defined by a vector of complex eigenvalues $\boldsymbol{\alpha} \in \mathbb{C}^N$, as follows:

$$|\boldsymbol{\alpha}\rangle = \bigotimes_{i=1}^{N} |\alpha_i\rangle \tag{4.32}$$

where $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)$.

It should be noted that coherent states are part of a broader set of quantum states known as Gaussian states. We will come to these states later in the present chapter.

4.3 Linear transformations

This section mostly follows references [52] and [66]. In quantum mechanics, the evolution of quantum states over time is described by unitary operators. Such an evolution is always reversible, as unitary operators verify the relation $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{\mathbb{1}}$. As a quantum state evolves, the expectation values it takes over observables change. There exist two different interpretation to that observation. Either we consider that the unitary acts on the quantum state, either we consider that the unitary acts on the observable. The former is the Schrödinger picture, the latter is the Heisenberg picture. We can compare these two interpretation in the following table.

	Schrödinger picture	Heisenberg picture
$\hat{\rho}'$	$\hat{U}\hat{\rho}\hat{U}^\dagger$	$\hat{\rho}$
$\hat{A}'$	$\hat{A}$	$\hat{U}^\dagger\hat{A}\hat{U}$
$\langle\hat{A}'\rangle_{\hat{\rho}'}$	$\mathrm{Tr}\left[\hat{A}\hat{\rho}'\right]$	$\mathrm{Tr}\left[\hat{A}'\hat{\rho}\right]$

The state $\hat{\rho}'$ corresponds to state $\hat{\rho}$ after the action of the unitary $\hat{U}$, and the observable $\hat{A}'$ corresponds to observable $\hat{A}$ after the action of the unitary $\hat{U}$. We see that even if $\hat{\rho}'$ and $\hat{A}'$ are different whether we are in the Schrödinger or Heisenberg picture, the expectation value $\langle\hat{A}'\rangle_{\hat{\rho}'}$ is equal in both interpretations.

In this section, we are going to present a set of quantum operators that are associated to linear phase-space transformations. These transformation constitute the building blocks to work in quantum optics. Note that the transformations that we are going to present do not always correspond to physically implementable transformations. However, they are interesting as they appear in other more complex physical transformations, as we are going to see. First let us define what we call a linear transformation.

Definition 4.1 (Linear transformation)**.** *A linear mode transformation (or Bogoliubov transformation) is an affine transformation that maps the quadrature operators onto a linear combination of themselves. It is associated to a quantum operator $\hat{U}$, an (invertible) matrix $\mathbf{U} \in \mathbb{R}^{2N\times 2N}$ and a vector $\mathbf{u} \in \mathbb{R}^{2N}$ which act on the quadrature opertaor as:*

$$\hat{U}^\dagger \hat{\mathbf{q}} \hat{U} \;=\; \mathbf{U}\hat{\mathbf{q}} + \mathbf{u} \tag{4.33}$$

We also associate it with a phase-space transformation $\mathcal{U}$ acting on the Wigner function as follows:

$$\mathcal{U}\left[W\right](\mathbf{q}) \;=\; \frac{1}{|\det \mathbf{U}|}\, W\left(\mathbf{U}^{-1}\left(\mathbf{q}-\mathbf{u}\right)\right) \tag{4.34}$$

The quantum operators of linear transformations are often associated to Gaussian unitaries, which are described by a unitary of the form $\exp\left(i\hat{H}/\hbar\right)$ where $\hat{H}$ is a second-order polynomial in the field operators. However, as we are going to see, our definition encompasses also operators that are not unitary.

A linear transformation respects the canonical commutation rules when the matrix $\mathbf{U}$ preserves the symplectic form, so that $\mathbf{U}\boldsymbol{\Omega}\mathbf{U}^\intercal = \boldsymbol{\Omega}$, and in such case, $\mathbf{U}$ is said to be *symplectic*. The inverse of a symplectic matrix can be computed as $\mathbf{U}^{-1} = -\boldsymbol{\Omega}\mathbf{U}^\intercal\boldsymbol{\Omega}$. The action of a linear transformation over the mode operators is described by the Bogoliubov transformation $\hat{U}\hat{\mathbf{a}}\hat{U}^\dagger = \mathbf{A}\hat{\mathbf{a}} + \mathbf{B}\hat{\mathbf{a}}^\dagger + \boldsymbol{\alpha}$ where $\mathbf{A},\mathbf{B} \in \mathbb{C}^{N\times N}$ and $\boldsymbol{\alpha} \in \mathbb{C}^N$. Note that in order to ensure that the bosonic commutations relations are preserved, we need the additional constraints that $\mathbf{A}\mathbf{B}^\intercal = \mathbf{B}\mathbf{A}^\intercal$ and $\mathbf{A}\mathbf{A}^\dagger = \mathbf{B}\mathbf{B}^\dagger + \mathbf{I}$. Let us also specify that any choice of $\mathbf{A},\mathbf{B},\boldsymbol{\alpha}$ defines unequivocally $\mathbf{U}$ and $\mathbf{u}$ and *vice versa*.

The point of view we have used until now is the Heisenberg picture. Within the Schrödinger picture, the evolution applies to the quantum state, and thus to its Wigner function. The value taken by the new Wigner function expressed in the new quadratures is related to the value of the old Wigner function in the old quadratures. Moreover, there is a normalization factor of $|\det \mathbf{U}|$ to ensure that the Wigner function remains normalized, so that we have the relation $W'(\mathbf{q}') = W(\mathbf{q})/|\det \mathbf{U}|$. When working in phase space, we are going to associate to each linear transformation operator $\hat{U}$ a phase space transformation $\mathcal{U}$, as presented in Equation (4.34).

4.3.1 Displacement

The displacement operator acting on mode $\hat{a}$ depends on the complex-valued parameter α. It is defined as follows:

$$\hat{D}_\alpha = \exp\left(\alpha\hat{a}^\dagger - \alpha^*\hat{a}\right) \tag{4.35}$$

In Heisenberg picture, the displacement operator act on the mode operator as $\hat{a} \to \hat{a} + \alpha$. As its name indicate, the displacement operator is an operator which has the effect of moving the Wigner function over phase space. When it comes to the quadrature operators, they evolves as $\hat{x} \to \hat{x} + \sqrt{2}\,\mathrm{Re}\,\alpha$ and $\hat{p} \to \hat{p} + \sqrt{2}\,\mathrm{Im}\,\alpha$. So that we can identify the symplectic matrix $\mathbf{D} = \mathbf{I}_2$ and the displacement vector $\mathbf{d}_\alpha = (\sqrt{2}\,\mathrm{Re}\,\alpha, \sqrt{2}\,\mathrm{Im}\,\alpha)^\intercal$:

$$\mathbf{d}_\alpha = \begin{pmatrix} \sqrt{2}\,\mathrm{Re}\,\alpha \\ \sqrt{2}\,\mathrm{Im}\,\alpha \end{pmatrix}. \tag{4.36}$$

In Schrödinger picture, the displacement operator can be associated to the phase-space transformation $\mathcal{D}$ defined as follows:

$$\mathcal{D}_\alpha\left[W\right](x,p) = W(x - \sqrt{2}\,\mathrm{Re}\,\alpha, p - \sqrt{2}\,\mathrm{Im}\,\alpha) \tag{4.37}$$

For convenience, we also define an alternative writing of the displacement operator. We use the convention that when $\mathcal{D}$ is provided with a couple of reals rather than a complex number it acts as follows:

$$\mathcal{D}_{(x_\alpha,p_\alpha)}\left[W\right](x,p) = W(x-x_\alpha, p-p_\alpha) \tag{4.38}$$

In our convention, we have the relation $\mathcal{D}_{\sqrt{2}\alpha} = \mathcal{D}_{(\mathrm{Re}\,\alpha,\mathrm{Im}\,\alpha)}$, so that α and x_α, p_α are related as $\alpha = (x_\alpha + ip_\alpha)/\sqrt{2}$.

Note that the displacement operator is formally equal to $\hat{D}_\alpha = \exp\left(i\sqrt{2}(\mathrm{Im}\,\alpha\hat{x} - \mathrm{Re}\,\alpha\hat{p})\right)$. This can be expressed as $\hat{D}_\alpha = \exp\left(i\hat{\mathbf{q}}^\intercal\boldsymbol{\omega}\mathbf{d}_\alpha\right)$. That operator can be extended to a multimode operator as:

$$\hat{D}_{\boldsymbol{\alpha}} = \exp\left(i\hat{\mathbf{q}}^\intercal\boldsymbol{\Omega}\mathbf{d}_{\boldsymbol{\alpha}}\right) \tag{4.39}$$

This is sometimes called the Weyl operator in the literature. It simply corresponds to a displacement over several modes.

4.3.2 Rotation

The rotation operator is a single mode operator acting on a quantum mode $\hat{a}$. It depends on a real parameter θ and is defined as:

$$\hat{R}_\theta = \exp\left(-i\theta\hat{a}^\dagger\hat{a}\right) \tag{4.40}$$

In Heisenberg picture, it acts on the mode operator as $\hat{a} \to \exp(-i\theta)\hat{a}$. Its action on the quadratures are $\hat{x} \to \cos\theta\hat{x} + \sin\theta\hat{p}$ and $\hat{p} \to -\sin\theta\hat{x} + \cos\theta\hat{p}$. This leads us to define the symplectic matrix $\mathbf{R}_\theta$ as follows:

$$\mathbf{R}_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{4.41}$$

The matrix $\mathbf{R}_\theta$ is a matrix that acts on the quadrature vector as $\hat{\mathbf{q}}' = \mathbf{R}_\theta\hat{\mathbf{q}}$ and should not be mistaken with the unitary operator $\hat{R}_\theta$. Notice that $\mathbf{R}_\theta^{-1} = \mathbf{R}_\theta^\intercal = \mathbf{R}_{-\theta}$. In the Schrödinger picture, the phase-space transformation associated to the rotation operator is the transformation $\mathcal{R}_\theta$ that we define as follows:

$$\mathcal{R}_\theta\left[W\right](x,p) = W(x\cos\theta - p\sin\theta, x\sin\theta + p\cos\theta). \tag{4.42}$$

In Heisenbeg picture, the operator $\hat{R}_\theta$ makes the axis (x,p) perform a rotation of angle θ. In Schrödinger picture, the Wigner function is performing a rotation of angle $-\theta$. We now detail a few examples of rotations which are particularly common.

$$\mathcal{R}_0\left[W\right](x,p) = W(x,p) \tag{4.43}$$
$$\mathcal{R}_{\frac{\pi}{2}}\left[W\right](x,p) = W(p,-x) \tag{4.44}$$
$$\mathcal{R}_{\pi}\left[W\right](x,p) = W(-x,-p) \tag{4.45}$$
$$\mathcal{R}_{-\frac{\pi}{2}}\left[W\right](x,p) = W(-p,x) \tag{4.46}$$

The rotation operator $\hat{R}_\theta$ plays a central role in quantum optics, as it corresponds to the evolution operator. Indeed, the bosonic Hamiltonian is $\hat{n} + 1/2$ so that the unitary evolution is governed by the operator $\exp\left(-i\hat{H}t\right)$ which is related to the rotation operator $\hat{R}_\theta$. As a consequence, the Wigner function of a bosonic light field in free evolution rotates and Fock states, which are the eigenstates of $\hat{H}$, are invariant under rotation.

4.3.3 Squeezing

The squeezing operator acting on mode $\hat{a}$ is defined as follows:

$$\hat{S}_z = \exp\left(\frac{1}{2}(z\hat{a}^{\dagger 2} - z^*\hat{a}^2)\right) \tag{4.47}$$

where z is a complex parameter such that $z = r\exp(i\varphi)$. The resulting action of $\hat{S}_z$ on mode $\hat{a}$ is $\hat{a} \to \hat{a}\cosh r + \hat{a}^\dagger \exp(i\varphi)\sinh r$. The symplectic matrix associated to the unitary operator $\hat{S}_z$ can be expressed more easily introducing the matrix $\mathbf{Z}_\varphi$:

$$\mathbf{S}_z = \cosh r \mathbf{I}_2 + \sinh r \mathbf{Z}_\varphi, \qquad \mathbf{Z}_\varphi = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}. \tag{4.48}$$

The parameter z is in general complex. Without loss of generality, we can restrict ourselves to the case of a real squeezing parameter. Indeed, one can see that $\hat{S}_z = \hat{R}^\dagger_{\varphi/2}\hat{S}_r\hat{R}_{\varphi/2}$. For that reason, we define the symplectic matrix $\mathbf{S}_r$:

$$\mathbf{S}_r = \begin{pmatrix} \exp(r) & 0 \\ 0 & \exp(-r) \end{pmatrix} \tag{4.49}$$

The phase-space transformation associated to squeezing is the following transformation:

$$\mathcal{S}_r\left[W\right](x,p) = W\left(\exp(-r)x, \exp(r)p\right). \tag{4.50}$$

The squeezing phase-space operator $\mathcal{S}_r$ squeezes the Wigner function in the p direction for positive values of r, so that the Wigner function becomes stretched alongside the x-axis.

4.3.4 Conjugation

Let us now introduce a phase-space transformation that does not corresponds to a physically implementable operation [14]. We definte the conjugation operator as an operator that maps the mode operator to its dagger, so that its action is $\hat{a} \to \hat{a}^\dagger$. Notice that this operator does not conserve the bosonic commutation relations, as the new commutator becomes $\left[\hat{a}', \hat{a}'^\dagger\right] = \left[\hat{a}^\dagger, \hat{a}\right] = -1$. Its action on the quadrature is to change to sign of $\hat{p}$, so that we have $\hat{x} \to \hat{x}$ and $\hat{p} \to -\hat{p}$. It is associated to the following matrix $\mathbf{\Pi}$:

$$\mathbf{\Pi} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.51}$$

Notice that $\mathbf{\Pi} = \mathbf{Z}_0$. The conjugation phase-space transformation is Π is defined as follows:

$$\Pi\left[W\right](x,p) = W(x,-p) \tag{4.52}$$

Despite being non physically implementable, the conjugation operator as the property to apply a physically acceptable state onto another physically acceptable state. The conjugation operator will be helpful in the following of this chapter.

4.3.5 Rescaling

We now introduce another non-physical operator, that we call the rescaling operator. In Chapter 2, we introduced the rescaling operator $\mathcal{L}_s$ (see (2.25)). We define the same transformation for Wigner function as the following:

$$\mathcal{L}_s\left[W\right](x,p) = \frac{1}{s^2} W\left(\frac{x}{s}, \frac{p}{s}\right) \tag{4.53}$$

The matrix acting on the quadrature can be identified as the following:

$$\mathbf{L}_s = \begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix} = s\mathbf{I} \tag{4.54}$$

The action of the rescaling operator over the quadratures is $\hat{x} \to s\hat{x}$ and $\hat{p} \to s\hat{p}$. Its action over the mode operator is $\hat{a} \to s\hat{a}$. Formally, we may identify a quantum operator $\hat{L}_s = \sqrt{s}$, so that indeed $\hat{L}_s^\dagger \hat{a} \hat{L}_s = s\hat{a}$. Note that $\hat{L}_s$ does not respect the bosonic commutation relation, as we have $\left[\hat{a}', \hat{a}'^\dagger\right] = \left[s\hat{a}, s\hat{a}^\dagger\right] = s^2$. Note also that $\hat{L}_s$ is not unitary, as $\hat{L}_s^\dagger \hat{L}_s = s$.

Similarly to the conjugation operator, the rescaling operator is not physically realizable. However, on the contrary to the conjugation, the rescaling operator does not apply in general a physical state onto another physical state.

4.3.6 Two-mode mixing

Two-mode mixing is performed by a beam-splitter acting on two modes, and we will use two-mode mixing and beam-splitting as synonyms. For clarity, we will label the two modes as $\hat{a}$ and $\hat{b}$ (rather than $\hat{a}_1$ and $\hat{a}_2$). The mode $\hat{b}$ should not be confused with the mode vector $\hat{\mathbf{b}} = \left(\hat{a}, \hat{a}^\dagger, \hat{b}, \hat{b}^\dagger\right)$.

$$\hat{B}_\xi = \exp\left(\xi \hat{a}^\dagger \hat{b} - \xi^* \hat{a} \hat{b}^\dagger\right) \tag{4.55}$$

where $\xi \in \mathbb{C}$ is such that $\xi = \theta \exp(i\varphi)$ with $\theta, \varphi \in \mathbb{R}$. The beam-splitting operator $\hat{B}_\xi$ acts on the mode operators as follows:

$$\begin{pmatrix} \hat{a}' \\ \hat{b}' \end{pmatrix} = \begin{pmatrix} \cos\theta & \exp(i\varphi)\sin\theta \\ -\exp(-i\varphi)\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} \tag{4.56}$$

Its action on the quadrature operators is then given by the following matrix:

$$\mathbf{B}_\xi = \begin{pmatrix} \cos\theta \mathbf{I}_2 & \sin\theta \mathbf{R}_\varphi \\ -\sin\theta \mathbf{R}_\varphi^\intercal & \cos\theta \mathbf{I}_2 \end{pmatrix}, \tag{4.57}$$

The matrix $\mathbf{B}_\xi$ can be used to find the new quadrature operators using the relation $\hat{\mathbf{q}}' = \mathbf{B}_\xi \hat{\mathbf{q}}$. The matrix $\mathbf{B}_\xi$ acts on the quadratures vector $\hat{\mathbf{q}}$ and should not be mistaken with the unitary operator $\hat{B}_\xi$.

Note that the parameter ξ is in general complex. However, can restrict ourselves to the study of a real valued ξ, because the added phase can be considered as the addition of a rotation over some modes. For that reason, we will consider the two-mode squeezing operator $\mathbf{B}_\theta$ where θ is a real number. The matrix associated to it is the following:

$$\mathbf{B}_\theta = \begin{pmatrix} \cos\theta \mathbf{I}_2 & \sin\theta \mathbf{I}_2 \\ -\sin\theta \mathbf{I}_2 & \cos\theta \mathbf{I}_2 \end{pmatrix} \tag{4.58}$$

Note that in a shorter form we can write $\mathbf{B}_\theta = \mathbf{R}_\theta \otimes \mathbf{I}_2$. Note the interesting property that $\hat{B}_{\theta_1} \hat{B}_{\theta_2} = \hat{B}_{\theta_1 + \theta_2}$.

Let us mention that a beam-splitter is sometimes defined by its transmittance parameter η, which is related to the parameter θ by the relation $\eta = \cos^2\theta$. From that relation it obviously appears that η is between 0 and 1.

Reduced output of a beam-splitter

Let us consider that the input state of a beam-splitter is a product state. It is characterized by a Wigner function of the form $W(x_1, p_1, x_2, p_2) = W_1(x_1, p_1) W_2(x_2, p_2)$. The output $W'(x_1, p_1, x_2, p_2)$ is then computed as:

$$\begin{aligned} W'(x_1, p_1, x_2, p_2) &= W_1(\cos\theta x_1 - \sin\theta x_2, \cos\theta p_1 - \sin\theta p_2) \\ &\times W_2(\sin\theta x_1 + \cos\theta x_2, \sin\theta p_1 + \cos\theta p_2) \end{aligned} \tag{4.59}$$

Tracing over mode 2 yields the single-mode output on mode 1:

$$\begin{aligned} W_1'(x_1, p_1) = \iint \quad & W_1(\cos\theta x_1 - \sin\theta x_2, \cos\theta p_1 - \sin\theta p_2) \\ & W_2(\sin\theta x_1 + \cos\theta x_2, \sin\theta p_1 + \cos\theta p_2) \quad \mathrm{d}x_2 \mathrm{d}p_2 \end{aligned} \tag{4.60}$$

Introducing the following change of variable, this gives:

$$\begin{cases} x' = \cos\theta \left(\cos\theta x_1 - \sin\theta x_2\right) \\ p' = \cos\theta \left(\cos\theta p_1 - \sin\theta p_2\right) \end{cases} \qquad \begin{cases} \mathrm{d}x' = -\cos\theta \sin\theta \mathrm{d}x_2 \\ \mathrm{d}p' = -\cos\theta \sin\theta \mathrm{d}p_2 \end{cases} \tag{4.61}$$

We can then rewrite the above equation as:

$$W_1'(x_1, p_1) = \iint \frac{1}{\cos^2\theta} W_1\left(\frac{x'}{\cos\theta}, \frac{p'}{\cos\theta}\right) \frac{1}{\sin^2\theta} W_2\left(\frac{x_1 - x'}{\sin\theta}, \frac{p_1 - p'}{\sin\theta}\right) \mathrm{d}x_2 \mathrm{d}p_2 \tag{4.62}$$

In our convention, the expression can be written in a compact way as:

$$W_1' = \mathcal{L}_{\cos\theta}[W_1] * \mathcal{L}_{\sin\theta}[W_2] \tag{4.63}$$

The same development can be done for the other output.

$$W(x_2,p_2) = \iint \begin{array}{l} W_1(\cos\theta x_1 - \sin\theta x_2, \cos\theta p_1 - \sin\theta p_2) \\ W_2(\sin\theta x_1 + \cos\theta x_2, \sin\theta p_1 + \cos\theta p_2) \end{array} \mathrm{d}x_1\mathrm{d}p_1 \tag{4.64}$$

We introduce the following change of variables

$$\begin{cases} x' = \cos\theta\,(\sin\theta x_1 + \cos\theta x_2) \\ p' = \cos\theta\,(\sin\theta p_1 + \cos\theta p_2) \end{cases} \qquad \begin{cases} \mathrm{d}x' = \cos\theta\sin\theta\mathrm{d}x_1 \\ \mathrm{d}p' = \cos\theta\sin\theta\mathrm{d}p_1 \end{cases} \tag{4.65}$$

So that we can write

$$W(x_2,p_2) = \iint \frac{1}{\sin^2\theta} W_1\left(\frac{x'-x_2}{\sin\theta}, \frac{p'-p_2}{\sin\theta}\right) \frac{1}{\cos^2\theta} W_2\left(\frac{x'}{\cos\theta}, \frac{p'}{\cos\theta}\right) \mathrm{d}x_1\mathrm{d}p_1 \tag{4.66}$$

In our notations, this corresponds to the shorter writing $W_2' = \mathcal{R}_\pi \circ \mathcal{L}_{\sin\theta}[W_1] * \mathcal{L}_{\cos\theta}[W_2]$. Note that we use the notation $\mathcal{R} \circ \mathcal{L}[W] = \mathcal{R}[\mathcal{L}[W]]$. We summarize the relations we have derived for the beam-splitter in the following table.

	$\hat{\rho}'$	W'	$\hat{a}'$
mode 1	$\mathrm{Tr}_2\left[\hat{B}_\theta\,(\hat{\rho}_1 \otimes \hat{\rho}_2)\,\hat{B}_\theta^\dagger\right]$	$\mathcal{L}_{\cos\theta}[W_1] * \mathcal{L}_{\sin\theta}[W_2]$	$\cos\theta\ \hat{a}_1 + \sin\theta\ \hat{a}_2$
mode 2	$\mathrm{Tr}_1\left[\hat{B}_\theta\,(\hat{\rho}_1 \otimes \hat{\rho}_2)\,\hat{B}_\theta^\dagger\right]$	$\mathcal{R}_\pi \circ \mathcal{L}_{\sin\theta}[W_1] * \mathcal{L}_{\cos\theta}[W_2]$	$-\sin\theta\ \hat{a}_1 + \cos\theta\ \hat{a}_2$

4.3.7 Two-mode squeezing

Two-mode squeezing is performed by a two-mode squeezer and corresponds to the following unitary:

$$\hat{\Sigma}_z = \exp\left(z\hat{a}^\dagger\hat{b}^\dagger - z^*\hat{a}\hat{b}\right) \tag{4.67}$$

where $z \in \mathbb{C}$ is such that $r\exp(i\varphi)$ with $r, \varphi \in \mathbb{R}$. Its action on the mode operator $\hat{a}$ and $\hat{b}$ is the following:

$$\hat{a}' = \hat{a}\cosh r + \hat{b}^\dagger \exp(i\varphi)\sinh r \tag{4.68}$$

$$\hat{b}' = \hat{b}\cosh r + \hat{a}^\dagger \exp(-i\varphi)\sinh r \tag{4.69}$$

From these relations, we can construct the two following symplectic matrices:

$$\mathbf{\Sigma}_z = \begin{pmatrix} \cosh r\mathbf{I}_2 & \sinh r\mathbf{Z}_\varphi \\ \sinh r\mathbf{Z}_\varphi & \cosh r\mathbf{I}_2 \end{pmatrix}, \qquad \mathbf{\Sigma}_r = \begin{pmatrix} \cosh r\mathbf{I}_2 & \sinh r\mathbf{Z}_0 \\ \sinh r\mathbf{Z}_0 & \cosh r\mathbf{I}_2 \end{pmatrix}. \tag{4.70}$$

The matrix $\mathbf{\Sigma}_r$ corresponds to $\mathbf{\Sigma}_z$ when z is real and equal to $r \in \mathbb{R}$. The new quadrature operators as then computed as $\hat{\mathbf{q}}' = \mathbf{\Sigma}_z\hat{\mathbf{q}}$.

Let us mention that a two-mode squeezer is sometimes defined by its gain g or parameter λ. They are related to the squeezing r by the relation $g = \cosh^2 r$ and $\lambda = \tanh^2 r$. Notice that the gain g is always greater or equal to 1, and the parameter λ is between 0 and 1.

Reduced output of a two-mode squeezer

Let us now consider that the input is a product state. We are going to compute the corresponding output in a two-mode squeezer.

$$\begin{aligned} W'(x_1,p_1,x_2,p_2) &= W_1(\cosh r x_1 - \sinh r x_2, \cosh r p_1 + \sinh r p_2) \\ &\times \; W_2(-\sinh r x_1 + \cosh r x_2, \sinh r p_1 + \cosh r p_2). \end{aligned} \tag{4.71}$$

Then, tracing over mode 2 gives the following:

$$\begin{aligned} W_1'(x_1,p_1) = \iint \quad & W_1(\cosh r x_1 - \sinh r x_2, \cosh r p_1 + \sinh r p_2) \\ & W_2(-\sinh r x_1 + \cosh r x_2, \sinh r p_1 + \cosh r p_2) \quad \mathrm{d}x_2 \mathrm{d}p_2. \end{aligned} \tag{4.72}$$

With the change of variables

$$\begin{cases} x' = \cosh r \left(\cosh r x_1 - \sinh r x_2\right) \\ p' = \cosh r \left(\cosh r p_1 + \sinh r p_2\right) \end{cases} \qquad \begin{cases} \mathrm{d}x' = -\cosh r \sinh r \mathrm{d}x_2 \\ \mathrm{d}p' = \cosh r \sinh r \mathrm{d}p_2 \end{cases} \tag{4.73}$$

this leads us to the following equation

$$W_1'(x_1,p_1) = \iint \frac{1}{\cosh^2 r} W_1\left(\frac{x'}{\cosh r}, \frac{p'}{\cosh r}\right) \frac{1}{\sinh^2 r} W_2\left(\frac{x_1 - x'}{\sinh r}, \frac{p' - p_1}{\sinh r}\right) \mathrm{d}x_2 \mathrm{d}p_2, \tag{4.74}$$

which we can rewrite in a simple notation as $W_1' = \mathcal{L}_{\cosh r}[W_1] * \Pi \circ \mathcal{L}_{\sinh r}[W_2]$. We now come to the second output W_2':

$$\begin{aligned} W_2'(x_2,p_2) = \iint \quad & W_1(\cosh r x_1 - \sinh r x_2, \cosh r p_1 + \sinh r p_2) \\ & W_2(-\sinh r x_1 + \cosh r x_2, \sinh r p_1 + \cosh r p_2) \quad \mathrm{d}x_1 \mathrm{d}p_1. \end{aligned} \tag{4.75}$$

We perform the following change of variables

$$\begin{cases} x' = \cosh r \left(-\sinh r x_1 + \cosh r x_2\right) \\ p' = \cosh r \left(\sinh r p_1 + \cosh r p_2\right) \end{cases} \qquad \begin{cases} \mathrm{d}x' = -\cosh r \sinh r \mathrm{d}x_1 \\ \mathrm{d}p' = \cosh r \sinh r \mathrm{d}p_1 \end{cases} \tag{4.76}$$

which leads us to the following relation

$$W_2'(x_2,p_2) = \iint \frac{1}{\sinh^2 r} W_1\left(\frac{x_2 - x'}{\sinh r}, \frac{p' - p_2}{\sinh r}\right) \frac{1}{\sinh^2 r} W_2\left(\frac{x'}{\cosh r}, \frac{p'}{\cosh r}\right) \mathrm{d}x_2 \mathrm{d}p_2, \tag{4.77}$$

that we can rewrite in a simple form as $W_2' = \Pi \circ \mathcal{L}_{\sinh r}[W_1] * \mathcal{L}_{\cosh r}[W_2]$. We can now resume the relations we have derived for the two-mode squeezer in the following table.

	$\hat{\rho}'$	W'	$\hat{a}'$
mode 1	$\mathrm{Tr}_2\left[\hat{\Sigma}_r \left(\hat{\rho}_1 \otimes \hat{\rho}_2\right) \hat{\Sigma}_r^\dagger\right]$	$\mathcal{L}_{\cosh r}[W_1] * \Pi \circ \mathcal{L}_{\sinh r}[W_2]$	$\cosh r \; \hat{a}_1 + \sinh r \; \hat{a}_2^\dagger$
mode 2	$\mathrm{Tr}_1\left[\hat{\Sigma}_r \left(\hat{\rho}_1 \otimes \hat{\rho}_2\right) \hat{\Sigma}_r^\dagger\right]$	$\Pi \circ \mathcal{L}_{\sinh r}[W_1] * \mathcal{L}_{\cosh r}[W_2]$	$\sinh r \; \hat{a}_1^\dagger + \cosh r \; \hat{a}_2$

Let us now resume our findings throughout the section. The following table presents the evolution according to a linear transformation in the Heisenberg picture and the Schrödinger picture.

Heisenberg picture	Schrödinger picture
$\hat{a} \to \exp(i\varphi)\hat{a}$	$W \to \mathcal{R}_\varphi[W]$
$\hat{a} \to s\hat{a}$	$W \to \mathcal{L}_s[W]$
$\hat{a} \to \hat{a} + \alpha$	$W \to \mathcal{D}_\alpha[W]$
$\hat{a} \to \hat{a}^\dagger$	$W \to \Pi[W]$
$\hat{a} \to \hat{a} + \hat{b}$	$W \to W_A * W_B$

4.4 Gaussian states

As their name indicates, Gaussian states are described by a Gaussian Wigner function. As such, they are approriately described by their two first statistical moments. In this section, we are first going to define the mean displacement and covariance matrix of a quantum state. We are going to highlight their relevance and show how they define Gaussian states. We will also present how symplectic transformations act on the covariance matrix of a quantum state.

4.4.1 Covariance matrix of a quantum state

At this point, it is interesting to define the covariance matrix of a quantum state. In what follows, we consider a N-mode state $\hat{\rho}$. We define the mean displacement of $\hat{\rho}$ as:

$$\bar{\mathbf{q}} = \langle \hat{\mathbf{q}} \rangle, \qquad (\bar{\mathbf{q}})_i = \langle \hat{q}_i \rangle. \tag{4.78}$$

So that we write $\bar{q}_i$ as the mean of the operator $\hat{q}_i$ over state $\hat{\rho}$. The vector $\bar{\mathbf{q}} \in \mathbb{R}^{2N}$ describes the first statistical moment of the Wigner function of $\hat{\rho}$. The information about the second statistical moments of the Wigner function of $\hat{\rho}$ is encoded in its covariance matrix $\mathbf{V}$, that we defined as follows:

$$(\mathbf{V})_{ij} = V_{ij} = \frac{1}{2}\langle \{\hat{q}_i - \langle \hat{q}_i \rangle, \hat{q}_j - \langle \hat{q}_j \rangle\} \rangle \tag{4.79}$$

$$= \frac{1}{2}\langle \{\hat{q}_i, \hat{q}_j\} \rangle - \langle \hat{q}_i \rangle \langle \hat{q}_j \rangle \tag{4.80}$$

Any quantum state $\hat{\rho}$ is associated to a mean displacement $\bar{\mathbf{q}}$ and covariance matrix $\mathbf{V}$. Note that to the contrary of Chapter 2, we are here considering distributions that can take negative values. As such, they are not probability distributions.

When considering bosonic quantum systems, the covariance matrix is particularly relevant as it is directly related to the energy of the state. Indeed, remember that the bosonic Hamiltonian of a N-mode quantum state reads as follows (using $\hbar = \omega = 1$):

$$\hat{H} = \sum_{k=1}^{N} \left(\hat{n}_k + \frac{1}{2} \right) = \frac{1}{2} \sum_{k=1}^{N} \hat{x}_k^2 + \hat{p}_k^2. \tag{4.81}$$

Using that relation, we can express the energy of the quantum state $\hat{\rho}$ as a function of its mean displacement $\bar{\mathbf{q}}$ and covariance matrix $\mathbf{V}$:

$$\langle \hat{H} \rangle = \mathrm{Tr}\left[\hat{H}\hat{\rho} \right] = \frac{1}{2} \left(\mathrm{Tr}\, \mathbf{V} + \|\bar{\mathbf{q}}\|^2 \right). \tag{4.82}$$

Note that this expression holds for any bosonic quantum state, Gaussian and non-Gaussian. In addition to this, the covariance matrix measures in a certain way the extent in phase space, so that it

can be related to the uncertainty principle. As a consequence, some instances of covariance matrix are not allowed. In general, any N-mode quantum state $\hat{\rho}$ is associated to covariance matrix $\mathbf{V} \in \mathbb{R}^{2N\times 2N}$ which is symmetric ($\mathbf{V}^\intercal = \mathbf{V}$) and positive-definite ($\mathbf{V} > 0$), and which respect the following condition:

$$\mathbf{V} + \frac{i}{2}\mathbf{\Omega} \geq 0. \tag{4.83}$$

Condition (4.83) is a direct manifestation of the uncertainty principle. In the case of single-mode states, (4.83) becomes the Schrödinger-Robertson uncertainty relation $\det \mathbf{V} \geq 1/4$, which is a stronger version of Heisenberg uncertainty relation since we have the relation $\sigma_x^2 + \sigma_p^2 \geq \det \mathbf{V}$.

In the very case of Gaussian states, the mean displacement and the covariance matrix are sufficient to describe the whole Wigner function, which then reads as follows:

$$W_{\hat{\gamma}}(\mathbf{q}) = \frac{1}{(2\pi)^N \sqrt{\det \mathbf{V}}} \exp\left(-\frac{1}{2}\left(\mathbf{q}-\bar{\mathbf{q}}\right)^\intercal \mathbf{V}^{-1}\left(\mathbf{q}-\bar{\mathbf{q}}\right)\right). \tag{4.84}$$

Let us finally mention that for any acceptable choice of covariance matrix $\mathbf{V}$ satisfying condition (4.83), there exists a Gaussian state with such covariance matrix $\mathbf{V}$. We are now going to see how it is possible to build quantum states which such a function.

4.4.2 Pure Gaussian states

Earlier in this chapter, we have introduced the set of pure states known as coherent states. The set of coherent state can be generated from vacuum with the displacement operator. As they are described by a Gaussian Wigner function, coherent states belong to the family of Gaussian pure states. Any single-mode Gaussian pure state can be generated from vacuum, together with the squeezing, rotation and displacement operators. A single-mode Gaussian state with squeezing r, rotation angle φ and displacement α is associated to the ket $|\gamma(r,\varphi,\alpha)\rangle = \hat{D}_\alpha \hat{R}_\varphi \hat{S}_r |0\rangle$.

$$\hat{\gamma}_{(r,\varphi,\alpha)} = \hat{D}_\alpha \hat{R}_\varphi \hat{S}_r |0\rangle \langle 0| \hat{S}_r^\dagger \hat{R}_\varphi^\dagger \hat{D}_\alpha^\dagger \tag{4.85}$$

The wave-function of a single-mode Gaussian pure states depends on the parameters $\alpha \in \mathbb{C}$, $\varphi \in [0, 2\pi)$ and $r \in \mathbb{R}_+$. Its wave-function can be computed as the following [26]:

$$\begin{aligned} \psi(x) = \langle x| \hat{D}_\alpha \hat{R}_\varphi \hat{S}_r |0\rangle = &\left(2\pi\left(\cosh(2r) - \cos(2\varphi)\sinh(2r)\right)\right)^{-\frac{1}{4}} \\ \times \quad &\exp\left(-\frac{(x-x_\alpha)^2}{2}\frac{\cosh r + \exp(2i\varphi)\sinh r}{\cosh r - \exp(2i\varphi)\sinh r}\right) \\ \times \quad &\exp(ip_\alpha\left(x - x_\alpha/2\right)) \end{aligned} \tag{4.86}$$

where we have used the parameters $x_\alpha, p_\alpha \in \mathbb{R}$ such that $\alpha = (x_\alpha + ip_\alpha)/\sqrt{2}$. Multi-mode Gaussian pure states are built by tensoring Gaussian pure states and applying any combination of beam-splitter and two-mode squeezer over the resulting state.

The mean displacement of a Gaussian pure state is $\mathbf{q} = (x_\alpha, p_\alpha) = \sqrt{2}\,(\mathrm{Re}\,\alpha, \mathrm{Im}\,\alpha)$. As an illustration, the covariance matrix of a Gaussian pure states is [26]:

$$\mathbf{V}_{(r,\varphi)} = \frac{1}{2}\begin{pmatrix} \cosh(2r) - 2\cos(2\varphi)\sinh(2r) & -\sin(2\varphi)\sinh 2r \\ -\sin(2\varphi)\sinh(2r) & \cosh(2r) + 2\cos(2\varphi)\sinh(2r) \end{pmatrix} \tag{4.87}$$

One should notice that the determinant of the covariance matrix is always equal to $\det \mathbf{V} = 1/4$. The covariance matrix of vacuum is $\mathbf{V}_0 = \mathbf{I}_2/2$. It is the same for any coherent state. The determinant of a multi-mode pure Gaussian state is $\det \mathbf{V} = 2^{-2N}$.

4.4.3 From thermal states to mixed Gaussian states

The set of quantum states with Gaussian Wigner functions is broader than the sole set of Gaussian pure states. Indeed, when considering mixed states, it is possible to construct Gaussian Wigner functions. A thermal state (or Gibbs state) is a statistical mixture that is in a thermal equilibrium. As such,

it can be related to the Hamiltonian as $\hat{\tau} \propto \exp\left(-\hat{H}/(k_B T)\right)$, where k_B is the Boltzman constant and T is the temperature. A thermal state is a Gibbs state [42]. In a quantum optics context, we introduce the parameter β as the inverse temperature, such that $\beta = \hbar\omega/(k_B T)$. A quantum optical thermal state can then be developed onto the basis of Fock states as:

$$\hat{\tau} = (1-y)\sum_{n=0}^{\infty} y^n \left|n\right\rangle \left\langle n\right| \tag{4.88}$$

where $y \in [0,1)$ is the thermal parameter and is related to the inverse temperature as $y = \exp(-\beta)$. A thermal state is equivalently described by its mean number of photons $\bar{n} = y/(1-y)$. It has an energy equal to $\bar{n} + 1/2$. The purity of a thermal state is $\mu = (2\bar{n}+1)^{-1}$. It has a von Neumann entropy equal to:

$$S(\hat{\tau}) = (\bar{n}+1)\ln(\bar{n}+1) - \bar{n}\ln\bar{n} = g(\bar{n}) \tag{4.89}$$

where we define the function $g(\bar{n})$ as the von Neumann entropy of a thermal state with mean number of photon $\bar{n}$. A thermal state maximizes the von Neumann entropy for a given energy. Conversely, it minimizes the energy for a given von Neumann entropy. A thermal state is unequivocally defined by its inverse temperature β, its thermal parameter y, its mean number of photon $\bar{n}$ or its von Neumann entropy. In the present thesis, we will use the mean number of photon $\bar{n}$ and the thermal parameter y to characterize a thermal state. The Wigner function of a thermal state can be computed to be a Gaussian distribution, radial and centered around the origin, with variance $V = \bar{n} + 1/2$:

$$W_{\hat{\tau}}(x,p) = \frac{1}{\pi\,(2\bar{n}+1)} \exp\left(-\frac{x^2+p^2}{2\bar{n}+1}\right) \tag{4.90}$$

It is a Gaussian distribution of zero mean displacement and with covariance $\mathbf{V} = (\bar{n}+1/2)\mathbf{I}_2 = (2\bar{n}+1)\mathbf{V}_0$. A thermal state can be understood as a vacuum that is displaced according to a normal distributed noise. Moreover, its Wigner function also corresponds to a rescaled vacuum. The Wigner function $W_{\hat{\tau}}$ of a thermal state $\hat{\tau}$ can be expressed as a function of vacuum as follows:

$$W_{\hat{\tau}} = W_0 * G_{\bar{n}} = \mathcal{L}_{\sqrt{2\bar{n}+1}}\left[W_0\right], \tag{4.91}$$

where $\bar{n}$ is the mean number of photons of $\hat{\tau}$ an $G_{\bar{n}}$ is a Gaussian distribution of mean displacement $(0,0)$ and covariance matrix $\bar{n}\mathbf{I}_2$.

From a thermal state, we may now construct any single-mode mixed Gaussian state in the same manner as Equation 4.85, as follows:

$$\hat{\gamma}_{(\bar{n},r,\varphi,\alpha)} = \hat{D}_\alpha \hat{R}_\varphi \hat{S}_r \;\hat{\tau}_{\bar{n}}\; \hat{S}_r^\dagger \hat{R}_\varphi^\dagger \hat{D}_\alpha^\dagger \tag{4.92}$$

The covariance matrix of state $\hat{\gamma}_{(\bar{n},r,\varphi,\alpha)}$ can be computed as $\mathbf{V}_{(\bar{n},r,\varphi)} = (2\bar{n}+1)\mathbf{V}_{(r,\varphi)}$.

4.4.4 Symplectic transformations

Previously in this chapter, we introduced a type of transformations that we called linear transformations (see Definition 4.34). Among these, we find in particular the set of transformations associated with Gaussian unitaries. Such a unitary operator$\hat{U}$ is associated to a symplectic matrix $\mathbf{U}$ which describes the evolution of the quadrature operators $\hat{\mathbf{q}}$ according to the transformation. As a consequence, the symplectic matrix $\mathbf{U}$ contains all the information needed in order to compute new mean displacement and new covariance matrix resulting from the transformation. For any quantum state that evolving according to $\hat{\rho} \longrightarrow \hat{U}\hat{\rho}\hat{U}^\dagger$, its mean displacement and covariance matrix evolve according to:

$$\bar{\mathbf{q}} \longrightarrow \mathbf{U}\bar{\mathbf{q}}, \qquad\qquad \mathbf{V} \longrightarrow \mathbf{U}\mathbf{V}\mathbf{U}^\intercal. \tag{4.93}$$

It should be noted that the above relations hold in particular for Gaussian states, but also in general to any quantum state. We call symplectic formalsim the framework in which we only consider the evolution of the mean displacement and covariance matrix of the quantum state. This tool appears particularly powerful because of the simplicity of the tools involved (a simple matrix product). Moreover it applies without distinction to Gaussian states and non-Gaussian states, even though for non-Gaussian states it does not provide a complete description of the state.

Chapter 5

Wigner-positivity

In this chapter, we are interested in a certain set of quantum states which have the particularity of being described by a Wigner function taking only non-negative values. We label this property as Wigner-positivity. This set is of particular interest to us in this work, because the states belonging to it all have the property that the usual measures used in probability theory are well defined. Let us also note that states with non-negative Wigner functions have the property of being classically simulatable, and therefore cannot give rise to a quantum advantage for computational tasks [47]. Determining precisely the set of states with a non-negative Wigner function is therefore also of interest in order to better understand which other states have a quantum advantage that can be exploited.

We begin this chapter with general considerations where we draw a parallel between the positive-semidefiniteness condition in state space and the Wigner-positivity condition in phase space. We then show that the set of states with non-negative Wigner functions forms a convex set. After that, we illustrate several ways to build Wigner-positive states, and present in particular a general setup using a balanced beam-splitter to build such states. In the last section of this chapter, we focus on the set of phase-invariant states to draw further conclusions. In particular, we analytically determine the set of Wigner-positive phase-invariant states containing at most two photons.

The content of this chapter has been published in the paper entitled *Quantum Wigner entropy* [64].

5.1 Introduction and preliminaries

5.1.1 Positive-definiteness versus Wigner-positivity

In Chapter 1, we have learned that any operator $\hat{A}$ can be represented in phase-space by a complex distribution $A(x,p)$. The mapping is embodied by the Weyl transform $\mathcal{T}_{\mathrm{w}}$ and it inverse $\mathcal{T}_{\mathrm{w}}^{-1}$, which are such that $A(x,p) = \mathcal{T}_{\mathrm{w}}\left[\hat{A}\right]$ and $\hat{A} = \mathcal{T}_{\mathrm{w}}^{-1}\left[A(x,p)\right]$. There is a one-to-one correspondence between quantum operators and complex phase-space distributions. Let us now restrict ourselves to Hermitian operators. From the definition of the Weyl transform, it is clear that such operators correspond to real-valued phase-space distributions. Similarly, there is a one-to-one correspondence between these two sets.

$$\hat{A} : \mathcal{H} \mapsto \mathcal{H}, \quad \hat{A}^\dagger = \hat{A} \qquad \longleftrightarrow \qquad A(x,p) : \mathbb{R}^2 \mapsto \mathbb{R} \tag{5.1}$$

In other words, any Hermitian quantum operator $\hat{A}$ corresponds to a unique real-valued phase-space distribution $A(x,p)$. Quantum states are such operators and correspond therefore to real-valued phase-space distributions, namely Wigner functions. However, there exist many real-valued phase-space distributions that are not acceptable Wigner functions. This is not only because of the normalization condition, which can be easily corrected. Indeed, many of such distributions do not correspond to positive semi-definite operators, and thus do not satisfy the physicality conditions.

It is in general a difficult task to determine whether a real-valued phase-space distribution $A(x,p)$ corresponds to a positive semi-definite operator. Proving that it is not positive semi-definite is easier, as we only need to provide a single physically acceptable Wigner function W which is such that $\iint W(x,p)A(x,p)\mathrm{d}x\mathrm{d}p < 0$ (see positive semi-definiteness condition (1.31)). To prove that it is positive

semi-definite, we would need to check that for *all possible* Wigner functions the result of the integral is non-negative. By proceeding in a more clever manner, we could limit ourselves to a complete basis, such as the Fock basis. Expressing the distribution $A(x,p)$ into the basis of Fock operators described by the distributions $W_{ij}(x,p)$ (see (4.23)), we could reconstruct a matrix and then verify that it is positive semi-definite. However, the Fock basis is infinite, and we would still require an infinite number of verifications.

Interestingly, one can consider the dual task of this problem as determining whether a given positive semi-definite Hermitian operator $\hat{A}$ is associated with a Weyl transform $A(x,p)$ taking non-negative values. Indeed, let us notice the apparent symmetry between these two problems:

- Given some $A(x,p)$, is the corresponding $\hat{A} \geq 0$? (5.2)
- Given some $\hat{A}$, is the corresponding $A(x,p) \geq 0$? (5.3)

It appears that problem (5.3) is as difficult to check as (5.2). Indeed, whereas we only need one single instance of (x,p) such that $A(x,p) < 0$ in order to prove that A is partly negative, the condition $A(x,p)$ has to be checked over all phase-space to ensure that $A(x,p) \geq 0$ for all x,p. When we are studying a phase-space distribution normalized to 1, problem (5.2) reduces to checking the physicality of the phase space distribution, i.e. checking that $A(x,p)$ is a physically acceptable Wigner function. Conversely, if we are provided with a normalized positive semi-definite operator $\hat{A}$ (which is then a density operator $\hat{\rho}$), problem (5.3) reduces to checking whether the density operator $\hat{A} = \hat{\rho}$ has a non-negative Wigner function. Throughout this chapter, we will be interested in the set of quantum states possessing that particular property, namely the states with non-negative Wigner functions. We define this property as follows:

Definition 5.1 (Wigner-positivity)**.** *A quantum state is said to be Wigner-positive if and only if it is described by a Wigner function that takes non-negative values everywhere in phase space, namely:*

$$W(\mathbf{x},\mathbf{p}) \geq 0 \qquad \forall \mathbf{x},\mathbf{p} \tag{5.4}$$

Conversely, we define Wigner-negative states as states which are not Wigner-positive. Wigner-negative states are such that there exists at least one phase-space coordinate $(\mathbf{x},\mathbf{p})$ such that $W(\mathbf{x},\mathbf{p}) < 0$.

5.1.2 The Wigner-positive set is a convex set

As we have explained, determining the set of Wigner-positive states seems to be a difficult task. In this chapter, we take advantage of a simple observation to address the problem, which is that the set of Wigner-positive states forms a convex set. Let us start by defining this concept precisely before going into details.

Definition 5.2 (Convex quantum set)**.** *The quantum set $\mathcal{A}$ is convex if and only if any mixture of states of $\mathcal{A}$ belongs to $\mathcal{A}$:*

$$\forall \hat{\rho}_1, \hat{\rho}_2 \in \mathcal{A}: \qquad p_1\hat{\rho}_1 + p_2\hat{\rho}_2 \in \mathcal{A} \tag{5.5}$$

where (p_1,p_2) is a probability vector.

A first set of states possessing this property of convexity that we can define is the set of all single-mode quantum states. We denote as $\mathcal{Q}$ the set of all single-mode density operators satisfying the three physicality conditions, *i.e.* normalization (1.16), Hermiticity (1.17) and positive semi-definiteness (1.18). This set is obviously convex from the fact that mixing physically acceptable quantum states yields another physically acceptable quantum state. Then, we define $\mathcal{Q}_+$ as the restriction of $\mathcal{Q}$ which is Wigner-positive, so that $\mathcal{Q}_+$ is the set of all single-mode Wigner-positive quantum states. The set $\mathcal{Q}_+$ is also convex, since mixing non-negative Wigner functions gives another non-negative Wigner function. The set of Wigner-negative states is then $\mathcal{Q} \setminus \mathcal{Q}_+$, which is not a convex set. Indeed, it is possible two mix two Wigner-negative states into a Wigner-positive state.

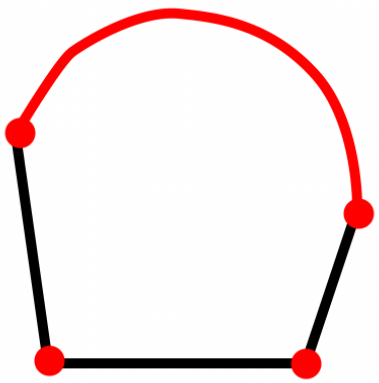

Figure 5.1: Schematic view of a convex set. The red points are extremal points of the set. Notice that some extremal points are isolated, while others form a continuum of points. Red points together with black points form the boundary of the convex set.

When considering convex sets, it is common to look at the minimum subset of states allowing us to generate the whole set by mixing them. Such states are called extremal states, and we define them hereafter:

Definition 5.3 (Extremal quantum state)**.** *The quantum state $\hat{\rho} \in \mathcal{A}$ is extremal with respect to $\mathcal{A}$ if and only if $\hat{\rho}$ cannot be obtained from a mixture of elements of $\mathcal{A} \setminus \{\hat{\rho}\}$*

$$\nexists \hat{\rho}_1, \hat{\rho}_2 \in \mathcal{A} \setminus \{\hat{\rho}\} : \qquad \hat{\rho} = p_1\hat{\rho}_1 + p_2\hat{\rho}_2 \tag{5.6}$$

where (p_1, p_2) is a probability vector.

In what follows, we will use the notation $\mathrm{Extr}\,(\mathcal{A})$ to denote the set of extremal states of $\mathcal{A}$. Figure 5.1 illustrates a convex set and its extremal points. In the same spirit, we define the convex hull of a quantum set as follows:

Definition 5.4 (Convex hull of a quantum set)**.** *The convex hull of the quantum set $\mathcal{A}$ is the smallest convex set that contains $\mathcal{A}$.*

An important property arising from that definition is that any compact convex set is equal to the convex hull of its extremal states. In practice, we will use the closure of the convex hull, which we will note as $\mathrm{Conv}(\mathcal{A})$. Note that the notation $\mathrm{Conv}(\mathcal{A})$ does not refer to the convex hull of $\mathcal{A}$, but to the *closure* of the convex hull of $\mathcal{A}$. This detail is important, especially when we consider mixtures composed of an infinite number of quantum states. Since $\mathrm{Conv}(\mathcal{A})$ is a closed set, it ensures that any mixture of states of $\mathcal{A}$ converges towards a state of $\mathrm{Conv}(\mathcal{A})$.

Now that we have introduced these different notions, let us focus on the case of $\mathcal{Q}$. The extremal states of $\mathcal{Q}$ are obviously the pure quantum states, since any mixed quantum state can be expressed as a convex mixture of quantum pure states, and since a pure state cannot be expressed as a mixture. We will denote the set of single-mode pure states as $\mathcal{Q}^*$, which obey the relations $\mathrm{Extr}\,(\mathcal{Q}) = \mathcal{Q}^*$ and $\mathrm{Conv}\,(\mathcal{Q}^*) = \mathcal{Q}$.

Let us now turn our attention to the set of extremal states of $\mathcal{Q}_+$. A natural candidate would be the set of pure states that are Wigner-positive, which we write as $\mathcal{Q}_+^*$. Indeed, each state of $\mathcal{Q}_+^*$ belongs to $\mathcal{Q}_+$ as it is Wigner-positive, but it is also necessarily an extremal state of $\mathcal{Q}_+$ because pure states are extremal to the quantum set $\mathcal{Q} \supset \mathcal{Q}_+$. However, things are more complicated as we will see. At this point, it seems important to us to introduce Hudson's theorem, which allows to define precisely the set $\mathcal{Q}_+^*$.

Definition 5.5 (Hudson theorem)**.** *A pure state is Wigner-positive if and only if it is a Gaussian pure state [34]. Equivalently, the set of Wigner-positive pure states $\mathcal{Q}_+^*$ is equal to the set of Gaussian pure states $\mathcal{G}^*$:*

$$\mathcal{Q}_+^* = \mathcal{G}^*. \tag{5.7}$$

We denote the set of single-mode Gaussian pure states as $\mathcal{G}^*$ and its convex hull as $\mathcal{G}$. The situation would then be very easily described if we had $\mathcal{Q}_+ = \mathcal{G}$, that is, if the extremal states of $\mathcal{Q}_+$ were solely the pure Wigner-positive states, i.e. $\text{Extr}\left(\mathcal{Q}_+\right) = \mathcal{G}^*$. However, the problem appears more complicated since there are mixed Wigner-positive states that cannot be constructed from Gaussian states. The convex hull of Gaussian pure states $\mathcal{G} = \text{Conv}\left(\mathcal{G}^*\right)$ thus only encompass a fraction of the set of Wigner-positve states $\mathcal{Q}_+$. Let us mention that the characterization of Wigner-positive mixed states has been attempted [46, 10], but the resulting picture is somehow complex.

5.2 Building Wigner-positive states

In this section, we will present several ways to construct Wigner-positive states. We restrict most of our considerations to single-mode states, but they can in general be easily extended to multi-mode states.

5.2.1 Classical states and Gaussian states

In Chapter 4, we introduced coherent states as pure quantum states corresponding to a displaced vacuum. We will denote the set of coherent states as $\mathcal{C}^*$ and to its convex hull as $\mathcal{C} = \text{Conv}\left(\mathcal{C}^*\right)$, which is usually called the set of classical states. Note that $\mathcal{C}^* \subset \mathcal{G}^*$, as coherent states are a Gaussian states with no squeezing. Any mixture of coherent states is naturally a Wigner-positive state. In this context, the Glauber-Sudarshan P-function has a particular interest in order to described convex mixture of coherent states. indeed, the P-function is defined as follows:

$$\hat{\rho} = \int P(\alpha) \left|\alpha\right\rangle \left\langle\alpha\right| \mathrm{d}^2\alpha, \tag{5.8}$$

so that for convex mixture of coherent states, $P(\alpha)$ simply corresponds to a probability distribution describing the mixture. Note that the P-function is defined for any quantum state (including states that do not belong to $\mathcal{C}$) . However, for states that are in $\mathcal{Q} \setminus \mathcal{C}$, the P-function does not correspond to a probability distribution [17, 28].

Remember now that the Wigner function of a coherent state is the Wigner function of a displaced vacuum. As a consequence, the Wigner function of a classical state can be expressed as a convolution between the P-distribution and W_0. To do so, we define the function $P(x,p) = P(x+ip)$ where x and p are real parameters. We can then write:

$$\begin{aligned} W(x,p) &= \iint P(x',p') W_0(x-\sqrt{2}x', p-\sqrt{2}p')\mathrm{d}x'\mathrm{d}p' && (5.9)\\ &= \iint \frac{1}{2} P\left(\frac{x''}{\sqrt{2}}, \frac{p''}{\sqrt{2}}\right) W_0(x-x'', p-p'')\mathrm{d}x''\mathrm{d}p'' && (5.10)\\ &= \iint \mathcal{L}_{\sqrt{2}}\left[P\right](x'',p'') W_0(x-x'', p-p'')\mathrm{d}x''\mathrm{d}p'' && (5.11)\\ &= \left(\mathcal{L}_{\sqrt{2}}\left[P\right] * W_0\right)(x,p) && (5.12) \end{aligned}$$

So that in a shorter writing, we can simply identify $W = \mathcal{L}_{\sqrt{2}}\left[P\right] * W_0$. Note that we can easily go a step further by allowing squeezed states in our mixture. In general, any mixture of Gaussian states can be expressed as follows:

$$\hat{\rho} = \int K(r,\varphi,\alpha)\ \hat{\gamma}_{(r,\varphi,\alpha)}\ \mathrm{d}r\mathrm{d}\varphi\mathrm{d}^2\alpha, \tag{5.13}$$

where $K(r,\varphi,\alpha)$ is a probability distribution which contains all the information about the mixture of Gaussian states. Any state constructed as (5.13) belongs to the convex hull of Gaussian states, that we note $\mathcal{G}$.

The two techniques we have presented here to build Wigner-positive states are trivial, since they simply consist in mixing pure Gaussian states which already have a non-negative Wigner function, according to Hudson theorem. Equation (5.13) actually enables us to construct any state which belongs to $\mathcal{G}$. In the following of this chapter we are going to present other techniques to build Wigner positive states which goes beyond the scope of $\mathcal{G}$.

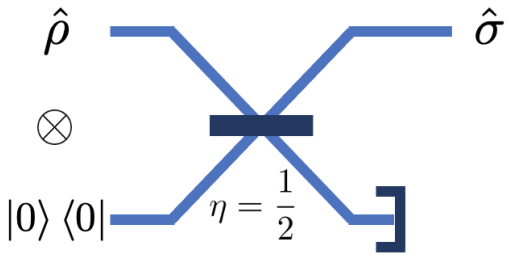

Figure 5.2: The state $\hat{\sigma}$ is the reduced output of a balanced beam-splitter of transmittance $\eta = 1/2$ fed by the input state $\hat{\rho}$, as described in Equation (5.22). In that particular setup, the Wigner function of $\hat{\sigma}$ coincides with the Husimi function of $\hat{\rho}$, so that $W_{\hat{\sigma}}(x,p) = Q_{\hat{\rho}}(x,p)$. As a consequence, any state $\hat{\sigma}$ built in this manner from any state $\hat{\rho}$ is Wigner-positive.

5.2.2 Husimi Q-function

In quantum optics, there exists a well-known distribution in quantum phase space that behaves as a genuine probability distribution, namely the Husimi Q-function [42]. It corresponds to the probability to measure state $\hat{\rho}$ in a coherent state $|\alpha\rangle$, and is defined as follows:

$$Q(\alpha) = \frac{1}{\pi} \langle\alpha| \hat{\rho} |\alpha\rangle \tag{5.14}$$

By a simple argument, we are going to show that the Husimi Q-distribution is also a physically acceptable Wigner function. We are going to show that the Husimi Q-distiribution $Q(x,p)$ of quantum state $\hat{\rho}$ is equal to the Wigner function $W(x,p)$ of another state $\hat{\sigma}$, in a particular setup. To that purpose, we define the distribution $Q(x,p) = Q(x+ip)$ where x and p are two real parameters.

$$\begin{aligned}
Q(x,p) &= \frac{1}{\pi} \langle x+ip| \hat{\rho} |x+ip\rangle && (5.15)\\
&= \frac{1}{\pi} \mathrm{Tr}\left[\hat{\rho} |x+ip\rangle \langle x+ip|\right] && (5.16)\\
&= \frac{1}{\pi} 2\pi \iint W(x',p') W_0(x'-\sqrt{2}x, p'-\sqrt{2}p) \mathrm{d}x' \mathrm{d}p' && (5.17)\\
&= 2 \iint W(x',p') W_0(\sqrt{2}x - x', \sqrt{2}p - p') \mathrm{d}x' \mathrm{d}p' && (5.18)\\
&= 2 \left(W * W_0\right)(\sqrt{2}x, \sqrt{2}p) && (5.19)\\
&= \mathcal{L}_{\frac{1}{\sqrt{2}}} \left[W * W_0\right](x,p) && (5.20)\\
&= \left(\mathcal{L}_{\frac{1}{\sqrt{2}}}[W] * \mathcal{L}_{\frac{1}{\sqrt{2}}}[W_0]\right)(x,p) && (5.21)
\end{aligned}$$

So that we can formally identify $Q = \mathcal{L}_{1/\sqrt{2}}[W] * \mathcal{L}_{1/\sqrt{2}}[W_0]$. This equation is familiar to us as we have introduced a similar pattern in Chapter 4. Indeed, it describes the action of a balanced beam-splitter acting on W and W_0. Let us define the state $\hat{\sigma}$ as follows:

$$\hat{\sigma} = \mathrm{Tr}_2\left[\hat{B}_{\frac{\pi}{4}} \left(\hat{\rho} \otimes |0\rangle\langle 0|\right) \hat{B}_{\frac{\pi}{4}}^\dagger\right]. \tag{5.22}$$

In that particular setup, which is pictured in Figure 5.2, we can identify the Husimi Q-function of the state $\hat{\rho}$ with the Wigner function of the state $\hat{\sigma}$, so that we can write $W_{\hat{\sigma}}(x,p) = Q_{\hat{\rho}}(x,p)$. As a consequence, it so happens that the state $\hat{\sigma}$ is Wigner-positive and that any state built in the manner described above is Wigner-positive. Any Husimi Q-function is thus an acceptable non-negative Wigner function.

At this point we should make an observation about the convolution of two Wigner functions.

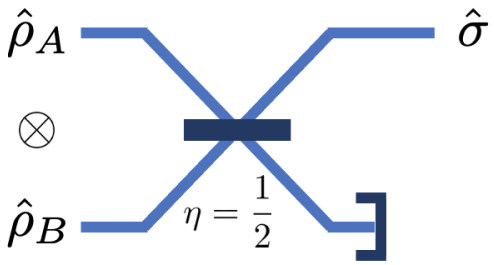

Figure 5.3: A beam-splitter of transmittance $\eta = 1/2$ is fed on the first input mode by $\hat{\rho}_A$ and on the second input mode by $\hat{\rho}_B$, so that $\hat{\rho}_A \otimes \hat{\rho}_B$ is a two-mode product state. After the action of the beam-splitter, the second output mode is discarded (traced out), and we are left on the first output mode with the single-mode state $\hat{\sigma}$. Any state $\hat{\sigma}$ constructed as depicted is Wigner-positive, i.e. it has a non-negative Wigner function. Note that we could have equivalently discarded the first output mode in place of the second output mode. By doing so, we would have obtained on the second output mode another state $\hat{\sigma}'$, which would also be Wigner-positive. This setup generalizes Figure 5.2.

Indeed, let us make the following development:

$$\begin{aligned}
(W_1 * W_2)(x,p) &= \iint W_1(x',p') W_2(x-x', p-p') \mathrm{d}x' \mathrm{d}p' && (5.23)\\
&= \iint W_1(x',p') \mathcal{R}_\pi [W_2](x'-x, p'-x) \mathrm{d}x' \mathrm{d}p' && (5.24)\\
&= \iint W_1(x',p') \mathcal{D}_{(x,p)} \circ \mathcal{R}_\pi [W_2](x',p') \mathrm{d}x' \mathrm{d}p' && (5.25)\\
&= \frac{1}{2\pi} \mathrm{Tr}\left[\hat{\rho}_1 \hat{D}_{(x,p)} \hat{R}_\pi \hat{\rho}_2 \hat{R}_\pi^\dagger \hat{D}_{(x,p)}^\dagger\right] && (5.26)
\end{aligned}$$

For each value of (x,p), the convolution of two Wigner functions can be linked to the overlap of two quantum states, which is always a non-negative quantity. As a consequence, the convolution of two Wigner function is always non-negative. Another interesting consequence arise from the fact that the overlap of two quantum states is lower or equal to 1, which means in turn that the convolution of two Wigner function is upper bounded by $1/(2\pi)$. The lower-bound is simply zero. We find the following inequality:

$$0 \leq (W_A * W_B)(x,p) \leq \frac{1}{2\pi}. \tag{5.27}$$

5.2.3 Beam-splitter states

In this subsection, we take advantage of relation (5.27) to present a generalization of the setup presented in Figure 5.2. Indeed, following the same development that we followed previously, we construct the state $\hat{\sigma}$ for two states $\hat{\rho}_A$ and $\hat{\rho}_B$ as follows:

$$\hat{\sigma} = \mathrm{Tr}_2\left[\hat{B}_{\frac{\pi}{4}} (\hat{\rho}_A \otimes \hat{\rho}_B) \hat{B}_{\frac{\pi}{4}}^\dagger\right]. \tag{5.28}$$

We call any state which can be written as such a *beam-plitter state*. See Figure 5.3 for an illustration. The Wigner function of $\hat{\sigma}$ can be obtained from the Wigner functions of $\hat{\rho}_A$ and $\hat{\rho}_B$, which we note respectively as W_A and W_B, as the following:

$$W_{\hat{\sigma}} = \mathcal{L}_{\frac{1}{\sqrt{2}}} [W_A * W_B]. \tag{5.29}$$

Thus, according to (5.27), we conclude that the resulting state $\hat{\sigma}$ is Wigner-positive. Note that the partial in (5.28) can equivalently be performed over mode 1. The two single-mode reduced outputs of the balanced beam-splitter are Wigner-positive. Let us mention that the observation that a balanced beam-splitter produces Wigner-positive states has also been highlighted in reference [4].

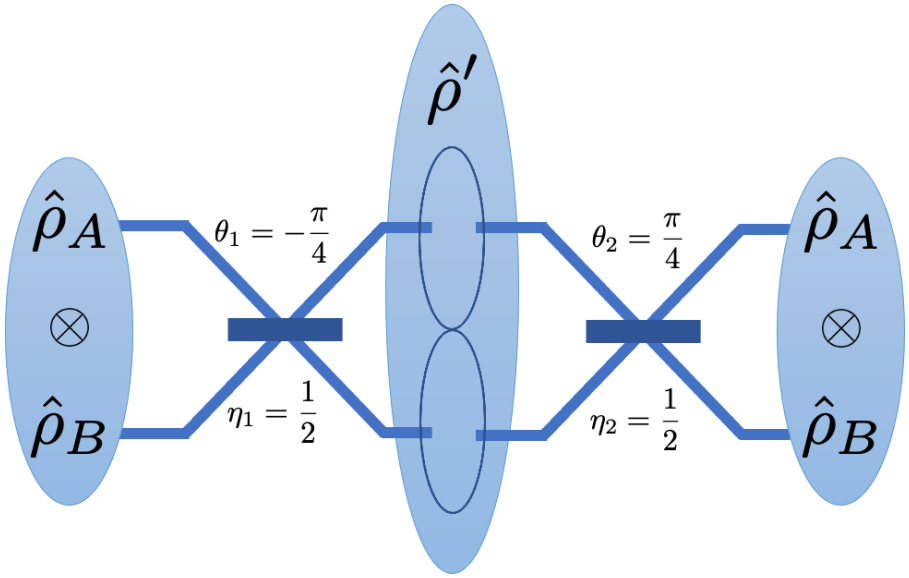

Figure 5.4: The setup described here is built from two successive balanced beam-splitters. The first beam-splitter is associated to the parameter $\theta_1 = -\pi/4$ (so that its transmittance is $\eta_1 = 1/2$) and to the unitary operator $\hat{B}_{\theta_1}$. The second beam-splitter is associated to the parameter $\theta_2 = \pi/4$ (so that its transmittance is $\eta_2 = 1/2$) and to the unitary operator $\hat{B}_{\theta_2}$. The successive application of the two beam-splitters corresponds to the unitary $\hat{B}_{\theta_2}\hat{B}_{\theta_1} = \hat{\mathbb{1}}$, since $\theta_1 + \theta_2 = 0$. As a consequence, the input of the first beam-splitter is identical to the output of the second beam-splitter. We choose the input of the first beam-splitter to be a two-mode product state $\hat{\rho}_A \otimes \hat{\rho}_B$. Between the two beam-splitters, we have the two-mode state $\hat{\rho}'$, which is in general not a product state. As soon as at least one of the two states $\hat{\rho}_A$ or $\hat{\rho}_B$ is Wigner-negative, the two-mode state $\hat{\rho}'$ is necessarily entangled since its output through a balanced beam-splitter is a Wigner-negative state. This naturally follows from the fact than any separable state yields a Wigner-positive state through the action of a balanced beam-splitter.

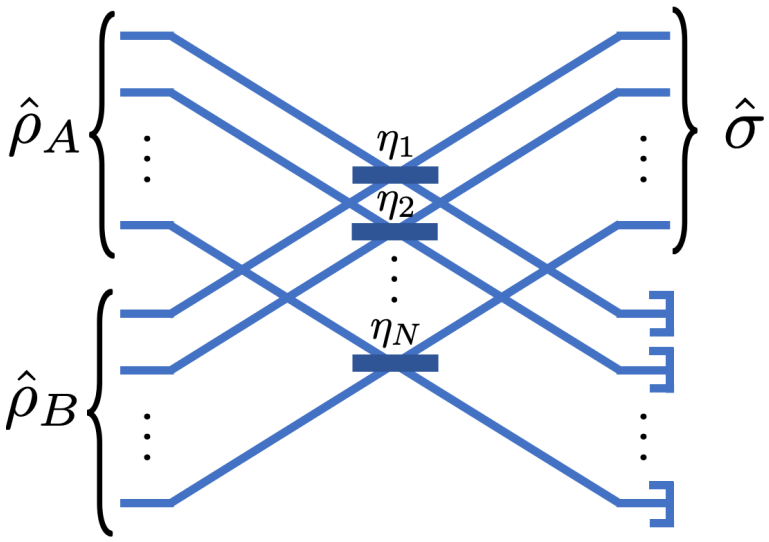

Figure 5.5: Optical implementation of a setup realizing multimode Wigner-positive states. The setup uses N beam-splitters with respective transmittance η_i. When the transmittance parameters η_i are all chosen to be equal to $1/2$, the N-mode output $\hat{\sigma}$ is Wigner-positive for any choice of $\hat{\rho}_A$ and ρ_B.

Note that in Equation (5.28), we can replace $\hat{\rho}_A \otimes \hat{\rho}_B$ whith any two-mode separable state and the output will still be Wigner-positive, since a mixture of Wigner-positive states remains Wigner-positive. An interesting consequence of that fact is that a balanced beam-splitter coupled with a Wigner-negative state always yields and entangled state. That observation is described in detail in Figure 5.4. Let us note that in this context it is appropriate to mention reference [27] which makes the link between entanglement and optical nonclassicality.

Let us now interest ourselves to a particular subset of beam-splitter states, which are built from pure states. We define the set $\mathcal{B}^*$ as the set of states which can be expressed as follows:

$$\hat{\sigma} = \mathrm{Tr}_2 \left[\hat{B}_{\frac{\pi}{4}} \left(|\psi_A\rangle \langle\psi_A| \otimes |\psi_B\rangle \langle\psi_B| \right) \hat{B}^\dagger_{\frac{\pi}{4}} \right]. \tag{5.30}$$

Note that until now we have been used to associate the superscript "$*$" with sets of pure states (such as $\mathcal{Q}^*$, $\mathcal{Q}^*_+$ or $\mathcal{G}^*$). The set $\mathcal{B}^*$, however, contains states which are in general not pure, but we still use the superscript "$*$" to stress that it is built from pure states. Notice that the convex hull of $\mathcal{B}^*$ encompass any state of the form (5.28), and even beyond as it also contains the set of beam-splitter states built from separable states. We define the set of beam-splitter states $\mathcal{B}$ as the convex hull of $\mathcal{B}^*$, so that $\mathcal{B} = \mathrm{Conv}\,(\mathcal{B}^*)$.

By a simple reasoning, we are going to show that the set $\mathcal{B}^*$ contains the set $\mathcal{G}^*$. To that purpose, let us consider any pure Gaussian state $\hat{\gamma} \in \mathcal{G}^*$. From the symplectic formalism introduced in Chapter 9, it can easily be shown that (when the displacement vector of $\hat{\gamma}$ is zero):

$$\hat{\gamma} = \mathrm{Tr}_2 \left[\hat{B}_\theta \left(\hat{\gamma} \otimes \hat{\gamma} \right) \hat{B}^\dagger_\theta \right], \tag{5.31}$$

so that $\hat{\gamma} \in \mathcal{B}^*$. In the case of a non-zero displacement vector, it is still possible to obtain $\hat{\gamma}$ with the appropriate choice of $\hat{D}_\alpha$. As a consequence, it appears that $\mathcal{G}^* \subset \mathcal{B}^*$ which also implies $\mathcal{G} \subset \mathcal{B}$.

Let us mention that beam-splitter states extends naturally to multimode states. The setup is pictured in Figure 5.5. The multimode state $\hat{\sigma}$ is built from the two multimode states $\hat{\rho}_A$ and $\hat{\rho}_B$ as follows:

$$\hat{\sigma} = \mathrm{Tr}_{(N+1)\cdots 2N} \left[\left(\bigotimes_{k=1}^{N} \hat{B}_{\frac{\pi}{4}} \right) \left(\hat{\rho}_A \otimes \hat{\rho}_B \right) \left(\bigotimes_{k=1}^{N} \hat{B}_{\frac{\pi}{4}} \right)^\dagger \right] \tag{5.32}$$

where each beam-splitter is coupling the mode i to the mode $N+i$ with $i = 1, ..., N$. It can then be shown that the Wigner function of $\hat{\sigma}$ is a convolution of the Wigner functions of $\hat{\rho}_A$ and $\hat{\rho}_A$:

$$W_{\hat{\sigma}} = \mathcal{L}_{\frac{1}{\sqrt{2}}} \left[W_A * W_B \right] \tag{5.33}$$

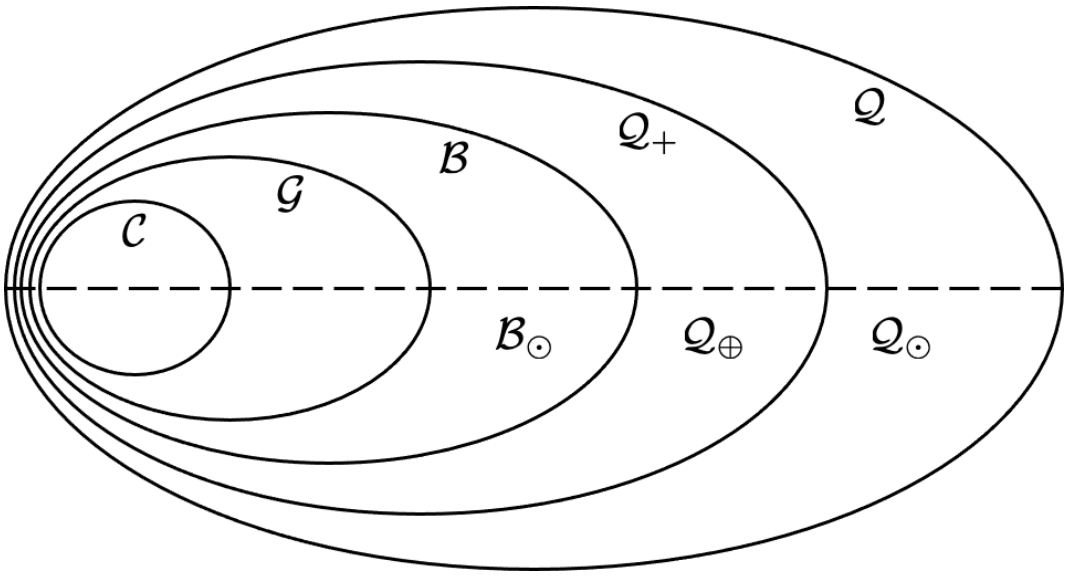

Figure 5.6: Pictorial representation of the various sets considered here. The full set of quantum states is denoted as $\mathcal{Q}$, while the set of Wigner-positive states is denoted as $\mathcal{Q}_+$. Then, $\mathcal{B}$ stands for the set of beam-splitter states, while $\mathcal{G}$ stands for the convex hull of Gaussian states. Further, $\mathcal{C}$ stands for the set of classical states. Within all these sets, we distinguish their phase-invariant restriction, which we label with the subscript "$\odot$" (or "$\oplus$" in the case of $\mathcal{Q}_+$). Any state in $\mathcal{Q}_\odot$ is associated to a probability vector of $\mathbb{S}$ and any state in $\mathcal{Q}_\oplus$ is associated to a probability vector in $\mathbb{S}_+$.

The convolution yields a Wigner-positive states, as we have the following relation:

$$(W_A * W_B)(\mathbf{x}, \mathbf{p}) = \frac{1}{(2\pi)^N} \operatorname{Tr}\left[\hat{\rho}_A \left(\bigotimes_{k=1}^{N} \hat{D}_{(x_k,p_k)}\hat{R}_\pi\right) \hat{\rho}_B \left(\bigotimes_{k=1}^{N} \hat{D}_{(x_k,p_k)}\hat{R}_\pi\right)^\dagger\right] \tag{5.34}$$

5.3 Focus on single-mode phase-invariant states

The methods we have presented so far allow us to construct positive Wigner quantum states of very different shapes. In this section, in order to make further progress in the characterization of the Wigner-positive set, we will restrict ourselves to single-mode quantum states with a phase-invariance symmetry. Such states possess a Wigner function with a radial symmetry, and are thus fully described by a radial function $W(r)$. We define the set $\mathcal{Q}_\odot$ as the set of single-mode phase-invariant states:

$$\mathcal{Q}_\odot = \{\hat{\rho} \in \mathcal{Q} : \quad \hat{R}_\theta \hat{\rho} \hat{R}_\theta^\dagger = \hat{\rho} \quad \forall\theta\}. \tag{5.35}$$

In accordance with the notation that we have introduced, we define $\mathcal{Q}_\odot^*$ as the set of pure phase-invariant states, so that $\mathcal{Q}_\odot^* = \{|n\rangle\langle n|\}$ is the set of Fock states. We define $\mathcal{Q}_\oplus$ as the set of Wigner-positive phase-invariant states, which is the set in which we are interested in now. Notice finally that the set of pure Wigner-positive and phase-invariant states $\mathcal{Q}_\oplus^*$ contains only vacuum. In general, when it comes to quantum sets, we associate the symbol "$\odot$" with phase-invariance and the symbol "$\oplus$" with both Wigner-positivity and phase-invariance. Conversely we define the corresponding sets of Wigner functions with the same symbols. In that manner, we define $\mathcal{W}_\odot = \mathcal{T}_\mathrm{w}[\mathcal{Q}_\odot]$ and $\mathcal{W}_\oplus = \mathcal{T}_\mathrm{w}[\mathcal{Q}_\oplus]$.

Remember that any state in $\mathcal{Q}_\odot$ can be expressed as a mixture of Fock states. As such, a phase-invariant state $\hat{\rho}$ is described by a probability vector $\mathbf{p} \in \mathbb{R}^\mathbb{N}$.

$$\hat{\rho} = \sum_{n=0}^{\infty} p_n |n\rangle\langle n| \tag{5.36}$$

The physicality conditions correspond to the non-negativity and normalization of $\mathbf{p}$:

$$p_n \geq 0 \quad \forall n, \qquad \sum_{n=0}^{\infty} p_n = 1. \tag{5.37}$$

We call $\mathbb{S}$ the restriction of $\mathbb{R}^{\mathbb{N}}$ satisfying the physicality conditions. Any vector $\mathbf{p}$ that belongs to $\mathcal{S}$ corresponds to a unique phase invariant state in $\mathcal{Q}_{\odot}$ with Wigner function in $\mathcal{W}_{\odot}$. Then, in order to determine whether $\hat{\rho}$ is Wigner-positive, we must check if its Wigner function is non-negative everywhere. The Wigner-positivity condition translates in a condition on $\mathbf{p}$ as follows:

$$W(r) \geq 0 \qquad \forall r \geq 0 \tag{5.38}$$

$$\Leftrightarrow \quad \sum_{n=0}^{\infty} p_n W_n(r) \geq 0 \qquad \forall r \geq 0 \tag{5.39}$$

$$\Leftrightarrow \quad \sum_{n=0}^{\infty} p_n (-1)^n L_n(2r^2) \exp\left(-r^2\right) \geq 0 \qquad \forall r \geq 0 \tag{5.40}$$

$$\Leftrightarrow \quad \sum_{n=0}^{\infty} p_n (-1)^n L_n(t) \geq 0 \qquad \forall t \geq 0 \tag{5.41}$$

In the above development, we have introduced the parameter $t = 2r^2$. Any phase-invariant state whose coefficients satisfy Equation (5.41) is Wigner-positive. We call the restriction of $\mathbb{S}$ satisfying the Wigner-positivity condition $\mathbb{S}_+$. Obviously, we have $\mathbb{S}_+ \subset \mathbb{S}$. Each value of $t \geq 0$ injected in the Wigner-positivity condition (5.41) gives the equation of a hyperplane dividing $\mathbb{S}$ in two halves, as $\mathbf{p}$ must be located on one side of the hyperplane to guarantee that $W(r) \geq 0$ for the corresponding r. Two hyperplane associated respectively to t and $t + \mathrm{d}t$ intersect in a lower-dimensional hyperplane which is at the boundary of the convex set $\mathbb{S}_+$. Mathematically, the condition that a point $\mathbf{p} \in \mathbb{S}$ belongs to the curved boundary of $\mathbb{S}_+$ is equivalent to the following condition:

$$\exists t \geq 0 : \quad \begin{cases} \displaystyle\sum_{k=0}^{\infty} p_k (-1)^k L_k(t) = 0 \\ \\ \displaystyle\sum_{k=0}^{\infty} p_k (-1)^k \frac{\mathrm{d}}{\mathrm{d}t} L_k(t) = 0 \end{cases} \tag{5.42}$$

Interestingly, a state satisfying the above condition is simply a state whose Wigner function has a local minimum which coincides with a zero. Note that since $\mathbb{S}_+$ is convex, all the points in its curved boundary are extremal points. However, other isolated extremal points will be shown to exist, as illustrated schematically in Figure 5.1.

5.3.1 Passive states

In the context of phase-invariant Wigner-positive states, we find it appropriate to introduce the set of passive states [41, 32]. Passive states are quantum states whose energy can only increase under the action of a unitary. As such, they cannot produce any useful work, which is why they are called passive. They are defined with respect to a given Hamiltonian $\hat{H}$, and a passive state with density operator $\hat{\rho}_p$ has the following property:

$$\mathrm{Tr}\left[\hat{H}\hat{\rho}_p\right] \leq \mathrm{Tr}\left[\hat{H}\hat{U}\hat{\rho}_p\hat{U}^{\dagger}\right] \qquad \forall \hat{U}. \tag{5.43}$$

It can be shown that passive states corresponds in general to decreasing mixtures of eigenstates of $\hat{H}$, in the sense that the probability coefficients should decrease considering increasing eigenvalues. This is consistent with the fact that unitaries preserve the eigenspectrum of a density operator, so that the optimal way to reduce the energy of the state is to give more probability to lower energy

states. In quantum optics, we are considering the bosonic Hamiltonian $\hat{H} = \hat{n} + 1/2$, so that passive states will be decreasing mixtures of Fock states:

$$\sum_{n=0}^{\infty} p_n \left|n\right\rangle \left\langle n\right| \qquad \text{with } p_n \geq p_{n+1} \tag{5.44}$$

It is interesting to define the set of extremal passive states, which correspond to equal mixture of the first Fock states. Extremal passive states are very useful because, as their name indicate, any passive state can be expressed as a convex mixture of extremal passive states. We refer to the density operator of the n^{th} extremal passive state as $\hat{\varepsilon}_n$ and to its Wigner functin as $E_n(x,p)$, which are defined as follows:

$$\hat{\varepsilon}_n = \frac{1}{n+1}\sum_{k=0}^{n} \left|k\right\rangle \left\langle k\right|, \qquad\qquad E_n(x,p) = \frac{1}{n+1}\sum_{k=0}^{n} W_k(x,p). \tag{5.45}$$

What justifies our interest in passive states in this chapter is that, in addition to being phase-invariant, they can be shown to possess a non-negative Wigner function. Indeed, starting from a formula linking Hermite polynomials to Laguerre polynomials [62], we can make the following derivation:

$$\begin{aligned}
& 2^n n! \sum_{k=0}^{n} (-1)^k L_k\left(2x^2+2y^2\right) = \sum_{k=0}^{n} \binom{n}{k} H_k(x)^2 H_{n-k}(y)^2 \\
\Leftrightarrow \quad & 2^n n! \frac{1}{\pi} \sum_{k=0}^{n} (-1)^k L_k\left(2x^2+2y^2\right) \exp\left(-x^2-y^2\right) \\
& \qquad = \frac{1}{\pi} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} H_k(x)^2 \exp\left(-x^2\right) H_{n-k}(y)^2 \exp\left(-y^2\right) \\
\Leftrightarrow \quad & \frac{1}{\pi} \sum_{k=0}^{n} (-1)^k L_k\left(2x^2+2y^2\right) \exp\left(-x^2-y^2\right) \\
& \qquad = \sum_{k=0}^{n} \left(\frac{1}{\sqrt{\pi} 2^k k!} H_k(x)^2 \exp\left(-x^2\right)\right) \left(\frac{1}{\sqrt{\pi} 2^{n-k}(n-k)!} H_{n-k}(y)^2 \exp\left(-y^2\right)\right)
\end{aligned} \tag{5.46}$$

So that we can finally identify the following relation:

$$\sum_{k=0}^{n} W_k(x,p) = \sum_{k=0}^{n} \psi_k(x)^2 \psi_{n-k}(p)^2 \tag{5.47}$$

where W_k and ψ_k respectively designate the Wigner function and wave-function of the k^{th} Fock state. From Equation (5.47), it appears that extremal states $\hat{\varepsilon}_n$ are Wigner-positive, so that the whole set of passive states is Wigner-positive, and we have $\{\hat{\varepsilon}_n\} \subset \mathcal{Q}_{\oplus}$.

5.3.2 Fock beam-splitter states

In the previous section, we have introduced the so-called beam-splitter states. Since in this section we are focusing on phase-invariant states, we are going to construct a set of phase-invariant Wigner-positive states by using a balanced beam-splitter. To do so, we are going to feed the beam-splitter with Fock states. For symmetry reasons, the output associated to such a set up is necessarily phase-invariant. Moreover, when the transmittance parameter of the beam-splitter is chosen to $\eta = 1/2$, the output is also Wigner-positive. With this in mind, we define the states $\hat{\sigma}(m,n)$ as follows:

$$\hat{\sigma}(m,n) = \mathrm{Tr}_2\left[\hat{B}_{\frac{\pi}{4}} \left(\left|m\right\rangle \left\langle m\right| \otimes \left|n\right\rangle \left\langle n\right|\right) \hat{B}_{\frac{\pi}{4}}^{\dagger}\right]. \tag{5.48}$$

These states, which we call Fock beam-splitter states, are obtained when feeding a balanced beam-splitter with respectively m and n photons. The expression of the state $\hat{\sigma}(m,n)$ can be decomposed

as follows:

$$\hat{\sigma}(m,n) = \frac{1}{m!n!2^m 2^n} \sum_{z=0}^{m+n} \sum_{i=\max(0,z-n)}^{\min(z,m)} \sum_{j=\max(0,z-n)}^{\min(z,m)} \binom{m}{i}\binom{n}{z-i}\binom{m}{j}\binom{n}{z-j}(-1)^{i+j} z!(m+n-z)!\,|z\rangle\langle z|\,. \tag{5.49}$$

Observe that $\hat{\sigma}(m,n) = \hat{\sigma}(n,m)$. Notice that since Equation (5.48) is a particular case of Equation (5.30), these Fock beam-splitter states belong to the set $\mathcal{B}^*$. Moreover, since they are also phase-invariant, they belong to the phase-invariant restriction of $\mathcal{B}^*$ that we label as $\mathcal{B}^*_\odot$. The convex hull of Fock beam-splitter states can be understood as the set of beam-splitter states that can be built from phase-invariant inputs. In the folllowing, we will define $\mathbb{S}_\mathrm{b}$ as the set of probability vectors which correspond to a mixture of Fock beam-splitter states. Obviously, we have $\mathbb{S}_\mathrm{b} \subseteq \mathbb{S}_+$ since mixtures of Fock beam-splitter states are both Wigner-positive and phase-invariant, and we will see later that the inclusion is actually strict.

Note that it is not clear that any state of $\mathcal{B}_\odot$ corresponds to a probability vector of $\mathbb{S}_\mathrm{b}$. Indeed, since it is possible to create beam-splitter states in the setup of Figure 5.3 starting from two input states that are not phase invariant (e.g., two squeezed states with orthogonal squeezing produce a thermal state), it might *a priori* be possible to build states within $\mathcal{B}_\odot$ that do not belong to $\mathbb{S}_\mathrm{b}$. Therefore we cannot at the moment state an equivalence between the set $\mathcal{B}_\odot$ and the set $\mathbb{S}_\mathrm{b}$, but we will however see later that they coincide when restricting ourselves to states containing up to two photons.

Let us now interest ourselves to the Wigner functions of these states $\hat{\sigma}(m,n)$. We define $S_{(m,n)}$ as the Wigner function of $\hat{\sigma}(m,n)$. Their Wigner function can be constructed as the mixture of Fock states with the coefficients of Equation (5.49). Also, we know that we can obtain them as the rescaled convolution of the Wigner functions Fock states:

$$S_{(m,n)} = \mathcal{L}_{\frac{1}{\sqrt{2}}}\left[W_m * W_n\right] \tag{5.50}$$

where W_m and W_n are respectivey the Wigner function of the m^{th} and n^{th} Fock state. As we will see later, these states generally correspond to extremal points of the set of Wigner-positive states, because their function generally has a zero. To illustrate this, let us now look at the value taken by $S_{(m,n)}$ at $(x,p) = (0,0)$.

$$S_{(m,n)}(0,0) = 2\left[W_m * W_n\right](0,0) \tag{5.51}$$

$$= 2\frac{1}{2\pi}\,\mathrm{Tr}\left[|m\rangle\langle m|\,\hat{D}_0\hat{R}_\pi\,|n\rangle\langle n|\,\hat{R}_\pi\hat{D}_0\right] \tag{5.52}$$

$$= \frac{1}{\pi}\,\mathrm{Tr}\left[|m\rangle\langle m|n\rangle\langle n|\right] \tag{5.53}$$

$$= \frac{1}{\pi}\delta_{mn} \tag{5.54}$$

where we have used the fact that Fock states are invariant under rotation and form an orthogonal basis. The Wigner function $S_{(m,n)}$ are plotted for several values of m,n in Figure 5.7.

Relation with extremal passive states

We are now going to highlight an interesting relation between extremal passive states $\hat{\varepsilon}_n$ and beam-splitter states $\hat{\sigma}(m,n)$. To that purpose, we define the two-mode state $\hat{\tau}(n)$ as an equal mixture of all the two-mode states with a total number of n photons ($\hat{\tau}(n)$ is not a thermal state). It can be written as follows:

$$\hat{\tau}(n) = \frac{1}{n+1}\sum_{k=0}^{n} |k\rangle\langle k| \otimes |n-k\rangle\langle n-k|\,. \tag{5.55}$$

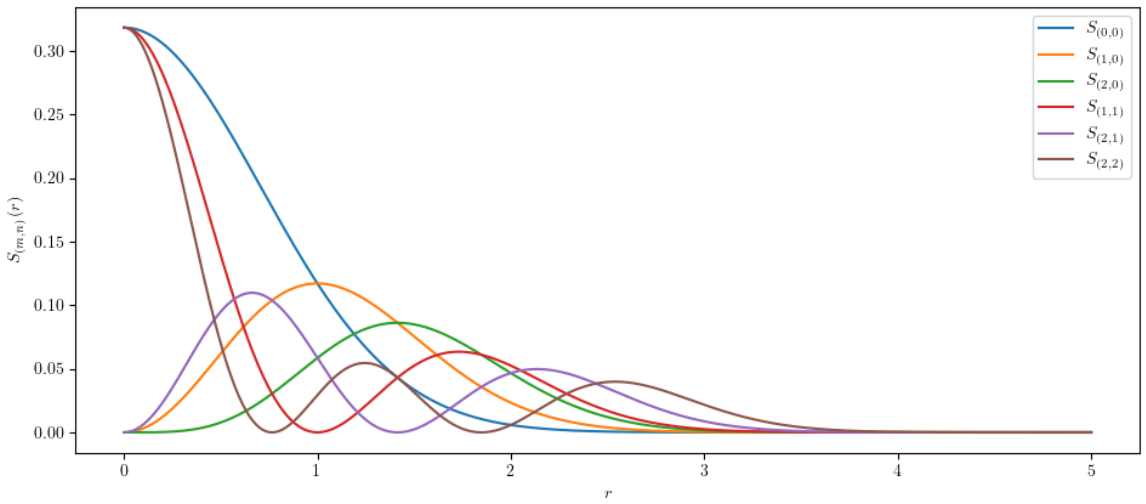

Figure 5.7: Radial functions of some states $\hat{\sigma}(m,n)$. As it appears, their Wigner function always cancels at least once (except for vacuum, which corresponds to $\hat{\sigma}(0,0)$).

By construction $\hat{\tau}(n)$ is a two-mode state with total energy that is fixed. It is maximally mixed over the set of two-mode quantum state with total number of photons n. For that reason, it is invariant under the action of any energy-preserving unitary. Hence, the action of an energy-preserving unitary will keep the state $\hat{\tau}(n)$ as a mixture of states with total number of photons n. Moreover, because unitary operators preserve the spectrum of eigenvalues, it will remain maximally mixed state over that subset. Applying these considerations to the particular case of a beam-splitter described by the unitary $\hat{B}_\theta$, we have the following equality:

$$\hat{\tau}_n = \frac{1}{n+1}\sum_{k=0}^{n} |k\rangle\langle k| \otimes |n-k\rangle\langle n-k| \tag{5.56}$$

$$= \frac{1}{n+1}\sum_{k=0}^{n} \hat{B}_\theta \left(|k\rangle\langle k| \otimes |n-k\rangle\langle n-k|\right) \hat{B}_\theta^\dagger \tag{5.57}$$

where the equality holds for any value of the parameter θ. Now, let us choose $\theta = \pi/4$ so that it corresponds to a balanced beam-splitter. After partial tracing the state over its second mode, we obtain the following relation:

$$\mathrm{Tr}_2\left[\hat{\tau}_n\right] = \frac{1}{n+1}\sum_{k=0}^{n} |k\rangle\langle k| = \frac{1}{n+1}\sum_{k=0}^{n} \hat{\sigma}(k,n-k). \tag{5.58}$$

That relation makes a link between the extremal passive state $\hat{\varepsilon}_n$ and the beam-splitter states $\hat{\sigma}(m,n)$. It expresses that extremal passive states can be obtained as a convex mixture of beam-splitter states $\hat{\sigma}(m,n)$. As a consequence, it implies the set of passive states is strictly included in the set of Fock beam-splitter states defined as the convex closure of $\hat{\sigma}(m,n)$. We can translate the above relation in terms of Wigner functions, which gives:

$$\sum_{k=0}^{n} W_k(x,p) = \sum_{k=0}^{n} S_{(k,n-k)}(x,p) \tag{5.59}$$

Equation (5.59) should be compared with Equation (5.47). Indeed, both relations gives a decomposition of the Wigner function of extremal passive states into a sum of non-negative functions. However, with Equation (5.47), the decomposition is carried out over functions which do not correspond to physically acceptable Wigner functions. Indeed, our numerical simulations have revealed that the function $\psi_k(x)^2\psi_{n-k}(p)^2$ does not correspond in general to a positive semi-definite operator. We then understand the advantage of Equation (5.59) which provides a decomposition of the Wigner function of an extremal passive state into a set of non-negative functions that correspond to physically acceptable Wigner functions.

5.3.3 Restriction up to two photons

Coming up with a precise description of the set of $\mathcal{Q}_\oplus$ appears to be difficult, even if that set is significantly easier to describe than the whole set $\mathcal{Q}_+$. In this subsection, we are going to restrict ourselves to a subset of $\mathcal{Q}_\odot$, which is the restriction of mixtures of Fock states up to two photons. To that purpose, let us denote by $\mathbb{S}^n$ and $\mathbb{S}^n_+$ the restriction of respectively $\mathbb{S}$ and $\mathbb{S}_+$ that have components $p_k = 0$ for $k > n$. Obviously, it appears that $\mathbb{S}^n_+ \subset \mathbb{S}^n$. Any probability vector of $\mathbb{S}^2$ is associated to a quantum state of $\mathcal{Q}_\odot$ through the following relation:

$$\hat{\rho} = (1 - p_1 - p_2) \left|0\right\rangle \left\langle 0\right| + p_1 \left|1\right\rangle \left\langle 1\right| + p_2 \left|2\right\rangle \left\langle 2\right| \tag{5.60}$$

The Wigner function of $\hat{\rho}$ can then be expressed as the corresponding mixture of Wigner function of Fock states. Let us recall the first Laguerre polynomials : $L_0(x) = 1$, $L_1(x) = -x + 1$ and $L_2(x) = (x^2 - 4x + 2)/2$. Using this, we find the Wigner function of $\hat{\rho}$ as the following:

$$W(r) = \frac{1}{\pi} \exp\left(-r^2\right) \left(2p_2 r^4 + (2p_1 - 4p_2)\, r^2 + 1 - 2p_1\right). \tag{5.61}$$

We can formulate the Wigner-positivity conditions using Equation (5.41). The state $\hat{\rho}$ is Wigner-positive if p_1, p_2 satisfy the following condition:

$$\frac{1}{2} p_2 t^2 + (p_1 - 2p_2)t + 1 - 2p_1 \geq 0 \qquad\qquad \forall t \geq 0 \tag{5.62}$$

We are now going to use this condition to characterize the set $\mathbb{S}^2_+$. First, we will determine what are the point of $\mathbb{S}^2$ which correspond to Wigner-positive states, which will give us the set $\mathbb{S}^2_+$. Then, will use Equation (5.42) to determine its extremal states.

Locus of Wigner-positivity

Eq. (5.62) is the equation of a parabola with a non-negative coefficient associated to t^2. We want that parabola to have non-negative values for $t \geq 0$. This is possible either if its discriminant Δ is non-positive ($\Delta \leq 0$), or if both its roots corresponds to $t \leq 0$. Let us examine the latter possibility first. If we have a parabola defined by $at^2 + bt + c = 0$, the sum of its roots is $-b/a$ and their product is c/a. The two roots are non-positive if their sum is non-positive and their product is non-negative. Applied to Eq. (5.62), this gives the following sufficient conditions:

$$\begin{cases} p_1 \geq 2p_2 \\ \\ p_1 \leq \dfrac{1}{2} \end{cases} \tag{5.63}$$

Condition (5.63) describes a locus which is the intersection of two half-planes. We define the region of the plane (p_1, p_2) satisfying that condition as region B. We now come to the discriminant condition. The discriminant is equal to $\Delta = 4p_2^2 - 2p_2 + p_1^2$, so that the condition $\Delta \leq 0$ can be written as:

$$\left(\frac{p_1}{1/2}\right)^2 + \left(\frac{p_2 - 1/4}{1/4}\right)^2 \leq 1. \tag{5.64}$$

Condition (5.64) describes an ellipse. We define the region of the plane (p_1, p_2) satisfying that condition as region A. Note that the union of the sets determined by conditions (5.63) and (5.64) alongside with the physicality conditions can be summarized as:

$$\begin{cases} p_1 \leq \dfrac{1}{2}, \\ \\ p_2 \leq \dfrac{1}{4} + \dfrac{1}{4}\sqrt{1 - 4p_1^2}, \end{cases} \tag{5.65}$$

with the additional constraint that $p_1, p_2 \geq 0$. Figure 5.8 illustrates the geometrical locus associated to the different conditions.

As we have explained previously, each value of t in Equation (5.62) gives the equation of a line which divides $\mathbb{S}^2$ in two half-planes. The situation is illustrated in Figure 5.9, where we have drawn these lines for several values of r (remember that $t = 2r^2$).

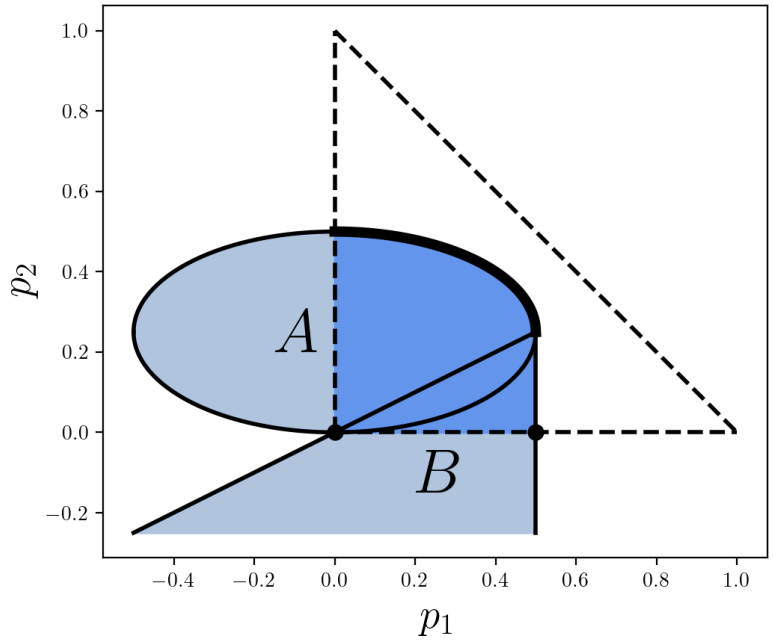

Figure 5.8: Geometrical locus of Wigner-positivity within the (p_1, p_2) plane, corresponding to the boundary of the dark blue region $\mathbb{S}_+^2$ satisfying Eq. (5.65). The boundary points and curve of $\mathbb{S}_+^2$ that are extremal are shown in bold. The ellipse (A) corresponds to the region where (p_1, p_2) is such that (5.62) is never negative. The semi-infinite triangular region (B) corresponds to values of (p_1, p_2) such that (5.62) becomes negative only for negative values of t. The dashed lines forming a triangle define the physicality limits, that is $p_1, p_2 \geq 0$ and $p_1 + p_2 \leq 1$. The union of (A) and (B) that belongs to the physicality triangle yields the dark blue region $\mathbb{S}_+^2$.

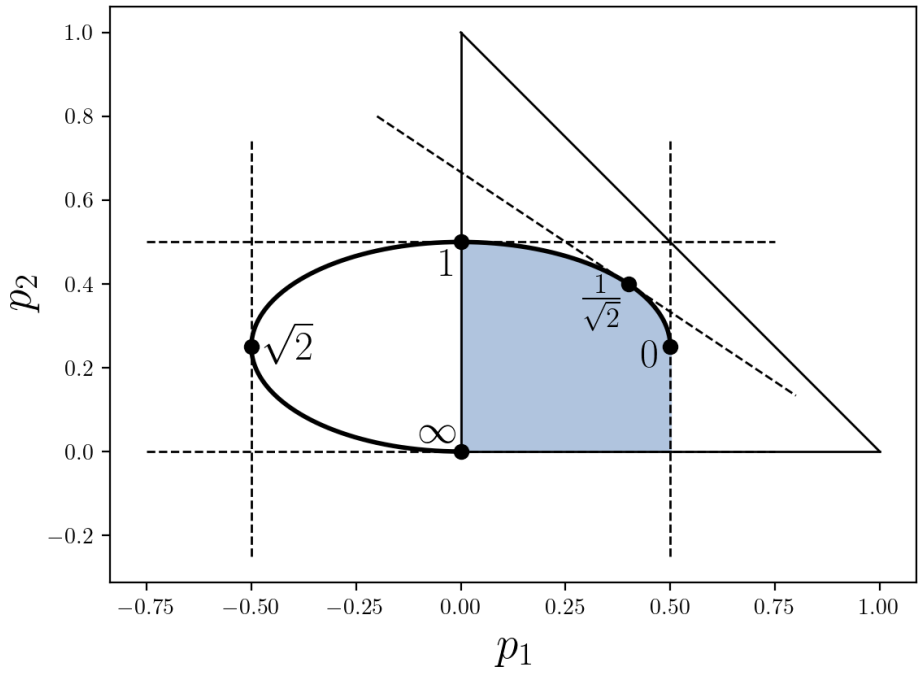

Figure 5.9: Expressing the positivity of the radial Wigner function $W(r)$ for increasing values of r corresponds to a continuum of straight lines, which are all tangents of ellipse (5.64). As an illustration, we plot as dashed lines the tangents associated with $W(r) = 0$ for $r = 0$, $r = 1/\sqrt{2}$, $r = 1$, $r = \sqrt{2}$, and $r \to \infty$. For instance, expressing $W(0) \geq 0$ implies $p_1 \leq 1/2$, while expressing $W(1) \geq 0$ implies $p_2 \leq 1/2$. For $r > 1$, the positivity condition becomes redundant, and, at the limit $r \to \infty$, it gives $p_2 \geq 0$, which is equivalent to the physicality condition.

Locus of extremal Wigner-positivity

The set of extremal Wigner-positive states in $\mathbb{S}^2_+$ can easily be identified from Figure 5.8 (where they appear in bold). Let us compare it with the set of states satisfying condition (5.42). Let us define $P(t)$ as the polynomial of the parabola described in Eq. (5.62). We refer to the first order derivative of $P(t)$ with respect to t as $P'(t)$.

$$\begin{aligned} P(t) &= \frac{1}{2}p_2 t^2 + (p_1 - 2p_2)t + 1 - 2p_1 \\ P'(t) &= p_2 t + p_1 - 2p_2 \end{aligned} \tag{5.66}$$

We are looking for the set of points (p_1, p_2) such that there exists one value of $t \geq 0$ which satisfies both $P(t) = 0$ and $P'(t) = 0$. The condition $P'(t) = 0$ is satisfied at $t = 2 - p_1/p_2$. Injecting that value of t in $P(t) = 0$ gives us the following equation $4p_2^2 - 2p_2 + p_1^2 = 0$, which corresponds to (5.64) with a strict equality. Remember that we have the additional constraint that $2p_2 \geq p_1$, since $t \geq 0$. This describes an arc of ellipse, that we can parameterize as follows:

$$\begin{cases} p_1 = \dfrac{1}{2}\sqrt{1 - a^2} \\ \\ p_2 = \dfrac{1}{4}(a + 1) \end{cases} \tag{5.67}$$

where the parameter a goes from 0 to 1. Injecting that parametrization in (5.61) yields the following expression:

$$W_a(r) = \frac{1}{\pi}\exp\left(-r^2\right)\frac{1}{2}(a+1)\left(r^2 - 1 + \sqrt{\frac{1-a}{1+a}}\right)^2. \tag{5.68}$$

$W_a(r)$ is the radial Wigner function of the extremal Wigner-positive states located on the arc of ellipse. As it appears, the condition 5.42 does not yields all the extremal states of $\mathbb{S}^2_+$, but only the extremal states which belong to the curved boundary of $\mathbb{S}^2_+$.

5.3.4 Discussion of the set $\mathbb{S}^2_+$

We have now properly determined the limits of the set $\mathbb{S}^2_+$ inside the set $\mathbb{S}^2$. Let us observe how the beam-splitter states $\hat{\sigma}(m, n)$ are located with respect to $\mathbb{S}^2_+$. We know that only the states $\hat{\sigma}(m, n)$ such that $m + n \leq 2$ will have at most 2 photons. For that reason, we can limit ourselves to look at the 4 different beam-splitter states that we define hereafter:

$$\hat{\sigma}_a \coloneqq \hat{\sigma}(0,0) = |0\rangle\langle 0| \tag{5.69}$$

$$\hat{\sigma}_b \coloneqq \hat{\sigma}(1,0) = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \tag{5.70}$$

$$\hat{\sigma}_c \coloneqq \hat{\sigma}(1,1) = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|2\rangle\langle 2| \tag{5.71}$$

$$\hat{\sigma}_d \coloneqq \hat{\sigma}(2,0) = \frac{1}{4}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| + \frac{1}{4}|2\rangle\langle 2| \tag{5.72}$$

Notice that $\hat{\sigma}_a$ is vacuum, and also corresponds to the extremal passive state $\hat{\varepsilon}_0$. The state $\hat{\sigma}_b$ corresponds to the extremal state $\hat{\varepsilon}_1$. The second extremal state is equal to $\hat{\varepsilon}_2 = (|0\rangle\langle 0| + |1\rangle\langle 1| + |2\rangle\langle 2|)/3$ and can be obtained as $(\hat{\sigma}_c + 2\hat{\sigma}_d)/3$. We label $\hat{\varepsilon}_2$ by the letter e in Figure 5.10, where we place the 4 beam-splitter states and the extremal state $\hat{\varepsilon}_2$. Each of these beam-splitter states corresponds to a precise point in $\mathbb{S}^2_+$. We define the convex hull of these points as the set $\mathbb{S}^2_{\mathrm{b}}$, which corresponds to all the states which can be obtained as a convex mixture from $\{\hat{\sigma}_a, \hat{\sigma}_b, \hat{\sigma}_c, \hat{\sigma}_d\}$.

Note that it is not a trivial observation to see that $\mathbb{S}^2_{\mathrm{b}}$ coincides with the two-photon restriction of $\mathcal{B}_\odot$ (i.e., the phase-invariant states with up to two photons within the convex hull of beam-splitter

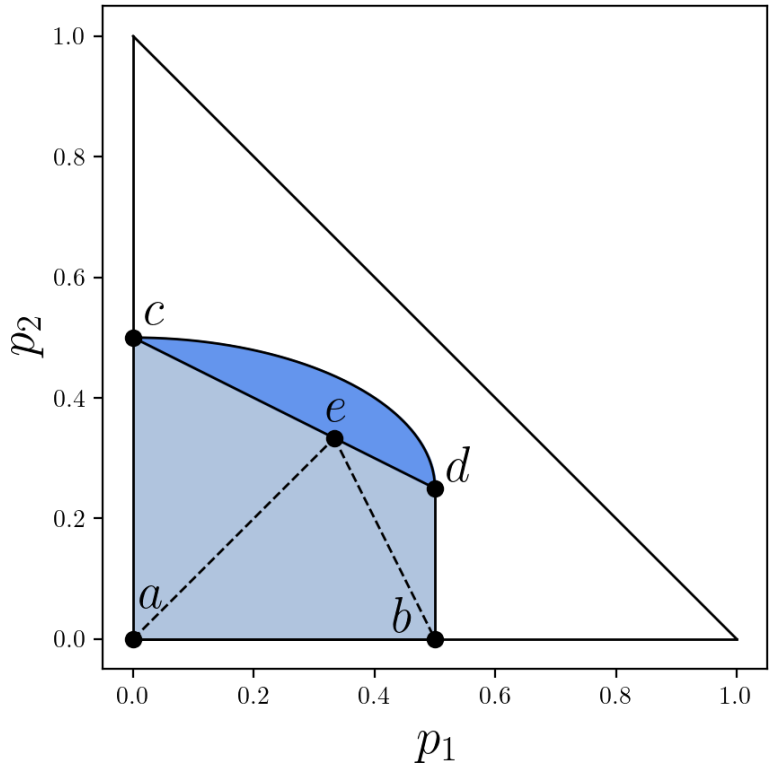

Figure 5.10: Graphic representation of the set $\mathbb{S}_+^2$ in $\mathbb{S}^2$. The outer triangle determines the region of physicality such that $p_1, p_2 \geq 0$ and $p_1 + p_2 \leq 1$, so that each point in the triangle corresponds to a quantum state in the restriction of $\mathcal{Q}_\odot$ with at most 2 photons. Each point in the blue region corresponds to a quantum state in the restriction of $\mathcal{Q}_\oplus$ with at most 2 photons. The light-blue region corresponds to the set $\mathbb{S}_\mathrm{b}$, i.e. the convex hull of Fock beam-splitter states. The dark-blue region corresponds to phase-invariant Wigner-positive states that cannot be expressed as a convex mixture of Fock beam-splitter states. The inner dashed triangle a-b-e encompasses the set of passive states while the trianlge a-b-d encompasses the states whose Wigner function coincides with the Husimi Q-function of a state.

states of $\mathcal{B}$). Since it is possible to create beam-splitter states in the setup of Fig. 5.3 that are phase-invariant starting from two input states that are not phase-invariant (*e.g.*, two squeezed states with orthogonal squeezing produce a thermal state), it might *a priori* be possible to build states within the two-photon restriction of $\mathcal{B}_\odot$ that do not belong to $\mathbb{S}_{\mathrm{b}}^2$. However, a simple argument convinces us otherwise. First, notice that we may restrict to pure input states without loss of generality. Since the output is a mixture with up to 2 photons, we must consider input states that are either in the form

$$|\psi\rangle = |0\rangle \otimes (a_0 |0\rangle + a_1 |1\rangle + a_2 |2\rangle)\,, \tag{5.73}$$

or

$$|\psi\rangle = (b_0 |0\rangle + b_1 |1\rangle) \otimes (c_0 |0\rangle + c_1 |1\rangle)\,. \tag{5.74}$$

In case (5.73), the first input is the vacuum, which is phase-invariant, so that the output state is phase-invariant only if the second input state is also phase-invariant. This is easy to understand given that the output Wigner function is a (scaled) convolution of the two input Wigner functions. In case (5.74), a straightforward calculation shows us that the output states is phase-invariant only if at least one of the coefficients b_0, b_1, c_0, or c_1 vanishes. This implies that one of the two input states must be phase-invariant, which in turns implies that the other input must be phase-invariant too in order to ensure the phase-invariance of the output. As a result, the two-photon restriction of $\mathcal{B}_\odot$ coincides with the set $\mathbb{S}_{\mathrm{b}}^2$ (it is unknown, however, whether this remains true for more than two photons, that is, whether $\mathcal{B}_\odot$ corresponds to the set $\mathbb{S}_{\mathrm{b}}$ in general). Since we have found phase-invariant Wigner-positive states outside $\mathbb{S}_{\mathrm{b}}^2$, this confirms that $\mathcal{B}$ is stricly included in $\mathcal{Q}_+$, as advertised earlier (see Fig. 5.6).

Several observations can be made from Figure 5.10. First, state $\hat{\sigma}_a$, which coincides with the vacuum state, is a trivial extremal state of $\mathbb{S}_+^2$ even if its Wigner function does not reach zero. As already mentioned, $\hat{\sigma}_b$, $\hat{\sigma}_c$, and $\hat{\sigma}_d$ are other extremal states of $\mathbb{S}_+^2$, as witnessed by the fact that their Wigner function vanishes at some location in phase space. The convex set $\mathbb{S}_+^2$ has three facets. Two of them correspond to the physicality conditions (5.37), i.e. $p_1 \geq 0$ and $p_2 \geq 0$. The third one corresponds to condition (5.41) where we have set $t = 0$, which gives us $p_0 + p_2 \geq 1/2$ or equivalently $p_1 \leq 1/2$. Note that the points in this third facet belong to the boundary of $\mathbb{S}_+^2$ but are not extremal. This can be easily understood as the segment in Fig. 5.10 connects $\hat{\sigma}_b$ and $\hat{\sigma}_d$, which both admit a zero of their Wigner function at the same location (i.e., the origin). Note also that, in general, the set $\mathbb{S}_+$ always has a facet corresponding to

$$\sum_{k \text{ even}} p_k = \frac{1}{2}. \tag{5.75}$$

which expresses the positivity of the Wigner function at $r = 0$ (recall that $t = 2r^2$). As pictured in Fig. 5.9, expressing the positivity of the radial Wigner function for increasing values of r yields a continuum of straight lines, whose locus of intersecting points forms an ellipse centered in $(0, 1/4)$, namely

$$\left(\frac{p_1}{1/2}\right)^2 + \left(\frac{p_2 - 1/4}{1/4}\right)^2 = 1\,. \tag{5.76}$$

Overall, Figure 5.10 shows that the subspace $\mathbb{S}_{\mathrm{b}}^2$, which is spanned by the extremal states $\hat{\sigma}_a$, $\hat{\sigma}_b$, $\hat{\sigma}_c$, and $\hat{\sigma}_d$, covers indeed a large region of $\mathbb{S}_+^2$ (indicated in light blue) so any point in this region can thus be generated by a convex mixture of them. However, $\mathbb{S}_+^2$ also includes a small region (indicated in dark blue) that is located under the ellipse defined by Eq. (5.76) and above the straight line c-d. This region is thus outside the polytope $\mathbb{S}_{\mathrm{b}}^2$ generated by the $\hat{\sigma}$ states, which confirms that $\mathbb{S}_+^2$ also admits a continuum of extremal points along this ellipse.

Finally, let us conclude this chapter by studying Figure 5.10 from the point of view of entanglement. Indeed, in Figure 5.4 we have highlighted how Wigner-negative states always produce entanglement with a balanced-beam splitter. A converse observation is that entanglement is required in order to produce Wigner-negative states with a balanced beam-splitter. With this in mind, let us now imagine that we have a balanced beam-splitter at our disposal and that we want to recreate the states of Figure 5.10. The white region inside the outer triangle corresponds to Wigner-negative states, which require thus an entangled input in order to be built. The light-blue region, which corresponds to the set $\mathbb{S}_{\mathrm{b}}$,

is associated with Wigner-positive states that can be constructed from separable inputs. In contrast, the dark-blue region corresponds to Wigner-positive states that cannot be built from separable inputs. Quantum states in that region have thus the rather surprising property that they require entanglement to be produced, even if their Wigner function is non-negative.

Chapter 6

Continuous majorization in quantum phase space

In this chapter, we focus on the central announced objective of this thesis, namely the application of the mathematical tool that is continuous majorization onto the Wigner functions which embody quantum states in phase space. As we have seen in Chapter 1, Wigner functions are distributions that go beyond the scope of probability distributions, due to their negativity originating from their quantum behavior. As such, the disorder of Wigner functions cannot be characterized only by a classical probabilistic approach. For this reason, we have in Chapter 3 focused our interest on the powerful theory of majorization, and we have shaped it for distributions taking both positive and negative values and defined over an infinite support. Chapters 1 and 3 therefore form the backgroundfor the current chapter. Note also that to some extent we will give a particular attention to non-negative Wigner functions. We will therefore also use several results derived in Chapter 5.

This chapter is structured in 4 sections. In the first section, we will properly define all the ideas of majorization in the framework of Wigner functions. We will develop the notions of level-functions, level-equivalence and rearrangements when the objects we consider are Wigner functions. This will allow us to write the conditions for a majorization relation between two Wigner functions.

Then, in the next section, we will ask ourselves the question of the existence of a state of least disorder in phase space. We will see in the rest of our reasoning that such a state cannot in all generality exist. However, restricting ourselves to Wigner-positive states, we will see that Gaussian states are very natural candidates to minimize the uncertainty over that subset. We will conclude the section by formulating a majorization conjecture applying to all Wigner-positive states.

Then, in the next section we will present a major result of our work, namely a proof of the majorization conjecture introduced in the previous section for a subset of Wigner-positive states. The set on which we will prove the conjecture is the set of phase-invariant states containing up to two photons, which is a set that we studied in detail in Chapter 5. The construction of our proof relies on several majorization lemmas that we prove separately.

Finally, we will conclude this chapter with a more prospective section, in which we will define slightly adapted versions of regular majorization. We will first define what we call radial-majorization, which is anticipated to play a role in quantum thermodynamics as it can, among other things, be related to a difference of energy between the two states involved. We will then define square-majorization, which applies only to pure states, and which will allow us to formulate an extended majorization conjecture.

To avoid any possible confusion, let us stress out that in the present chapter we only deal with continuous majorization between Wigner functions. At no time in this chapter is there any reference to a discrete majorization relation (except for Fock-majorization in the last section). When we say that a quantum state majorizes another one, we mean that their Wigner functions are related by a continuous majorization relation.

The content of this chapter corresponds to the paper entitled *Continuous majorization in quantum phase space* [65].

6.1 Definitions and preliminaries

In this section, we apply the notion introduced in Chapter 3 to distributions represented by Wigner functions of quantum states. To that purpose, we recall the mathematical objects introduced in the context of majorization and particularize them to Wigner functions. Wigner functions are infinite-domain continuous distributions. As such, they enter the scope of majorization that we have introduced in Chapter 3 in the particular case of continuous distributions defined over an infinite domain. In this chapter, we will consider single-mode Wigner functions. However, the definition and results can easily be generalized to multi-mode Wigner functions.

6.1.1 Level-equivalence in phase-space

The level-function of a Wigner function is the function which, in the prism of majorization, contains all the information we need to know. Wigner function are normalized distributions defined over an infinite domain. They take in general both positive and negative values. We define the upper level-function and lower level-function of a Wigner function according to Definition 3.9.

$$m_W^+(t) = \nu\left(\{(x,p) : W(x,p) \geq t\}\right), \qquad m_W^-(t) = \nu\left(\{(x,p) : W(x,p) \leq t\}\right). \tag{6.1}$$

As we remember from Chapter 1, the Wigner function is upper and lower bounded. This implies that its level-function takes the value zero outside of these bounds. Therefore, the upper level-function $m_W^+(t)$ will be zero for all $t \geq 1/\pi$, while the lower-level function $m_W^-(t)$ will be zero for all $t \leq -1/\pi$. Moreover, in the limit $t \to 0$, both the upper and lower level-functions tends towards infinity. it follows from the normalization condition that applies on the Wigner function that it has an infinite domain with value arbitrarily close to zero. The domain of interest of the upper and lower level-functions will then be respectively $(0, 1/\pi]$ and $[-1/\pi, 0)$. We define

$$\mu_W(t) = \begin{cases} -\dfrac{\mathrm{d}}{\mathrm{d}t} m_W^+(t) & t > 0 \\ \\ \dfrac{\mathrm{d}}{\mathrm{d}t} m_W^-(t) & t < 0 \end{cases} \tag{6.2}$$

so that level-density function has an infinite singularity in $t = 0$, so that when $t \to 0$ we have $\mu_W(t) \to \infty$. The domain of interest of level-density function $\mu_W(t)$ is $[-1/\pi, 1/\pi] \setminus \{0\}$. The level-density function μ_W can be used to compute any integral of the following form:

$$\iint \varphi\left(W(x,p)\right) \mathrm{d}x\mathrm{d}p = \int_{-\frac{1}{\pi}}^{\frac{1}{\pi}} \varphi(t)\mu_W(t)\mathrm{d}t \tag{6.3}$$

where the infinite singularity of the level-density function μ_W implies that the integration converges only if $\varphi(0) = 0$. Two Wigner functions are level-equivalent when their upper and lower level-functions are equal, or equivalently when their level-density function are equal. In that case, we write $W_A \equiv W_B$.

6.1.2 Rearrangements of a Wigner function

In order to define the rearrangements of a Wigner function, we use some notions defined in Chapter 3. Namely, we use the volume function $V_{\mathbb{A}}(s)$ of a set $\mathbb{A}$ (see (3.81)), as well as the functions M_W^+ and M_W^- which are the respective inverse of the upper and lower level-functions (see (3.83)). Single-mode Wigner functions are defined over $\mathbb{R}^2$, which is of infinite measure. The volume function of the set $\mathbb{R}^2$ is $V_{\mathbb{R}^2}(s) = \pi s^2$. The knowledge of the upper and lower level-functions of W enables us to build the functions M_W^+ and M_W^- which are the respective inverse of m_W^+ and m_W^- and can be obtained as $M_W^+(u) = \max\left(\{t : m_W^+(t) \geq u\}\right)$ and $M_W^-(u) = \min\left(\{t : m_W^-(t) \geq u\}\right)$. The decreasing and increasing rearrangements of W are then build. Wigner functions have in general two rearrangements. The decreasing rearrangements $W^\downarrow$ and the increasing rearrangement $W^\uparrow$, which are defined as follows:

$$W^\downarrow(x,p) = M_W^+(\pi(x^2+p^2)), \qquad W^\uparrow(x,p) = M_W^-(\pi(x^2+p^2)). \tag{6.4}$$

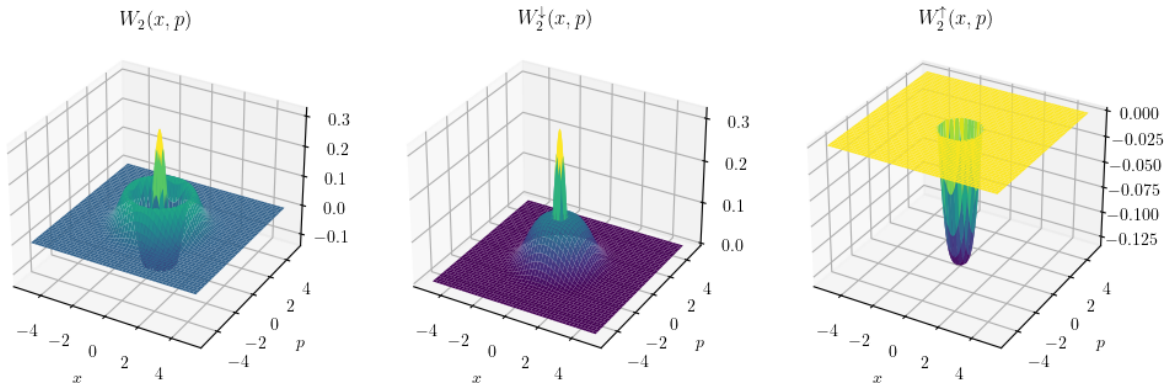

Figure 6.1: Example of decreasing and increasing rearrangements of a partly negative Wigner function. $W_2(x,p)$ is the Wigner function of the Fock state $n=2$. The decreasing rearrangement $W_2^{\downarrow}(x,p)$ is radial and decreasing. It is always non-negative, and normalized to some value greater than 1. The increasing rearrangement $W_2^{\uparrow}(x,p)$ is radial and increasing. It is always non-positive.

Since the decreasing and increasing rearrangements are radial functions that only depend on the distance r from the origin, we will simply refer to them as $W^{\downarrow}(r) = M_W^{+}(\pi r^2)$ and $W^{\uparrow}(r) = M_W^{-}(\pi r^2)$.

Let us now introduce the positive volume and negative volume as the following functionals:

$$\mathrm{Vol}_+(W) = \iint [W(x,p)]^+ \,\mathrm{d}x\mathrm{d}p, \qquad \mathrm{Vol}_-(W) = -\iint [W(x,p)]^- \,\mathrm{d}x\mathrm{d}p. \tag{6.5}$$

The positive volume $\mathrm{Vol}_+(W)$ corresponds to the volume of the positive parts of W, and the negative volume $\mathrm{Vol}_-(W)$ corresponds to the volume of the negative parts of W. They are both non-negative quantities. Notice that because of the normalization of W, we have $\mathrm{Vol}_+(V) - \mathrm{Vol}_-(W) = 1$.

The decreasing rearrangement $W^{\downarrow}$ takes its maximum value at $(x,p) = (0,0)$ and is then radially decreasing. It reaches 0 towards infinity. The norm of the decreasing rearrangement is equal to $\mathrm{Vol}_+(W)$ which is in general greater than zero. Conversely, the increasing rearrangement $W^{\uparrow}$ takes its minimum value at $(x,p) = (0,0)$ and is then radially increasing and reaches 0 towards infinity. The norm of the increasing rearrangement is equal to $-\mathrm{Vol}(W)$, which is negative. As we have explained in Chapter 3, $W^{\downarrow} \equiv [W]^+$ and $W^{\uparrow} \equiv [W]^-$, that is the decreasing rearrangement of W is level-equivalent to the positive parts of W, and the increasing rearrangement of W is level-equivalent to the negative parts of W. Figure 6.1 illustrate the decreasing and increasing rearrangements of a partly negative Wigner-function.

Let us now address the particular case of Wigner-positive states. Since the Wigner function of a Wigner-positive never takes negative values, its increasing rearrangement is zero everywhere: for Wigner-positive states, we have the relation $W^{\uparrow}(x,p) = 0$. Their decreasing rearrangement, however, is level-equivalent ot the original Wigner function and we have:

$$\forall W \in \mathcal{W}_+ : \qquad W^{\downarrow} \equiv W. \tag{6.6}$$

At this point it is interesting to ask us whether the decreasing rearrangement of a Wigner-positive state corresponds in general to a physically acceptable Wigner function. The question has been addressed in a previous work [63], and we came to the conclusion than it is in general not the case. Obviously, there are cases where the decreasing rearrangement indeed corresponds to a physical Wigner function. Indeed, if we consider thermal states for example, we see that their Wigner function is non-negative and radial decreasing, which implies that the Wigner-function is equal to its decreasing rearrangement. We have also found many examples of non-negative Wigner functions whose decreasing rearrangement is a distribution with abrupt variations of derivative (see [63] for more details).

Cumulative integrals

In Chapter 3, we have defined the cumulative integral of a distribution as the integration of the distribution over a ball with definite volume centered around the origin (see Definition 3.13). The

cumulative integral depends on a real parameter which defines the size of the ball. For convenience, the parameter we have used in Chapter 3 is the n-dimensional volume of the ball, as it was more suitable to the formulation of majorization. In the present chapter, however, we are going to use a parameter which corresponds to the radius of the ball. This will not affect the rest of our considerations and will facilitate the expression of our developments. We define the decreasing cumulative integral and increasing cumulative integral of W as follows:

$$S_s^{\downarrow}(W) = \int_0^s W^{\downarrow}(r) 2\pi r \mathrm{d}r, \qquad S_s^{\uparrow}(W) = \int_0^s W^{\uparrow}(r) 2\pi r \mathrm{d}r. \tag{6.7}$$

6.1.3 Majorization relations between Wigner functions

The definition of majorization directly follows from Chapter 3. Notice that by construction, Wigner functions are normalized. As a consequence, the statement $W_A \succ W_B$ is equivalent to any of the following conditions:

$$\iint \varphi\left(W_A(x,p)\right) \mathrm{d}x\mathrm{d}p \geq \iint \varphi\left(W_B(x,p)\right) \mathrm{d}x\mathrm{d}p \qquad \forall \varphi \text{ convex} \tag{6.8}$$

$$\begin{cases} \Phi_t^+(W_A) \geq \Phi_t^-(W_B) & \forall t \in \left[0, \frac{1}{\pi}\right] \\ \Phi_t^-(W_A) \leq \Phi_t^-(W_B) & \forall t \in \left[-\frac{1}{\pi}, 0\right] \end{cases} \tag{6.9}$$

$$\begin{cases} S_s^{\downarrow}(W_A) \geq S_s^{\downarrow}(W_B) & \forall s \geq 0 \\ S_s^{\uparrow}(W_A) \leq S_s^{\uparrow}(W_B) & \forall s \geq 0 \end{cases} \tag{6.10}$$

Remember that Φ_t^+ and Φ_t^- are the functions defined in Chapter 3, which are respectively built from the convex function γ_t^+ and concave function γ_t^- (see (3.13) and (3.73)).

Let us now interest ourselves to the particular case of Wigner-positive states. Indeed, in that case, the majorization condition can be simplified. The increasing rearrangement of a non-negative Wigner function is identically zero, so that we have $S_s^{\uparrow}(W) = 0$ for all $W \in \mathcal{W}_+$. In general, for $W_A, W_B \in \mathcal{W}_+$, the statement $W_A \succ W_B$ is equivalent to the following conditions:

$$\Phi_t^+(W_A) \geq \Phi_t^+(W_B) \qquad \forall t \in \left[0, \frac{1}{\pi}\right] \tag{6.11}$$

$$S_s^{\downarrow}(W_A) \geq S_s^{\downarrow}(s)(W_B) \qquad \forall s \geq 0 \tag{6.12}$$

6.1.4 Symplectic invariance

Let us see how the level-functions of the Wigner function evolve under the linear phase space transformations that we have described in Chapter 4. Let us first consider a symplectic transfromation described by the matrix $\mathbf{U}$. Its upper level-function can be derived as:

$$\begin{aligned} m_{W'}^+(t) &= \int \Theta\left(W'(\mathbf{r}') - t\right) \mathrm{d}\mathbf{r}' && (6.13) \\ &= \int \Theta\left(\frac{W(\mathbf{r})}{|\det \mathbf{U}|} - t\right) |\det \mathbf{U}| \mathrm{d}\mathbf{r} && (6.14) \\ &= \int \Theta\left(W(\mathbf{r}) - t\right) \mathrm{d}\mathbf{r} && (6.15) \\ &= m_W^+(t) && (6.16) \end{aligned}$$

where we have used the property the symplectic matrices have a determinant with absolute value equal to 1. A very similar development yields that the lower level-function m_W^- also remains unchanged under

symplectic transformations. This concludes that Wigner functions are level-equivalent to their result from the transformations $\mathcal{R}_\theta$, $\mathcal{D}_\alpha$ and $\mathcal{S}_r$. This result also extends to the conjugation transformation Π as the determinant of $\mathbf{\Pi}$ is equal to -1, so that its absolute value is 1.

$$\mathcal{R}_\theta[W] \equiv W, \qquad \mathcal{D}_\alpha[W] \equiv W, \qquad \mathcal{S}_z[W] \equiv W, \qquad \Pi[W] \equiv W. \tag{6.17}$$

6.2 Towards a state of least disorder in phase space

As we have explained, a majorization relation allows us to compare two distributions, in this case Wigner functions, in terms of disorder. It is thus a natural question to ask us whether there exists a Wigner function W_{sup} that majorizes every other Wigner function, and another Wigner function W_{inf} that is majorized by every other functions.

$$\exists?\quad W_{\text{sup}} \in \mathcal{W}: \qquad \forall W \in \mathcal{W}, \qquad W \prec W_{\text{sup}} \tag{6.18}$$

$$\exists?\quad W_{\text{inf}} \in \mathcal{W}: \qquad \forall W \in \mathcal{W}, \qquad W \succ W_{\text{inf}} \tag{6.19}$$

Let us examine proposition (6.19) first. Suppose there exists a state $W_{\text{inf}} \in \mathcal{W}$ which is majorized by every other Wigner function of $\mathcal{W}$. Using any of the symplectic transformations of (6.17), we can construct a different state W'_{inf} such that $W'_{\text{inf}} \equiv W_{\text{inf}}$. Then, we can construct a new state W_{new} as a convex combination of W_{inf} and W'_{inf}, so that $W_{\text{new}} = (W_{\text{inf}} + W'_{\text{inf}})/2$. At this point, we would like to highlight an important property that we will use later in this chapter. Let $W \succ W_A$ and $W \succ W_B$, then any convex combination of W_A and W_B is majorized by W:

$$\begin{cases} W \succ W_A \\ W \succ W_B \end{cases} \quad \Rightarrow \quad W \succ p_1 W_A + p_2 W_B \tag{6.20}$$

where (p_1, p_2) is a probability vector. That property is easily shown by exploting Jensen's inequality for convex and concave functions in relation (6.9). As a consequence of property (6.20), we see that $W_{\text{new}} \prec W_{\text{inf}}$, which shows that a state W_{inf} cannot exist. Note however that the proof is not complete, as it could happen that $W_{\text{new}} \equiv W_{\text{inf}}$. Nevertheless, it seems reasonable to argue that with an appropriate choice of symplectic transformation, it is always possible to build a state W'_{inf} such that $W_{\text{new}} = (W_{\text{inf}} + W'_{\text{inf}})/2 \not\equiv W_{\text{inf}}$.

Let us now come to proposition (6.18). Does there exist a Wigner function W that majorizes every other Wigner function ? The question seems natural, as Wigner functions are subject to the uncertainty principle. As a consequence, Wigner functions must contain some disorder because of their quantum nature.

We will see that there is in fact no unique state minimizing disorder from the point of view of majorization. This conclusion will follow from the observation that pure states are in general incomparable. We will then see that by restricting ourselves to Wigner-positive states it is possible to formulate a minimal disorder conjecture.

6.2.1 Pure states are in general incomparable

Pure states are states without any statistical uncertainty. For that reason, they appear as the most natural candidates to minimize disorder in phase space. Indeed, remember that purity corresponds in phase-space to a Schur-convex function:

$$\mu(W) = 2\pi \iint (W(x,p))^2 \,\mathrm{d}x\mathrm{d}p. \tag{6.21}$$

As a consequence, the purity should be maximized for a state of least disorder, and it should then be equal to 1.

Building Schur-convex functionals

In this subsection, we make a small digression to present a general way to build Schur-convex functionals. We know that the set of function $S_s^{\downarrow}$ and $S_s^{\uparrow}$ are respectively Schur-convex and Schur-concave. As a consequence, any convex combination of $S_s^{\downarrow}$ is also Schur-convex, and any convex combination of $S_s^{\uparrow}$ is Schur-concave. Let us imagine a distribution $k(s) : \mathbb{R}_+ \mapsto \mathbb{R}_+$ which is always non-negative. Since a sum of Schur-convex functional is Schur-convex, it appears that the following functional $\Phi(W)$ is Schur-convex:

$$\Phi(W) = \int_0^\infty \mathrm{d}s k(s) S_s^{\downarrow}(W) \tag{6.22}$$

$$= \int_0^\infty \mathrm{d}s k(s) \int_0^s 2\pi r \mathrm{d}r W^{\downarrow}(r) \tag{6.23}$$

$$= \int_0^\infty \mathrm{d}s \int_0^\infty 2\pi r \mathrm{d}r \quad \Theta(s-r)\, k(s) W^{\downarrow}(r) \tag{6.24}$$

$$= \int_0^\infty 2\pi r \mathrm{d}r \; W^{\downarrow}(r) \int_0^\infty \mathrm{d}s \; \Theta(s-r)\, k(s) \tag{6.25}$$

$$= \int_0^\infty 2\pi r \mathrm{d}r W^{\downarrow}(r) \int_r^\infty k(s) \mathrm{d}s \tag{6.26}$$

$$= \int_0^\infty 2\pi r \mathrm{d}r \; W^{\downarrow}(r) \tilde{k}(r) \tag{6.27}$$

Remember that the only assumption we have made on k is that it is always non-negative. As a consequence, the only assumption that we have to make on $\tilde{k}$ is that it is non-negative and non-increasing. What we have shown here is that the functional of the form $\Phi(W) = \int W^{\downarrow}(r)\varphi(r)2\pi r \mathrm{d}r$ is Schur-convex provided that $\varphi : \mathbb{R}_+ \mapsto \mathbb{R}$ is non-negative and non-increasing. Conversely, a similar development yields the conclusion any functional of the form $\Phi(W) = \int W^{\uparrow}(r)\varphi(r)2\pi r \mathrm{d}r$ is Schur-convex provided that $\varphi : \mathbb{R}_+ \mapsto \mathbb{R}_-$ is non-positive and non-increasing. As a consequence, any functional Φ built as follows is Schur-convex:

$$\Phi(W) = \int W^{\downarrow}(r)\varphi_1(r)2\pi r \mathrm{d}r + \int W^{\uparrow}(r)\varphi_2(r)2\pi r \mathrm{d}r \tag{6.28}$$

under the condition that $\varphi_1 : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is non-negative non-increasing and $\varphi_2 = \mathbb{R}_+ \mapsto \mathbb{R}_-$ is non-positive non-decreasing. Remember now that the decreasing rearrangement of a distribution is itself described by a non-increasing distribution, while the increasing rearrangement is described by a non-decreasing distribution.

Schur-convex functional associated to a pure state

We are now going to use the previous observation to define a Schur-convex function associated to any pure state. Let $W \in \mathcal{W}^*$ be the Wigner function of a pure state. We define the associated function

$\Phi_W : \mathbb{R}^{\mathbb{R}^2} \mapsto \mathbb{R}$ as follows:

$$\Phi_W(A) = 2\pi \int \left(W^{\downarrow}(r) A^{\downarrow}(r) 2\pi r \mathrm{d}r + W^{\uparrow}(r) A^{\uparrow}(r) 2\pi r \mathrm{d}r \right) \tag{6.29}$$

$$= 2\pi \left(\iint W^{\downarrow}(x,p) A^{\downarrow}(x,p) \mathrm{d}x\mathrm{d}p + \iint W^{\uparrow}(x,p) A^{\uparrow}(x,p) \mathrm{d}x\mathrm{d}p \right) \tag{6.30}$$

$$= 2\pi \left(\int\limits_{\mathbb{R}^2} W^{\downarrow}(\mathbf{q}) A^{\downarrow}(\mathbf{q}) \mathrm{d}\mathbf{q} + \int\limits_{\mathbb{R}^2} W^{\uparrow}(\mathbf{q}) A^{\uparrow}(\mathbf{q}) \mathrm{d}\mathbf{q} \right) \tag{6.31}$$

$$= 2\pi \left(\int\limits_{\mathbb{R}^2} \left(W^{\downarrow} A^{\downarrow} \right)(\mathbf{q}) \mathrm{d}\mathbf{q} + \int\limits_{\mathbb{R}^2} \left(W^{\uparrow} A^{\uparrow} \right)(\mathbf{q}) \mathrm{d}\mathbf{q} \right) \tag{6.32}$$

$$= 2\pi \int\limits_{\mathbb{R}^2 \oplus \mathbb{R}^2} \left(W^{\downarrow} A^{\downarrow} \oplus W^{\uparrow} A^{\uparrow} \right)(\mathbf{q}) \mathrm{d}\mathbf{q} \tag{6.33}$$

$$= 2\pi \int\limits_{\mathbb{R}^2 \oplus \mathbb{R}^2} \left(W^{\downarrow} \oplus W^{\uparrow} \right) \left(A^{\downarrow} \oplus A^{\uparrow} \right)(\mathbf{q}) \mathrm{d}\mathbf{q} \tag{6.34}$$

We introduce the notation $W^{\updownarrow} = W^{\downarrow} \oplus W^{\uparrow}$, and the scalar product:

$$\langle W_A, W_B \rangle = 2\pi \iint W_A(x,p) W_B(x,p) \mathrm{d}x\mathrm{d}p. \tag{6.35}$$

We can then simply write $\Phi_W(W_A) = \left\langle W^{\downarrow}, W_A^{\downarrow} \right\rangle + \left\langle W^{\uparrow}, W_A^{\uparrow} \right\rangle = \left\langle W^{\updownarrow}, W_A^{\updownarrow} \right\rangle$. Notice that the function Φ_W benefits of interesting properties:

- Linear

  The functional Φ_W is linear for non-negative scalars:

  $$\Phi_W(k \cdot W_A) = k \cdot \Phi_W(W_A) \qquad \forall k \in \mathbb{R}_+ \tag{6.36}$$

  where $k \in \mathbb{R}_+$.

- Schur-convex

  This is a direct consequence of the construction of $\Phi_W(W_A)$. We have the relation:

  $$W_A \succ W_B \qquad \Rightarrow \qquad \Phi_W(W_A) \geq \Phi_W(W_B) \tag{6.37}$$

- Definite positive

  From its construction, the functional $\Phi_W(A)$ cannot take negative values. Moreover, it is zero only if the distribution A is zero everywhere:

  $$\Phi_W(A) \geq 0, \qquad \Phi_W(A) = 0 \quad \Leftrightarrow \quad A = 0 \tag{6.38}$$

- Definite upper-bounded

  It directly follows from Cauchy-Schwarz inequality that it is upper-bounded. Indeed, we can write $\left\langle A^{\updownarrow}, B^{\updownarrow} \right\rangle^2 \leq \left\langle A^{\updownarrow}, A^{\updownarrow} \right\rangle \left\langle B^{\updownarrow}, B^{\updownarrow} \right\rangle$. Moreover, we have the inquality $\left\langle W^{\updownarrow}, W^{\updownarrow} \right\rangle = \langle W, W \rangle = \mu(W)$. From that, we can conclude that $\Phi_W(A) \leq 1$. Also, Cauchy-Schwarz inequality is saturated only for proportional distributions, which means that Φ_W takes the value 1 only for Wigner-functions level-equivalent to W.

  $$\Phi_W(W_A) \leq 1, \qquad \Phi_W(W_A) = 1 \quad \Leftrightarrow \quad W_A \equiv W \tag{6.39}$$

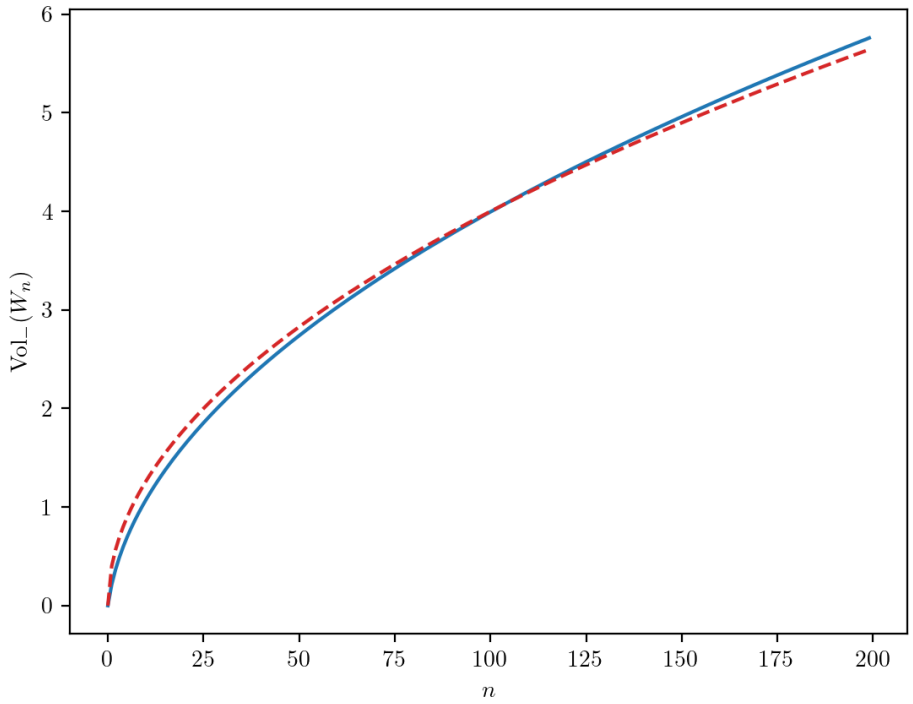

Figure 6.2: The blue curve represents the negative volume of Fock states $\mathrm{Vol}_-(W_n)$ as a function of n. The red dashed curve corresponds to the function $0.4\sqrt{n}$, which appears to be a good approximation of $\mathrm{Vol}_-(W_n)$. Note that this graph can be found almost identically in reference [39].

A direct consequence of the existence of the Schur-convex function Φ_W for any pure state $W \in \mathcal{W}^*$ is that pure states are in general incomparable. Indeed, let $W_A, W_B \in \mathcal{W}^*$. Then the Schur-convex functions $\Phi_{W_A}(W_A) = 1 \geq \Phi_{W_A}(W_B)$ and $\Phi_{W_B}(W_B) = 1 \geq \Phi_{W_B}(W_A)$. For W_A to majorizes W_B, it must that any Schur-convex function takes a greater or equal value on W_A than on W_B. The only possiblity to do so is that $\Phi_{W_B}(W_A) = 1$, in which case it must that $W_A \equiv W_B$. As a consequence, the Wigner function of any two single-mode pure states are either incomparable either level-equivalent.

$$\forall W_A, W_B \in \mathcal{W}^* : \qquad W_A \equiv W_B \quad \text{or} \quad W_A \not\gtrless W_B \tag{6.40}$$

where we use the notation $W_A \not\gtrless W_B$ to say that W_A and W_B are incomparable.

6.2.2 The negative volume of Wigner functions is unbounded

In this subsection, we present an interesting observation over the maximum negative volume that Wigner functions can exhibit. As it can be seen, the negative volume Vol_- corresponds to the functional $-\Phi_t^-$ for $t = 0$, so that it is a Schur-convex functional. As such, it should take greater values for a states of low disorder.

It appears that the negative volume is unbounded and can take arbitrarily large values. For example, let us consider the set of Fock states. Figure 6.2 plots the negative volume of Fock states. We observe that it is strictly increasing as a function of n, and that it is unbounded:

$$\lim_{n\to\infty} \mathrm{Vol}_-(W_n) = \infty \tag{6.41}$$

The fact that the negative volume is unbounded can be understood as another proof that there exists no state W_{sup} of least disorder in phase space. It is impossible to find a Wigner function that majorizes every other functions, since it will always possible to find a Wigner function with a greater negative volume. It is thus vain to look for a Wigner function which majorizes every other Wigner functions. However, the problem becomes interesting if we restrict ourselves to Wigner functions whose negative volume is bounded. It is in this perspective that we introduce the following conjecture.

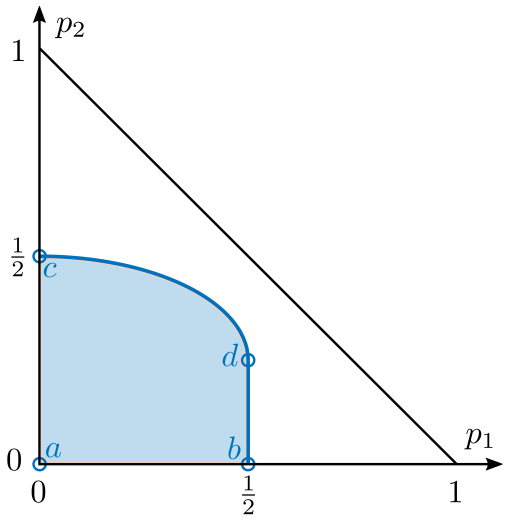

Figure 6.3: Two-dimensional representation (white triangle) of quantum states of the form (6.44). The blue region included in the triangle corresponds to the Wigner-positive states. The points a, b, c, and d are associated with the Wigner functions W_a, W_b, W_c, and W_d, while the points in the segment of an ellipse connecting c and d have a Wigner function V_t where t takes any value between 0 and 1.

6.2.3 Disorder conjecture over the Wigner-positive set

The set of non-negative Wigner functions $\mathcal{W}_+$ appear as an appropriate subset to lay a least disorder conjecture. Indeed, the negative volume of Wigner-positive states is bounded as it is zero. Pure states then appear as the natural candidates to minimize the disorder. As a consequence of Hudson's theorem, $\mathcal{W}_+^*$ only contains Gaussian pure states and actually contains all of them. Remarkably, the Wigner functions of all Gaussian pure states are level-equivalent since they are all related by symplectic transformations in phase space. Indeed, we have the relation:

$$\forall W_A, W_B \in \mathcal{W}_+^* : \qquad W_A \equiv W_B \tag{6.42}$$

Hence, all Gaussian pure states have a Wigner function that is level-equivalent to W_0, making them all equivalent to W_0 from the point of view of majorization, namely $W_0 \equiv W$ for all $W \in \mathcal{W}_+^*$. With this in mind, we state the following conjecture:

Conjecture 6.1 (Continuous majorization for Wigner-positive states)**.** *The Wigner function of a Wigner-positive state is majorized by the Wigner function of vacuum.*

$$W \prec W_0 \qquad\qquad \forall W \in \mathcal{W}_+ \tag{6.43}$$

This expresses that, in the sense of majorization theory, the most fundamental (Wigner-positive) state is the vacuum state, *i.e.*, the ground state of the Hamiltonian of the harmonic oscillator. Note that Conjecture 6.1 goes beyond the scope of quantum optical states and applies to the phase space associated with any canonical pair (x, p). Furthermore, it is unrelated to the Hamiltonian of the system : the (positive) Wigner function of any state of the system must always be majorized by the Wigner function of a pure Gaussian state.

6.3 Restricted proof of the conjecture

In this section, we make a first step towards solving conjecture (6.1) by considering a particular subset of quantum states, namely phase-invariant states that are restricted to two photons at most,

$$\hat{\rho} = (1 - p_1 - p_2) \left|0\right\rangle \left\langle 0\right| + p_1 \left|1\right\rangle \left\langle 1\right| + p_2 \left|2\right\rangle \left\langle 2\right|, \tag{6.44}$$

where $p_1, p_2 \geq 0$ and $p_1 + p_2 \leq 1$. These states form a convex set given by the area whose outer boundaries are the p_1-axis, the p_2-axis and the line satisfying $p_1 + p_2 = 1$ as pictured on Fig. 6.3. We will prove Conjecture 6.1 for the subset of Wigner-positive states of the form (6.44), denoted as $\mathcal{W}_+^2$, which corresponds to the blue area in Figure 6.3. That set has previously been studied in Chapter 5. It is obvious to see that $\mathcal{W}_+^2$ forms a convex set as well since any convex mixture of Wigner-positive states is Wigner-positive, but the boundary of this set is nonetheless non trivial, see Figure 6.3. At the same time, this set is simple enough to enable a fully analytical proof of Conjecture 6.1.

As shown in Chapter 5, the boundary of the set $\mathcal{W}_+^2$ comprises the extremal states $\hat{\rho}_a$, $\hat{\rho}_b$, $\hat{\rho}_c$ and $\hat{\rho}_d$, represented by the corresponding letters in Figure 6.3, as well as the segment of an ellipse connecting $\hat{\rho}_c$ to $\hat{\rho}_d$. Thus, any state in $\mathcal{W}_+^2$ can be written as a convex mixture of these extremal states. Note that $\hat{\rho}_a = |0\rangle\langle 0|$, which lies at the origin in Figure 6.3, is simply the vacuum state which will be proven to majorize every other state. The expressions of the Wigner functions of the first four extremal states (as a function of the parameter r) read as follows (see Chapter 5):

$$\begin{aligned} W_a(r) &= W_0(r) = \frac{1}{\pi}\exp\left(-r^2\right), \\ W_b(r) &= \frac{1}{\pi}\exp\left(-r^2\right) r^2, \\ W_c(r) &= \frac{1}{\pi}\exp\left(-r^2\right)\left(r^2-1\right)^2, \\ W_d(r) &= \frac{1}{\pi}\exp\left(-r^2\right)\frac{1}{2}r^4. \end{aligned} \tag{6.45}$$

In addition to these, there is a continuum of extremal states located on the segment of an ellipse connecting point c to point d in Figure 6.3. Using a parameter $t \in [0,1]$, the Wigner function of these states can be parametrized as follows (see Chapter 5):

$$V_t(r) = \frac{t+1}{2\pi}\exp\left(-r^2\right)\left(r^2 - 1 + \sqrt{\frac{1-t}{1+t}}\right)^2. \tag{6.46}$$

Note that for $t = 0$, V_t coincides with W_d, while for $t = 1$, it coincides with W_c.

We are now going to prove that Conjecture 6.1 holds for all states contained in the convex set $\mathcal{W}_+^2$. To do so, it is sufficient to prove that the Wigner functions of all the extremal states are majorized by the Wigner function of the vacuum W_0. As a consequence of Equation (6.20), this will indeed automatically imply that the same majorization relation holds for all convex mixtures of extremal states, hence for all states in $\mathcal{W}_+^2$. In order to prove our result, we begin by showing that a majorization relation on radial functions in $\mathbb{R}^n$ (here, we only need $n = 2$) is equivalent to a majorization relation on specific functions defined on the non-negative real line. This is the content the following lemma, which we prove in Appendix A.

Lemma 1. *If f and g are two n-dimensional radial distributions defined on $\mathbb{R}^n$ such that $f(\mathbf{r}) = f_{\mathrm{R}}(\|\mathbf{r}\|)$ and $g(\mathbf{r}) = g_{\mathrm{R}}(\|\mathbf{r}\|)$ with f_{R} and g_{R} defined on $\mathbb{R}_+$, then $f \succ g$ is equivalent to $\tilde{f} \succ \tilde{g}$, where $\tilde{f}$ and $\tilde{g}$ are 1-dimensional distributions defined on $\mathbb{R}_+$ as $\tilde{f}(x) = f_{\mathrm{R}}(\sqrt[n]{x})$ and $\tilde{g}(x) = g_{\mathrm{R}}(\sqrt[n]{x})$.*

Lemma 1 implies that a majorization relation between any two Wigner functions picked from W_0, W_b, W_c, W_d and V_t is equivalent to a majorization relation between the corresponding 1-dimensional functions picked from f_0, f_b, f_c, f_d and g_t, which are defined on $\mathbb{R}_+$ as

$$\begin{aligned} f_0(x) &= \exp(-x), \\ f_b(x) &= \exp(-x)\, x, \\ f_c(x) &= \exp(-x)\,(x-1)^2, \\ f_d(x) &= \exp(-x)\frac{1}{2}x^2, \end{aligned} \tag{6.47}$$

and

$$g_t(x) = \exp(-x)\frac{1}{2}(t+1)\left(x - 1 + \sqrt{\frac{1-t}{1+t}}\right)^2. \tag{6.48}$$

Thus, we need to prove now that f_0 majorizes f_b, f_c, f_d, and g_t. Our proof relies on the following lemma, which we prove in Appendix A for completeness, as we could not find it in the literature.

Lemma 2. *Consider two probability distributions f and g defined on the same domain $\mathbb{A}$. If there exists a collection of level-equivalent distributions $f^{(\alpha)}$ on $\mathbb{A}$ depending on the parameter α with $f^{(\alpha)} \equiv f$ for all α such that*

$$g(\mathbf{r}) = \int k(\alpha)\, f^{(\alpha)}(\mathbf{r})\, \mathrm{d}\alpha, \quad \forall\, \mathbf{r} \in \mathbb{A}, \tag{6.49}$$

where k is a probability density distribution, then $f \succ g$.

Lemma 2 enables us to prove that $f \succ g$, provided that we can build g as some convex mixture of distributions that are level-equivalent to f.

Case of f_b and f_d

Let us first prove that f_0 majorizes f_b and f_d. In order to make use of Lemma 2, we are going to build an appropriate collection of level-equivalent functions to f_0. One simple way to generate level-equivalent functions is simply by shifting the original function to the right. Starting from f_0, we define the functions $f_0^{(\alpha)}$ labelled by the non-negative shift parameter α as

$$f_0^{(\alpha)}(x) = \exp(-x + \alpha)\, \Theta(x - \alpha), \tag{6.50}$$

where $\Theta(z)$ represents the Heaviside step function. We obviously have that $f_0^{(\alpha)} \equiv f_0$ for all $\alpha \geq 0$. Now, define the probability densities $k_b(\alpha) = \exp(-\alpha)$ and $k_d(\alpha) = \alpha \exp(-\alpha)$, with $\alpha \in \mathbb{R}_+$. It is trivial to verify that $k_{b(d)}(\alpha) \geq 0$ for all $\alpha \in \mathbb{R}_+$ and $\int k_{b(d)}(\alpha)\, \mathrm{d}\alpha = 1$. Furthermore, it can easily be shown that

$$f_{b(d)}(x) = \int_0^{+\infty} k_{b(d)}(\alpha)\, f_0^{(\alpha)}(x)\, \mathrm{d}\alpha. \tag{6.51}$$

Lemma 2 then directly implies that $f_0 \succ f_b$ and $f_0 \succ f_d$.

Case of f_c and g_t

The same method is not directly applicable to prove that f_0 majorizes f_c and g_t because the latter functions are non-zero at the origin. The trick, however, is to exploit the fact that f_c looks like a rescaled version of f_d in the domain $[1, \infty)$. We can then "split" f_c into two parts and prove the majorization relation separately for each part. This is possible as a consequence of the following lemma, which we prove in Appendix A.

Lemma 3. *Consider four functions f_1, f_2, g_1, and g_2 defined on the same domain $\mathbb{A}$ and such that f_1 and f_2 do not both take non-zero values in the same element of $\mathbb{A}$, and similarly g_1 and g_2 do not both take non-zero values in the same element of $\mathbb{A}$. If the functions satisfy $f_1 \succ g_1$ and $f_2 \succ g_2$, then $(f_1 + f_2) \succ (g_1 + g_2)$.*

In light of Lemma 3, define the two functions f_c^- and f_c^+ on $\mathbb{R}_+$ as

$$f_c^-(x) = \begin{cases} f_c(x), & \text{for } 0 \leq x \leq 1, \\ 0, & \text{else}, \end{cases} \tag{6.52}$$

and

$$f_c^+(x) = \begin{cases} 0, & \text{for } 0 \leq x \leq 1, \\ f_c(x), & \text{else}. \end{cases} \tag{6.53}$$

Obviously, we have $f_c^- + f_c^+ = f_c$. In order to prove that $f_0 \succ f_c$ by using Lemma 3, we also need to "split" f_0 into two parts f_0^- and f_0^+ such that $f_0^- + f_0^+ = f_0$. Moreover, in order to be able to apply

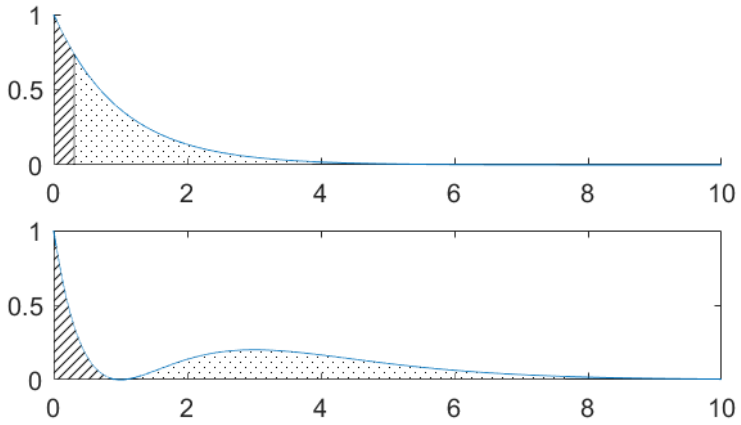

Figure 6.4: Illustration of the "split" of f_0 (top) and f_c (bottom) into two parts. The hashed parts are f_0^- and f_c^-, while the dotted parts are f_0^+ and f_c^+. The two hashed areas are equal, while the two dotted areas are equal. This allows us to treat $f_0^- \succ f_c^-$ and $f_0^+ \succ f_c^+$ separately, in order to conclude finally that $f_0 \succ f_c$.

majorization on each part, f_0^- must have the same normalization as f_c^-, and similarly for f_0^+ and f_c^+. Define $x^* = 1 - \ln 2$, and note that

$$\begin{aligned} \int_0^1 f_c(x)\mathrm{d}x &= \int_0^{x^*} f_0(x)\mathrm{d}x, \\ \int_1^\infty f_c(x)\mathrm{d}x &= \int_{x^*}^\infty f_0(x)\mathrm{d}x. \end{aligned} \tag{6.54}$$

With this in mind, the functions f_0^- and f_0^+ on $\mathbb{R}_+$ are

$$f_0^-(x) = \begin{cases} f_0(x), & \text{for } 0 \le x \le x^*, \\ 0, & \text{else}, \end{cases} \tag{6.55}$$

and

$$f_0^+(x) = \begin{cases} 0, & \text{for } 0 \le x \le x^*, \\ f_0(x), & \text{else}. \end{cases} \tag{6.56}$$

It follows from (6.54) that f_c^- and f_0^- have the same normalization (and similarly for f_c^+ and f_0^+). The distributions f_0^-, f_0^+, f_c^-, and f_c^+ are represented in Fig. 6.4.

The last step now is to prove that $f_0^- \succ f_c^-$ as well as $f_0^+ \succ f_c^+$. Starting with the latter relation, we define the two functions $\tilde{f}_0^+$ and $\tilde{f}_c^+$ on $\mathbb{R}_+$ by respectively shifting to the left f_0^+ by an amount x^* and f_c^+ by 1, namely

$$\tilde{f}_0^+(x) = f_0^+(x + 1 - \ln 2) = 2\exp(-1)\exp(-x), \tag{6.57}$$

and

$$\tilde{f}_c^+(x) = f_c^+(x + 1) = \exp(-1)\exp(-x)x^2. \tag{6.58}$$

Note that $\tilde{f}_0^+(x)$ and $\tilde{f}_c^+(x)$ are proportional to $f_0(x)$ and $f_d(x)$, respectively, with the same proportionality factor of $2\exp(-1)$. Since we have already shown that $f_0 \succ f_d$, it follows that $\tilde{f}_0^+ \succ \tilde{f}_c^+$, which is equivalent to $f_0^+ \succ f_c^+$ since $\tilde{f}_0^+ \equiv f_0^+$ and $\tilde{f}_c^+ \equiv f_c^+$.

In order to prove that $f_0^- \succ f_c^-$, we note that f_0^- and f_c^- are both monotonically decreasing functions, so they coincide with their decreasing rearrangements, namely $f_0^- = (f_0^-)^\downarrow$ and $f_c^- = (f_c^-)^\downarrow$.

Therefore, their cumulative integrals are simply given by

$$S_s(f_0^-) = \int_0^s f_0^-(x)\,\mathrm{d}x \quad \text{and} \quad S_s(f_c^-) = \int_0^s f_c^-(x)\,\mathrm{d}x. \tag{6.59}$$

In order to prove the majorization relation, we will now show that $S_s(f_0^-) \geq S_s(f_c^-)$ for all $s \in \mathbb{R}_+$. Since f_0^- and f_c^- are both monotonically decreasing functions, since $f_0^-(x) = 0$ for all $x > x^*$, and since $x^* < 1$, it is sufficient to show that $f_0^-(x) \geq f_c^-(x)$ for all $x \in [0, x^*]$. Indeed, the ratio $f_c(x)/f_0(x) = (x-1)^2 \leq 1$ for $x \in [0, x^*]$. Since f_0^- and f_c^- are non-negative distributions normalized to the same value, we get $f_0^- \succ f_c^-$. From Lemma 3, we conclude that $f_0 \succ f_c$.

Finally, the same "splitting" technique can be used to prove that $f_0 \succ g_t$ for all values of $t \in [0,1]$, which of course includes $f_0 \succ f_d$ and $f_0 \succ f_c$ as limiting cases for $t = 0$ and 1, respectively. We point the interested reader to Appendix B for such a proof. In summary, we have thus shown that all functions (i.e., f_b, f_c, f_d and g_t for all $t \in [0,1]$) are majorized by f_0.

Using Lemma 1, this translates into the fact that the Wigner functions of all extremal states (i.e., W_b, W_c, W_d, and V_t for all $t \in [0,1]$) are majorized by W_0. Hence, any convex mixture of these extremal Wigner functions is also majorized by W_0 as a consequence of Eq. (6.20). This concludes the proof of Conjecture (6.1) for all Wigner-positive states in $\mathcal{W}_+^2$.

6.4 Alternative continuous majorizations

Let us conclude this chapter by defining two adapted versions of continuous majorization. These are slightly different to the definition of a regular majorization relation. We first present radial-majorization, and show its particular connection with energy. Then we present square-majorization, which enables us to formulate a conjecture over the set of pure states.

6.4.1 Radial-majorization

Any single-mode Wigner function $W(x,p)$ can be equivalently expressed in a polar coordinate system $W(r,\theta)$, such that $r = \sqrt{x^2+p^2}$ and $\theta = \arctan(p/x)$. The cumulative integral S_s of a Wigner function can then be expressed as follows:

$$S_s(W) = \iint_{\|(x,p)\| \leq s} W(x,p)\mathrm{d}x\mathrm{d}p = \int_0^s \mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta\; W(r,\theta) \tag{6.60}$$

We say that the Wigner function W_A radial-majorizes the Wigner function W_B, written $W_A \succ_\mathrm{r} W_B$ if and only if the following condition holds:

$$S_s(W_A) \geq S_s(W_B) \qquad \forall s \geq 0. \tag{6.61}$$

Radial-majorization can be understood as the continuous analog of Fock-majorization [37]. Indeed, Fock-majorization corresponds to an adapted version of discrete majorization between two density operators, where the eigenvalues are sorted by inceasing energy, rather than decreasing order. Indeed, rememebr that the radius $r = \sqrt{x^2+p^2}$ can be related to the energy of the state as $\hat{H} = \left(\hat{x}^2 + \hat{p}^2\right)/2$ for bosonic systems. With radial-majorization, we use the cumulative integral which perform the integration starting with the lowest values of r, and thus by increasing energy.

Let us now consider that $W_A \succ_\mathrm{r} W_B$. As a consequence, we have the relation $S_s(W_A) \geq S_s(W_B)$

for all $s \geq 0$. If we now introduce a non-negative function $k(s)$, we can write the following implication:

$$\int_0^\infty \mathrm{d}s k(s) S_s(W_A) \geq \int_0^\infty \mathrm{d}s k(s) S_s(W_B) \tag{6.62}$$

$$\Leftrightarrow \quad \int_0^\infty \mathrm{d}s k(s) \int_0^s r\mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta W_A(r,\theta) \geq \int_0^\infty \mathrm{d}s k(s) \int_0^s r\mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta W_B(r,\theta) \tag{6.63}$$

$$\Leftrightarrow \quad \int_0^\infty \mathrm{d}s k(s) \int_0^\infty r\mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta \Theta(s-r) W_A(r,\theta) \geq \int_0^\infty \mathrm{d}s k(s) \int_0^\infty r\mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta \Theta(s-r) W_B(r,\theta) \tag{6.64}$$

$$\Leftrightarrow \quad \int_0^\infty r\mathrm{d}r \int_0^\infty \mathrm{d}s k(s) \Theta(s-r) \int_0^{2\pi} \mathrm{d}\theta W_A(r,\theta) \geq \int_0^\infty r\mathrm{d}r \int_0^\infty \mathrm{d}s k(s) \Theta(s-r) \int_0^{2\pi} \mathrm{d}\theta W_B(r,\theta) \tag{6.65}$$

$$\Leftrightarrow \quad \int_0^\infty r\mathrm{d}r \underbrace{\int_r^\infty \mathrm{d}s k(s)}_{\tilde{k}(r)} \int_0^{2\pi} \mathrm{d}\theta W_A(r,\theta) \geq \int_0^\infty r\mathrm{d}r \underbrace{\int_r^\infty \mathrm{d}s k(s)}_{\tilde{k}(r)} \int_0^{2\pi} \mathrm{d}\theta W_B(r,\theta) \tag{6.66}$$

$$\Leftrightarrow \quad \int_0^\infty r\mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta \tilde{k}(r) W_A(r,\theta) \geq \int_0^\infty r\mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta \tilde{k}(r) W_B(r,\theta) \tag{6.67}$$

Remember that the only condition we have laid on $k(s)$ is that it should be non-negative. For that reason, the only condition that applies on $\tilde{k}(r) = \int_r^\infty k(r)\mathrm{d}r$ is that it is non-negative and non-increasing. Note also that the non-negativity of $\tilde{k}(r)$ can be relaxed, since adding or subtracting a constant on both sides of the inequality gives an equivalent inequality. One last thing to note is that if choosing $\tilde{k}(r)$ to be $\Theta(s-r)$, the condition yields the inequality $S_s(W_A) \geq S_s(W_B)$. We can thus write the following equivalence:

$$W_A \succ_\mathrm{r} W_B \quad \Leftrightarrow \quad \int_0^\infty \int_0^{2\pi} W_A(r,\theta)\varphi(r)\mathrm{d}\theta r\mathrm{d}r \geq \int_0^\infty \int_0^{2\pi} W_B(r,\theta)\varphi(r)\mathrm{d}\theta r\mathrm{d}r \qquad \forall \varphi \text{ non-increasing} \tag{6.68}$$

$$\Leftrightarrow \quad \int_0^\infty \int_0^{2\pi} W_A(r,\theta)\varphi(r)\mathrm{d}\theta r\mathrm{d}r \leq \int_0^\infty \int_0^{2\pi} W_B(r,\theta)\varphi(r)\mathrm{d}\theta r\mathrm{d}r \qquad \forall \varphi \text{ non-decreasing} \tag{6.69}$$

Remember now that the bosonic Hamiltonian is $\hat{H} = (\hat{x}^2+\hat{p}^2)/2$. It is associated to the distribution $\mathcal{T}_\mathrm{w}\left[\hat{H}\right] = H(x,p) = (x^2+p^2)/2 = r^2/2$ where we use the parameter $r = \sqrt{x^2+p^2}$. As a consequence, we can write:

$$W_A \succ_\mathrm{r} W_B \quad \Rightarrow \quad \langle \hat{H} \rangle_A \leq \langle \hat{H} \rangle_B \tag{6.70}$$

Radial-majorization is expected to play an important role in quantum thermodynamics with continuous variables. Let us mention that we have good reasons to believe that a relation of radial-majorization between two quantum states is preserved when the states evolves through Gaussian phase-insensitive channels (which we will introduce in Chapter 8). This way is still under research for the moment.

6.4.2 Square-majorization

The next alternative version of majorization that we present is square-majorization. In this different version, we consider the square of the Wigner function rather than the Wigner function itself, which allows us to consider only non-negative distributions. The origin of this idea is that, for states of same purity, the squared Wigner functions are normalized to the same value. Hence, square-majorization is a perfect tool to compare states with same purity. Let $W_A, W_B \in \mathcal{W}^*$ be the Wigner functions of two pure states. We say that W_A square-majorizes W_B if the following holds:

$$W_A \succ_{\mathrm{s}} W_B \quad \Leftrightarrow \quad (W_A)^2 \succ (W_B)^2 \tag{6.71}$$

$$\Leftrightarrow \quad \iint \varphi\left(W_A(x,p)^2\right) \mathrm{d}x\mathrm{d}p \geq \iint \varphi\left(W_B(x,p)^2\right) \mathrm{d}x\mathrm{d}p \qquad \forall \varphi \text{ convex} \tag{6.72}$$

Square-majorization has the advantage of providing a way to compare pure states. Indeed, we have seen earlier that pure state were in general incomparable as regards regular majorization. This is not true anymore with square-majorization. Indeed, our numerical simulations have shown for example the existence of a square-majorization chain for Fock states:

$$W_n \succ_{\mathrm{s}} W_{n+1} \tag{6.73}$$

where W_n is the Wigner function of the n^{th} Fock state. This is consistent with the fact that Fock states are increasingly widespread throughout phase-space for increasing values of n. In addition to this chain of square-majorization for Fock states, we have also found numerical evidence that the Wigner function of vacuum square-majorizes the Wigner function of any pure state. This leads us to state the following conjecture:

Conjecture 6.2 (Square-majorization for pure states)**.** *The Wigner function of any pure state is square-majorized by the Wigner function of a pure Gaussian state.*

$$W \prec_{\mathrm{s}} W_0 \qquad \forall W \in \mathcal{W}^* \tag{6.74}$$

Note that, similarly to regular majorization, pure Gaussian states are equivalent regarding square-majorization. We will show some numerical evidence supporting this conjecture in the next chapter.

Chapter 7

Measuring phase-space uncertainty

In the previous chapter, we laid the foundation for continuous majorization in phase space. As we know, a majorization relation is a very strong statement as it implies in turn an infinite number of inequalities over the broad class of Schur-convex and -concave functions. Majorization theory, however has its limits, since in some cases it cannot decide between two distributions, in which case they are said to be incomparable. In addition to that, a majorization relation never allows to quantify disorder, it is only a comparison between two distributions. It is well known that in physics in general, quantifying a quantity is an essential step in order to be able to establish equations that can be used in practice. This is the context for this chapter, and we will now turn to particular measures of disorder and study their properties.

First, we will focus on Schur-concave measures that are compatible with Wigner-negative states. We will indeed notice that some measures are adapted to measure the uncertainty of non-negative distributions, but become inadequate when we consider instead distributions taking both positive and negative values. We will consider in that section p-norms and Rényi entropies.

Then, we will devote a section of this chapter to the definition of a quantity which appears very naturally in the framework of Wigner-positive states, namely the Shannon differential entropy. We define that quantity as the Wigner entropy of a Wigner-positive state. This section will therefore echo Chapter 5, and we will identify the properties of the Wigner entropy that make it a remarkable physical quantity and worthy of particular interest. We will also present a conjecture on a very natural lower bound for the Wigner entropy. Then, we will present an important result of our work, namely an analytical proof that this lower bound is satisfied for the large set of Wigner-positive passive states. The main part of this chapter has been published in the paper entitled *Quantum Wigner entropy* [64].

Then, in the next section, we will extend our considerations by defining the Wigner-Renyi entropy. We will use these measures to illustrate the validity of the conjectures introduced in Chapter 6. To do so, we will numerically and randomly simulate quantum states and measure their respective Wigner-Rényi entropy.

7.1 Building a measure of uncertainty in phase space

The measures of disorder that interest us in this work are the measures compatible with the theory of majorization. These measure are therefore the Schur-convex or Schur-concave functionals. Such functionals can be symmetric, so that their value shouldn't change under transformation that keep the level-function unchanged. A simple way to construct a symmetric functional over phase-space is to integrate a function $\varphi : \mathbb{R} \mapsto \mathbb{R}$ over the whole domain of the Wigner function. With this in mind, we introduce the functions φ_p and functional Φ_p as follows:

$$\Phi_p(W) = \iint \varphi_p(W(x,p))\mathrm{d}x\mathrm{d}p, \qquad \text{where} \quad \varphi_p(x) = |x|^p. \tag{7.1}$$

Let us observe the graph of the function $\varphi(x) = |x|^p$. When $p \geq 1$, the function is convex over $\mathbb{R}$. When $0 \leq p < 1$, the function is concave over $\mathbb{R}_+$ and $\mathbb{R}_-$, but not over $\mathbb{R}$. This is illustrated

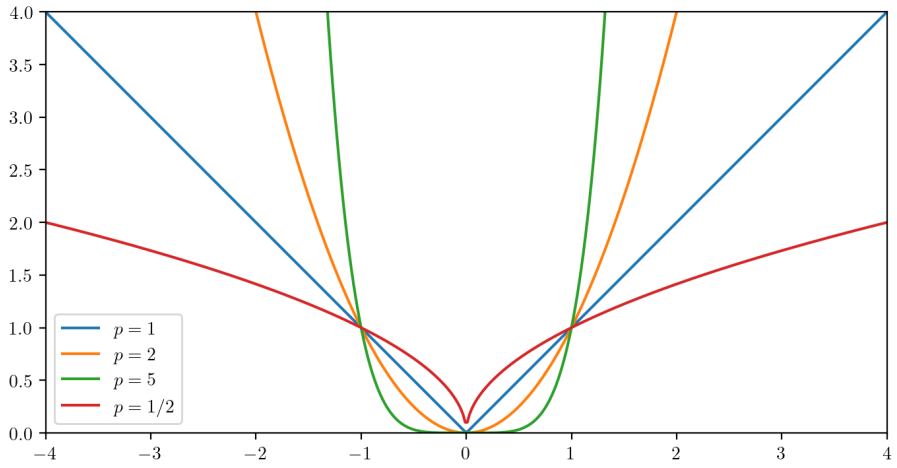

Figure 7.1: Graphs of the functions $\varphi_p(x) = |x|^p$ for different values of p. As it appears, φ_p is convex on $\mathbb{R}$ for $p \geq 1$. For $0 \leq p < 1$, it is concave on $\mathbb{R}_-$ and $\mathbb{R}_+$, but not on $\mathbb{R}$.

on Figure 7.1. As a consequence, the functional Φ_p behaves differently whether considering Wigner-positive states and Wigner-negative states. For Wigner-positives states, Φ_p is convex for $p \geq 1$ and concave for $p \leq 1$. For Wigner-negative states, Φ_p is convex for $p \geq 1$, but it is not concave (neither convex) for $0 \leq p < 1$. The functions Φ_p can somehow be understood as the building blocks of the p-norms and the Rényi entropies. Indeed, both functionals can be expressed as follows:

$$\|W\|_p = (\Phi_p(W))^{\frac{1}{p}}, \qquad h_\alpha(W) = \frac{1}{1-\alpha} \ln\left(\Phi_\alpha(W)\right). \tag{7.2}$$

It is a straightforward observation to notice the following. A non-decreasing function of a Schur-convex function (resp. Schur-concave) gives a Shur-convex function (resp. Schur-concave). Conversely, a non-increasing function of a Schur-convex function (resp. Schur-concave) gives a Schur-concave function (resp. Schur-convex). Also, a convex non-decreasing function of a convex function is convex, and a concave non-decreasing function of a concave function is concave. These observations are summarized in the following table.

$\Phi : \mathcal{W} \mapsto \mathbb{R}$	$\varphi : \mathbb{R} \mapsto \mathbb{R}$	$\varphi \circ \Phi : \mathcal{W} \mapsto \mathbb{R}$
Convex	Convex non-decreasing	Convex
Convex	Concave non-increasing	Concave
Concave	Concave non-decreasing	Concave
Concave	Convex non-increasing	Convex
Schur-convex	Non-decreasing	Schur-convex
Schur-convex	Non-increasing	Schur-concave
Schur-concave	Non-decreasing	Schur-concave
Schur-concave	Non-increasing	Schur-convex

Now, observe that the function $\varphi(x) = x^{\frac{1}{p}}$ is concave increasing over $\mathbb{R}_+$ for $p \geq 1$, and convex increasing over $\mathbb{R}_+$ for $p \leq 1$. The function $\varphi(x) = \ln(x)/(1-\alpha)$ is convex decreasing for $\alpha > 1$ and

concave increasing for $0 \leq \alpha < 1$. From that, we can conclude that for $p \geq 1$, the p-norm is a concave increasing function of a convex function, which gives a Schur-convex function. For $0 \leq p < 1$, the p-norm is a convex increasing function of a concave function, which gives a Schur-concave function. For $\alpha > 1$, the Rényi entropy is a convex decreasing function of a convex function, which gives a Schur-concave function. For $0 \leq \alpha < 1$, the Rény entropy is a concave increasing function of a concave function, which gives a concave function. Note that the case $\alpha = 1$ requires some clarification. For Wigner-positive states, we know that it corresponds to the Shannon differential entropy of the Wigner function (see Chapter 2):

$$\lim_{\alpha \to 1} \; h_\alpha(W) = h(W). \tag{7.3}$$

Let us look at the limit for Wigner-negative states.

$$\begin{aligned}
\lim_{\alpha \to 1} \quad h_\alpha(W) &= \lim_{\alpha \to 1} \quad \frac{\alpha}{1-\alpha} \ln\left(\|W\|_\alpha\right) && (7.4)\\
&= \lim_{\alpha \to 1} \quad \frac{\alpha}{1-\alpha} \ln\left(1 + 2\mathrm{Vol}_-(W)\right) && (7.5)\\
&= \ln\left(1 + 2\mathrm{Vol}_-(W)\right) \quad \lim_{\alpha \to 1} \quad \frac{\alpha}{1-\alpha} && (7.6)\\
&= \pm\infty && (7.7)
\end{aligned}$$

where we have used the fact that $\|W\|_1 = \mathrm{Vol}_+(W) + \mathrm{Vol}_-(W) = 1 + 2\mathrm{Vol}_-(W)$, which is greater than 1 for Wigner-negative states. For a Wigner-negative state described by a Wigner funciton W, the in the limit $\alpha \to 1^-$, the Rényi entropy $h_\alpha(W)$ tends towards $+\infty$. In the limit $\alpha \to 1^+$, the Rényi entropy $h_\alpha(W)$ tends towards $-\infty$. We now summarize our observations in the following table.

Functional	Parameter	$W \in \mathcal{W}_+$	$W \in \mathcal{W}$
$\Phi_p(W)$	$p \geq 1$	Convex	Convex
	$0 \leq p < 1$	Concave	$\times$
$\|W\|_p$	$p \geq 1$	Schur-convex	Schur-convex
	$0 \leq p < 1$	Schur-concave	$\times$
$h_\alpha(W)$	$\alpha > 1$	Schur-concave	Schur-concave
	$0 \leq \alpha \leq 1$	Concave	$\times$

Remember that every functional that we consider here is symmetric. As a consequence, any of the above mentioned functional that is convex (resp. concave) is also Schur-convex (resp. Schur-concave). However, the functionals labeled as Schur-convex (or Schur-concave) in the above table are not convex (neither concave).

We can understand that the p-norm for $p = 0$ is not concave when considering Wigner-negative states by a simple example. Remember that for $p = 0$, $\|W\|_0$ simply corresponds to the size of the support of W.. Consider now a function W with a non-zero support. Then the function $-W$ has the same support. However, the size of the support of $(W + (-W))/2$ is equal to zero, which is lower than the sum of the support of W and $-W$.

We have now properly identified several functionals compatible with a measure of disorder depending whether we were considering Wigner-positive or Wigner-negative states. Among these functionals,

there is the Shannon differential entropy, that we have studied in Chapter 2. As it appears, the measure converges to a useful value only for Wigner-positive states. For Wigner-negative states, the measure diverges, towards $-\infty$ (when $\alpha \to 1^+$) or $+\infty$ (when $\alpha \to 1^-$). This makes that quantity unusable for Wigner-negative states. However, because of its properties, we are going to devote a section of the present chapter to the Shannon entropy of Wigner-positive states.

7.2 Wigner entropy of Wigner-positive states

Earlier in Chapter 5, we have introduced the Husimi Q-function of a state $\hat{\rho}$ as the probability to measure $\hat{\rho}$ in a coherent state $|\alpha\rangle$. As such, it corresponds to a true probability distribution, hence it has a properly defined entropy. The Shannon differential entropy of the Husimi function is indeed known as the Wehrl entropy and is defined as $h(Q) = -\iint Q(x,p) \ln Q(x,p) \mathrm{d}x\mathrm{d}p$. This entropy is at the core of the Wehrl conjecture [67], later proven by Lieb [43, 44], which states that the Wehrl entropy is lower-bounded by $\ln \pi + 1$ and that the only minimizers of $h(Q)$ are the coherent states.

Interestingly, we shown in Chapter 5 that the Wigner function can in some particular setups be considered as the Husimi Q-function of a quantum state (see Figure 5.2). It indeed appears that the Wehrl entropy of quantum state $\hat{\rho}$ can in general be understood as the entropy of the Wigner function of another state $\hat{\sigma}$, where $\hat{\rho}$ and $\hat{\sigma}$ are related by the setup described in Figure 5.2. A natural question then arises : can we give an intrinsic meaning to the entropy of a Wigner function independently of this particular setup? In this chapter, we will answer by the affirmative. We define the Wigner-entropy of a Wigner-positive state and state several of its properties hereafter.

7.2.1 Definition and properties

Definition 7.1 (Wigner entropy)**.** *The Wigner entropy of a Wigner-positive state is the Shannon differential entropy of its Wigner function:*

$$h(W) = -\iint W(\mathbf{x},\mathbf{p}) \ln W(\mathbf{x},\mathbf{p}) \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p} \tag{7.8}$$

Let us now highlight several of its properties which make it an interesting physical quantity.

- Positivity

  The Wigner entropy of Wigner-positive states is always positive. This can easily be deduced from the fact that the Wigner function of Wigner-positive states takes values between 0 and $1/\pi$. Indeed, the value of the function $\varphi(z) = -z \ln z$ is greater than zero for z between 0 and 1. Note that the positivity of the Wigner entropy relies on a quantum property of phase space distributions, which ensures that they are not too peaked. Indeed, classically, it would be possible to make up a distribution with a negative differential entropy.

- Symplectic invariance

  It is invariant under symplectic transformations (displacement, rotation, and squeezing) in phase space. Indeed, consider the symplectic transformation $\hat{\mathbf{q}} \mapsto \hat{\mathbf{q}}' = \mathbf{U}\hat{\mathbf{q}} + \mathbf{u}$ and let us denote as W and W' the Wigner function of the input and output states, respectively. The change of variables corresponding to this transformation gives $W'(x',p') = W(x,p)/|\det \mathbf{U}|$, which indeed implies that

$$\begin{aligned} h(W') &= -\iint W'(x',p') \ln W'(x',p') \,\mathrm{d}x' \,\mathrm{d}p' \\ &= -\iint W(x,p) \ln \left(\frac{W(x,p)}{|\det \mathbf{U}|} \right) \mathrm{d}x \,\mathrm{d}p \\ &= h(W) + \ln |\det \mathbf{S}| \\ &= h(W) \end{aligned} \tag{7.9}$$

Note that this invariance can also be understood as a sole consequence of the fact that symplectic transformations conserve areas in phase space since $\det \mathbf{S} = 1$. Indeed, for any functional F, we have

$$\begin{aligned}\iint F\big(W'(x',p')\big)\,\mathrm{d}x'\,\mathrm{d}p' \\ &= \iint F\left(\frac{W(x,p)}{|\det \mathbf{S}|}\right)|\det \mathbf{S}|\,\mathrm{d}x\,\mathrm{d}p \\ &= \iint F\big(W(x,p)\big)\,\mathrm{d}x\,\mathrm{d}p \end{aligned} \tag{7.10}$$

In contrast, $h(Q)$ is greater for squeezed states than for coherent states. As it can be understood from Fig. 5.2, this preference simply originates from the fact that one input of the balanced beam-splitter is itself a coherent state.

- Additivity

  As we have seen in Chapter 2, the entropy of a a tensor product of two probability distributions is equal to the sum of the entropy of each probability distribution. In terms of Wigner-entropy, this means that when constructing a mutlimode state $\hat{\rho}$ from two Wigner-positive states $\hat{\rho}_1$ and $\hat{\rho}_2$, so that $\hat{\rho} = \hat{\rho}_1 \otimes \hat{\rho}_2$, we have the relation $h(W) = h(W_1) + h(W_2)$. The Wigner entropy of $\hat{\rho}$ is simply the sum of the Wigner entropies of $\hat{\rho}_1$ and $\hat{\rho}_2$.

- Subadditivity

  An important property of Shannon entropy is that the entropy of a joint probability distribution is always lower than the sum of the entropies of the marginal distributions. We have proven that property earlier in Chapter 2. As a consequence, the Wigner entropy $h(W)$ can be related to the entropy of the marginal distributions $h(\rho_x)$ and $h(\rho_p)$, but also encompasses the x-p correlations. Shannon information theory establishes a relation between the entropy of a joint distribution and its marginal entropies, namely $h(x,p) = h(x) + h(p) - I$, where $I \geq 0$ is the mutual information [16]. Applied to the Wigner entropy, this gives the inequality:

  $$h(W) \leq h(\rho_x) + h(\rho_p)\,. \tag{7.11}$$

  This means that a lower bound on the Wigner entropy implies in turn a lower bound on the sum of the marginal entropies. More generally, this can be used to express that the Wigner entropy of a multi-mode quantum system is lower or equal to the sum of the Wigner entropy of each subsystems taken separately:

  $$h(W) \leq h(W_1) + h(W_2), \tag{7.12}$$

  where W is the Wigner function of a Wigner-positive two-mode state $\hat{\rho}$, while W_1 and W_2 are respectively the Wigner functions of $\mathrm{Tr}_2[\hat{\rho}]$ and $\mathrm{Tr}_1[\hat{\rho}]$.

- Concavity

  Finally, let us recall that entropy is a concave functional. This implies that mixing Wigner-positive states can only result in an increase of their Wigner-entropy, which mathematically translates as follows:

  $$h\left(\sum p_i W_i\right) \geq \sum p_i h(W_i) \tag{7.13}$$

  where W_i are non-negative Wigner functions.

7.2.2 Wigner-entropy power inequality

Since the Wigner entropy is the Shannon differential entropy of the Wigner function, viewed as a genuine probability distribution, it inherits all its key features. For example, we may easily extend to Wigner entropies the celebrated entropy power inequality [16], which relates to the entropy of the

convolution of probability distributions. With this in mind, we define the Wigner-entropy power of a Wigner-positive state has the entropy power of its Wigner function W:

$$N(W) = \frac{1}{2\pi e} \exp(h(W)) \tag{7.14}$$

Consider now an arbitrary transformation between two optical modes, which can be expressed in the general expression:

$$W' = \mathcal{L}_a [W_A] * \mathcal{L}_b [W_B] \tag{7.15}$$

where W' is the Wigner function of the state $\hat{\rho}'$, which is built from the states $\hat{\rho}_A$ and $\hat{\rho}_B$ respectively described by the Wigner functions W_A and W_B. Relation (7.15) can for example represent the action of a beam-splitter of transmittance η if we choose $a = \sqrt{\eta}$ and $b = \sqrt{1-\eta}$.

Let us restrict to the special case where both $\hat{\rho}_A$ and $\hat{\rho}_B$ are Wigner-positive states, which of course implies that $\hat{\rho}'$ is Wigner-positive as well. Thus, $\hat{\rho}_A$, and $\hat{\rho}_B$, and $\hat{\rho}'$ all have a well-defined Wigner entropy, which we denote respectively as $h(W_A)$, $h(W_B)$, and $h(W')$. Hence, the entropy power inequality directly applies to the Wigner entropy. Defining the Wigner entropy-power of the two input states $\hat{\rho}_A$ and $\hat{\rho}_B$ respectively as $N(W_A)$ and $N(W_B)$, and the Wigner entropy-power of the output $\hat{\rho}'$ as $N(W')$, we obtain the *Wigner-entropy power inequality*:

$$N(W') \geq a^2 N(W_A) + b^2 N(W_B). \tag{7.16}$$

This is equivalent to a nontrivial lower bound on the Wigner entropy of the output state $\hat{\rho}'$, namely $h(W') \geq h(W_{\hat{\tau}})$, where $\hat{\tau}$ denotes the Gaussian output state obtained if each input state is replaced by the phase-invariant Gaussian state (thermal state) with the same Wigner entropy. This illustrates the physical significance of the Wigner entropy.

As we have explained, relation (7.15) can represent a beam-splitter for the adapted choice of a and b. For a beam-splitter of transmittance η, we have the following Wigner-entropy power inequality:

$$N(W') \geq \eta N(W_A) + (1-\eta) N(W_B). \tag{7.17}$$

Moreover, relation (7.15) can also be used to derive a Wigner-entropy power inequality in a two-mode squeezer. Indeed, the action of a two-mode squeezer corresponds to (7.15) except that we should add the action of a conjugation operator Π. However, the conjugation operator is a transformation which does not change the Wigner-entropy of a state, and therefore neither its Wigner-entropy power. The Wigner-entropy power in a two-mode squeezer of gain g reads as follows:

$$N(W') \geq g N(W_A) + (g-1) N(W_B) \tag{7.18}$$

Note that these relations only hold for Wigner-positive states.

7.2.3 Conjectured lower-bound

In the light of these considerations, we introduce a conjecture on the Wigner entropy, which resembles the Wehrl conjecture. As anticipated in [29], we conjecture the following statement.

Conjecture 7.1 (Wigner entropy)**.** *The Wigner entropy of a single-mode Wigner-positive quantum state is lower-bounded by* $\ln \pi + 1$*:*

$$h(W) \geq \ln \pi + 1 \qquad \forall W \in \mathcal{W}_+ \tag{7.19}$$

It implies (but is stronger than) the entropic uncertainty relation of Białynicki-Birula and Mycielski [7], namely $h(\rho_x) + h(\rho_p) \geq \ln \pi + 1$. Indeed, this follows from the property of subadditivity that we have highlighted previously. Importantly, Conjecture 7.1 also implies the Wehrl conjecture since we have shown that the Husimi function of any state $\hat{\rho}$ is the Wigner function of some Wigner-positive state $\hat{\sigma}$ in a particular setup, see Fig. 5.2. However, the converse is not true as there exist Wigner-positive

states whose Wigner function cannot be written as the Husimi function of a physical state. Note that the conjecture extends straightforwardly to Wigner-positive multimode state as $h(W) \geq N(\ln \pi + 1)$ where N is the number of modes. Notice also that Conjecture 7.1 is a direct consequence of Conjecture 6.1, since Shannon differential entropy is Schur-concave over the set of non-negative Wigner functions. We can summarize the chain of implications as follows:

$$W \prec W_0 \qquad \Rightarrow \qquad h(W) \geq \ln \pi + 1 \qquad \Rightarrow \qquad \begin{cases} h(Q) \geq \ln \pi + 1 \\ \\ h(\rho_x) + h(\rho_p) \geq \ln \pi + 1 \end{cases} \tag{7.20}$$

The special case of Gaussian states is very easy to deal with. A straightforward calculation shows that the Wigner entropy of a Gaussian state $\hat{\rho}$ is given by

$$h(W) = \ln\left(2\pi\sqrt{\det \mathbf{V}}\right) + 1 = \ln(\pi/\mu) + 1 \tag{7.21}$$

where $\mu = \mathrm{Tr}\left[\hat{\rho}^2\right] = 1/(2\sqrt{\det \mathbf{V}}) \leq 1$ stands for the purity of the state. All Gaussian states that are connected with a symplectic transformation obviously conserve their purity since $\det \mathbf{V}' = \det(\mathbf{U}\mathbf{V}\mathbf{U}^\intercal) = \det \mathbf{V}$, which confirms that their Wigner entropy is invariant. The lowest value of $h(W)$ among Gaussian states is then reached for pure states ($\mu = 1$) and is given by $\ln \pi + 1$, which appears consistent with the Hudson theorem [34]. This is the value of the Wigner entropy of all coherent states and squeezed states (regardless the squeezing parameter, squeezing orientation, and coherence vector). Accordingly, the Gaussian pure states would be the minimum-Wigner-uncertainty states.

Note that a direct consequence of Equation (7.21) together with the property of concavity of Wigner entorpy is that the conjecture is satisfied over the whole convex hull of Gaussian states, i.e. $\mathcal{G}$. Indeed, any state of $\mathcal{G}$ is built as a convex mixture of quantum states with Wigner entropy satisfying the lower-bound. The difficult task remains, however, to prove that non-Gaussian Wigner-positive states that do not belong to the convex hull of Gaussian states cannot violate this lower bound.

Provided this conjecture is valid, the Wigner function of any Wigner-positive state can be classically simulated from the Wigner function of the vacuum state (or any other Gaussian pure state). More precisely, information theory tells us that the difference $\Delta = h(W) - \ln \pi - 1$ can be viewed as the number of independent equiprobable random bits that are needed, on average, to generate deterministically one random (x, p) instance drawn from the Wigner function of state $\hat{\rho}$ from one random (x, p) instance drawn from the Wigner function of the vacuum state (or any Gaussian pure state). Of course, this results holds at the asymptotic limit only, that is, around $N \times \Delta$ bits of extra randomness are needed for converting N random instances of $(x, p) \sim W_0$ into N random instances of $(x, p) \sim W$ by deterministic means when $N \to \infty$.

7.2.4 Proof of the lower-bound over the set of passive states

Let us prove Conjecture 7.1 for the subset of passive states $\hat{\rho}_p$. As we have seen in Chapter 5, passive states indeed possess non-negative Wigner functions, so that their Wigner entropy is well defined. Also, notice that, as a consequence of the concavity of entropy, it is sufficient to prove the Wigner conjecture for extremal passive states $\hat{\varepsilon}_n$.

The main tool that we will use to carry out our proof is the formula (5.47) which makiesa nontrivial link between the Wigner functions and wave functions of the first n Fock states. Let us denote the x and p probability densities of the n^{th} Fock state as $\rho_n(x) = |\psi_n(x)|^2$ and $\rho_n(p) = |\psi_n(p)|^2$. Their corresponding Shannon differential entropy is defined as $h\left(\rho_k(x)\right) = -\int \rho_k(x) \ln \rho_k(x)\, \mathrm{d}x$ and $h\left(\rho_k(p)\right) = -\int \rho_k(p) \ln \rho_k(p)\, \mathrm{d}p$. In the following, we refer to these quantities as $h\left(\rho_k\right) \equiv h\left(\rho_k(x)\right) = h\left(\rho_k(p)\right)$. We are now ready to lower bound the Wigner entropy of the n^{th} extremal passive state $\hat{\varepsilon}_n$

by using Eq. (5.47):

$$\begin{aligned} h\left(E_n\right) &= h\left(\frac{1}{n+1}\sum_{k=0}^{n} W_k(x,p)\right) \\ &= h\left(\frac{1}{n+1}\sum_{k=0}^{n} \psi_k(x)^2\psi_{n-k}(p)^2\right) \\ &\geq \frac{1}{n+1}\sum_{k=0}^{n} h\big(\rho_k(x)\rho_{n-k}(p)\big) \\ &= \frac{1}{n+1}\sum_{k=0}^{n} \big(h\left(\rho_k\right) + h\left(\rho_{n-k}\right)\big) \\ &= \frac{2}{n+1}\sum_{k=0}^{n} h\left(\rho_k\right) \\ &\geq \ln \pi + 1 \end{aligned} \tag{7.22}$$

The first inequality in (7.22) results from the concavity of the entropy. Then, we use the fact that the entropy of a product distribution is the sum of the marginal entropies. Finally, we apply the entropic uncertainty relation of Białynicki-Birula and Mycielski [7] on Fock states, namely $2\,h\left(\rho_k\right) \geq \ln \pi + 1$, $\forall k$. We have thus proven the conjecture for all extremal passive states and this proof naturally extends to the whole set of passive states.

Let us now prove that a slightly tighter lower bound can be derived for the Wigner entropy of passive states. To that purpose, we exploit the fact that passive states can be expressed as a convex mixture of extremal passive states $\hat{\varepsilon}_n$ (in place of decreasing mixtures of Fock states). Indeed, any passive state can equivalently be expressed as:

$$\hat{\rho}_p = \sum_{k=0}^{\infty} p_k \left|k\right\rangle \left\langle k\right| = \sum_{k=0}^{\infty} e_k \hat{\varepsilon}_k \tag{7.23}$$

where p_k are probabilities such that $p_k \geq p_{k+1}$ and the coefficients e_k are related to p_k as $e_k = (k+1)(p_k - p_{k+1})$. We denote the Wigner function of the passive state $\hat{\rho}_p$ as $W_P(x,p)$ and bound its Wigner entropy as follows:

$$\begin{aligned} h\left(W_P\right) &= h\left(\sum_{k=0}^{+\infty} e_k\, E_k(x,p)\right) \\ &\geq \sum_{k=0}^{\infty} e_k\, h\big(E_k(x,p)\big) \\ &= \sum_{k=0}^{\infty} (k+1)\left(p_k - p_{k+1}\right) h\big(E_k(x,p)\big) \\ &\geq \sum_{k=0}^{\infty} (k+1)\left(p_k - p_{k+1}\right) \frac{2}{k+1}\sum_{j=0}^{k} h\left(\rho_j\right) \\ &= 2\sum_{k=0}^{\infty}\sum_{j=0}^{k} \left(p_k - p_{k+1}\right) h\left(\rho_j\right) \\ &= 2\sum_{j=0}^{\infty}\sum_{k=j}^{\infty} \left(p_k - p_{k+1}\right) h\left(\rho_j\right) \\ &= 2\sum_{j=0}^{\infty} p_j\, h\left(\rho_j\right) \end{aligned} \tag{7.24}$$

The first inequality in (7.24) comes from the concavity of entropy over the convex set of extremal states, while the second inequality is obtained from Eq. (7.22). The final expression is a stronger lower bound on the Wigner entropy of any passive state which reads as

$$h\left(\sum_{k=0}^{\infty} p_k\, W_k\right) \geq 2\sum_{k=0}^{\infty} p_k\, h\left(\rho_k\right). \tag{7.25}$$

and is valid as soon as the probabilities p_k are decreasing, that is, $p_k \geq p_{k+1}$.

It is tempting to extrapolate that the bound (7.25) remains valid beyond the set of passive states. We know indeed that there exist phase-invariant Wigner-positive states that are not passive states (in Figure 5.10, these are the states within the blue region, which do not belong to the triangle a-b-e). As long as the coefficients p_k are such that the corresponding state is Wigner-positive, it has a well-defined Wigner entropy and we may infer that the lower bound (7.25) applies. Unfortunately, our numerical simulations have shown that relation (7.25) does not hold in general for non-passive (Wigner-positive) states. Of course, we maintain that Conjecture 7.1 does hold for such states and we have not found any counterexample.

7.3 Wigner-Rényi entropies

As we know, the Shannon differential entropy is an uncertainty measure that belongs to a broader family, known as Rényi differential entropies. Moreover, as we explained in the first section, the Renyi entropies have the particularity that they can be defined for negative distributions. With this in mind, we define the Rényi-Wigner entropy of a (Wigner-positive or Wigner-negative) quantum state as follows:

Definition 7.2 (Wigner-Rényi entropy)**.** *The Wigner-Rényi entropy a quantum state is the Rényi entropy of its Wigner function:*

$$h_\alpha(W) = \frac{1}{1-\alpha}\ln\left(\iint |W(\mathbf{x},\mathbf{p})|^\alpha \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p}\right) \tag{7.26}$$

where $\alpha \in \mathbb{R}_+ \setminus \{1\}$.

Interestingly, some values of α are endowed with a special meaning. Denoting as $\mathrm{supp}(W)$ the part of the domain of $W(x,p)$ where $W(x,p) \neq 0$ and denoting as ν the Lebesgue measure, we have:

$$h_0(W) = \ln(\nu[\mathrm{supp}(W)]) \tag{7.27}$$

This diverges when applied to any Wigner function W since the size of the support of W is infinite. In the limit $\alpha \to 1$, h_α tends to the Shannon differential entropy, so that $h_1(W)$ coincides with $h(W)$ for Wigner-positive states.

$$\lim_{\alpha\to 1} h_\alpha(W) = h(W) \qquad \forall W \in \mathcal{W}_+ \tag{7.28}$$

However the limit $\alpha \to 1$ diverges for Wigner-negative states as we have seen in the first section. The Rényi entropy with parameter $\alpha = 2$ is sometimes called the collision entropy, and, applied to a Wigner function W, it is related to the purity of the corresponding state through the relation:

$$h_2(W) = \ln\left(\frac{2\pi}{\mu}\right) \tag{7.29}$$

Finally, the case $\alpha \to \infty$ can be related to the maximum value of W as

$$h_\infty(W) = -\ln\left(\max_{x,p}|W(x,p)|\right). \tag{7.30}$$

Note that, following the same reasoning as (7.10), we observe that the Wigner-Rényi entropy is invariant under symplectic transformations in phase space (i.e., Gaussian unitaries in state space).

It is interesting to note that it is possible to exploit Young's convolution inequality introduced in Chapter 2 in order to derive a relation between the Rényi-Wigner entropy of the input and output of a beam-splitter or two-mode squeezer. Let us consider a general transformation defined as (7.15), where the parameter a and b can be chosen appropriately to represent either a beam-splitter or a two-mode squeezer. The state $\hat{\rho}'$ is obtained from the states $\hat{\rho}_A$ and $\hat{\rho}_B$, and Equation (7.15) describes how their respective Wigner functions W', W_A and W_B are related to each other. We can apply Young's convolution inequality to (7.15) to get an inequality in terms of p-norms:

$$\|W'\|_r \leq \|\mathcal{L}_a[W_A]\|_p \cdot \|\mathcal{L}_b[W_B]\|_q \tag{7.31}$$

$$= |a|^{2\frac{1-p}{p}} |b|^{2\frac{1-q}{q}} \|W_A\|_p \cdot \|W_B\|_q \tag{7.32}$$

where $p, q, r \geq 1$ are related through the relation $1/p + 1/q = 1/r + 1$. Then, we can use the above expression to translate it in terms of Wigner-Rényi entropies, so that we find:

$$\frac{r-1}{r} h_r(W') \geq \frac{p-1}{p}\left(h_p(W_A) + \ln a^2\right) + \frac{q-1}{q}\left(h_q(W_B) + \ln b^2\right). \tag{7.33}$$

We can particularize that relation to the case of a beam-splitter with transmittance η, in which case we obtain the following inequality:

$$\frac{r-1}{r} h_r(W') \geq \frac{p-1}{p}\left(h_p(W_A) + \ln \eta\right) + \frac{q-1}{q}\left(h_q(W_B) + \ln(1-\eta)\right). \tag{7.34}$$

In the case of a two-mode squeezer of gain g, we obtain the following inequality:

$$\frac{r-1}{r} h_r(W') \geq \frac{p-1}{p}\left(h_p(W_A) + \ln g\right) + \frac{q-1}{q}\left(h_q(W_B) + \ln(g-1)\right). \tag{7.35}$$

7.3.1 Conjectured lower-bound for Wigner-positive states

In the same spirit as Conjecture 7.1, we lay a lower-bound over the value taken by the Wigner-Rényi entropy of a Wigner-positive state. Namely, we state that the Wigner-Rényi entropy of any Wigner-positive state is greater or equal to the value it takes for a pure Gaussian state. For the vacuum state (or any pure Gaussian state), the Wigner-Rényi entropy gives

$$h_\alpha(W_0) = \ln \pi + \frac{\ln \alpha}{\alpha - 1}. \tag{7.36}$$

Using this, we lay the following conjecture.

Conjecture 7.2 (Wigner-Rényi entropy of Wigner-positive states)**.** *The Wigner-Rényi entropy of any Wigner-positive state is lower bounded by the value it takes for a pure Gaussian state:*

$$h_\alpha(W) \geq \ln \pi + \frac{\ln \alpha}{\alpha - 1} \qquad \forall W \in \mathcal{W}_+ \tag{7.37}$$

Of course, Conjecture 7.2 coincides with Conjecture 7.1 when $\alpha \to 1$. Notice also that it is a straightforward implication of Conjecture 6.1, since Rényi entropies are Schur-concave $\forall \alpha \geq 0$ over the set of non-negative Wigner functions. Let us examine this Wigner-Rényi conjecture for other special values of the parameter α, namely

$$h_2(W) \geq \ln 2\pi, \tag{7.38}$$

$$h_\infty(W) \geq \ln \pi. \tag{7.39}$$

For $\alpha = 2$, the fact that the purity μ is upper bounded by 1 implies Eq. (7.38). Also, the Wigner function of any sate is upper bounded by $1/\pi$, which implies Eq. (7.39) for $\alpha \to \infty$. Furthermore, for $\alpha = 0$, the Wigner-Rényi conjecture implies that the support of any Wigner function is unbounded, which is a well-know fact. These elements support the validity of the Wigner-Rényi conjecture and especially conjecture 7.1 when $\alpha \to 1$.

We plot on Figure 7.2 the Wigner-Rényi entropies of several Wigner-positive states for different values of α. As it appears from the figure, it is always greater than the value it takes for vacuum.

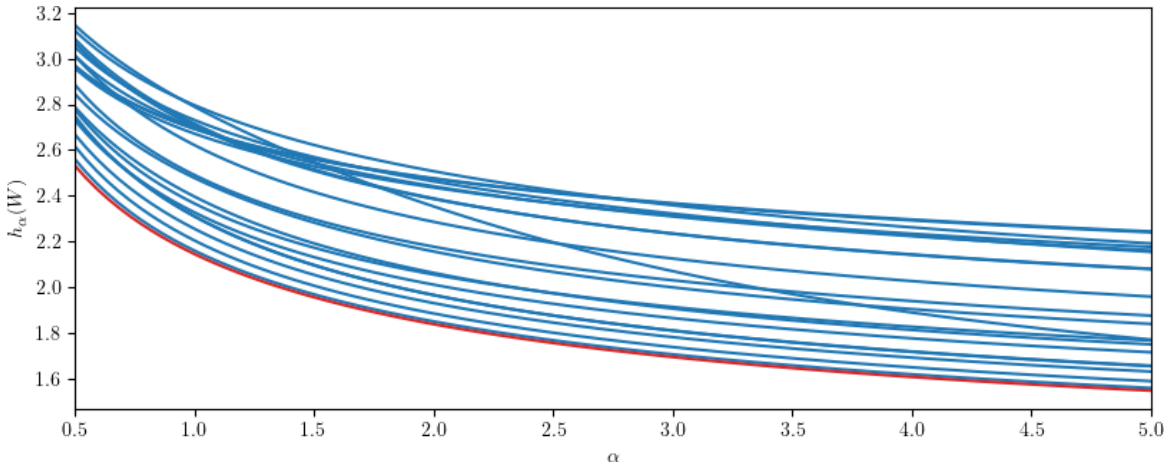

Figure 7.2: We plot here the Wigner-Rényi entropies $h_\alpha(W)$ of several Wigner-positive states as a function of α. Each blue line corresponds to a (mixed) Wigner-positive state with Wigner function $W \in \mathcal{W}_+$. The red line corresponds to vacuum (or any pure Gaussian state), and can be computed analytically from Equation 7.36. As conjectured in Conjecture 7.2, each blue line lies above the red line. Note that in the limit $\alpha \to 0^+$, the Wigner-Rényi entropy tends towards $+\infty$.

7.3.2 Conjectured lower-bound for Wigner-negative states

At the end of Chapter 3, we have introduce an alternative version of majorization that we called square-majorization. Along the way, we formulated Conjecture 6.2, which states than vacuum (or any pure Gaussian state) square-majorizes the Wigner function of any pure state.

$$W_0 \succ_{\mathrm{s}} W \quad \Leftrightarrow \quad \iint \varphi\left((W_0(x,p))^2\right) \mathrm{d}x\mathrm{d}p \geq \iint \varphi\left((W(x,p))^2\right) \mathrm{d}x\mathrm{d}p \qquad \forall \varphi \text{ convex} \tag{7.40}$$

Let us now particularize that to the convex function $\varphi_m(x) = |x|^m$ where $m \geq 1$. Conjecture 6.2 thus implies an inequality in terms of the function φ_m. We can write the following development:

$$\iint \varphi_m\left((W_0(x,p))^2\right) \mathrm{d}x\mathrm{d}p \geq \iint \varphi_m\left((W(x,p))^2\right) \mathrm{d}x\mathrm{d}p \tag{7.41}$$

$$\Leftrightarrow \quad \iint \left(|W_0(x,p)|\right)^{2m} \mathrm{d}x\mathrm{d}p \geq \iint \left(|W(x,p)|\right)^{2m} \mathrm{d}x\mathrm{d}p \tag{7.42}$$

$$\Leftrightarrow \quad \ln\left(\iint \left(|W_0(x,p)|\right)^{2m} \mathrm{d}x\mathrm{d}p\right) \geq \ln\left(\iint \left(|W_0(x,p)|\right)^{2m} \mathrm{d}x\mathrm{d}p\right) \tag{7.43}$$

$$\Leftrightarrow \quad \frac{1}{1-2m} \ln\left(\iint \left(|W_0(x,p)|\right)^{2m} \mathrm{d}x\mathrm{d}p\right) \leq \frac{1}{1-2m} \ln\left(\iint \left(|W(x,p)|\right)^{2m} \mathrm{d}x\mathrm{d}p\right) \tag{7.44}$$

$$\Leftrightarrow \quad h_{2m}(W_0) \leq h_{2m}(W) \tag{7.45}$$

Remember now that in order for φ_m to be convex, we have chosen $m \geq 1$. As a consequence, we see that Conjecture 6.2 implies the following inequality in terms of Wigner-Rényi entropies:

$$W_0 \succ_{\mathrm{s}} W \qquad \Rightarrow \qquad h_\alpha(W_0) \leq h_\alpha(W) \qquad \forall \alpha \geq 2 \tag{7.46}$$

This leads us to state the following Wigner-Rényi conjecture applying to pure states.

Conjecture 7.3 (Lower-bound on the Wigner-Rényi entropy of a pure state)**.** *For $\alpha \geq 2$, the Wigner-Rényi entropy of a pure state is lower-bounded by the Wigner-Rényi entropy of a pure Gaussian state.*

$$h_\alpha(W) \geq h_\alpha(W_0) \qquad \forall \alpha \geq 2, \quad \forall W \in \mathcal{W}^* \tag{7.47}$$

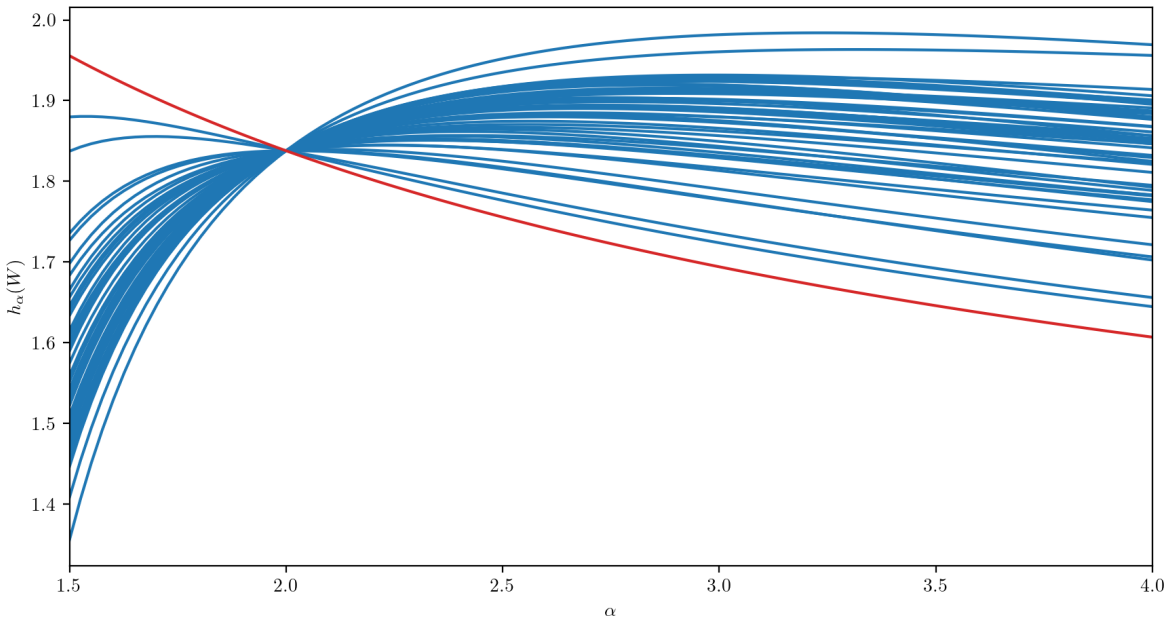

Figure 7.3: We plot here the Wigner-Rényi entropies $h_\alpha(W)$ of several pure states as a function of α. Each blue line corresponds to a non-Gaussian pure state with Wigner function $W \in \mathcal{W}^*$. The red line corresponds to vacuum (or any pure Gaussian state), and can be computed analytically from Equation 7.36. At $\alpha = 2$, all the lines meet since we are only considering pure states, which are such that $h_2(W) = \ln(2\pi)$. As conjectured in Conjecture 7.3, each blue line lies above the red line for $\alpha > 2$. Moreover, we observe that for $1 < \alpha < 2$, the situation is reversed and each blue line lies below the red line. Note that we have restriced the plot to values of α larger than 1 since the Wigner-Rényi entropy is Schur-concave over the set $\mathcal{W}$ only when $\alpha > 1$. For any non-Gaussian pure state (which is thus Wigner-negative), the Wigner-Rényi entropy tends towards $-\infty$ in the limit $\alpha \to 1^+$.

We plot on Figure 7.3 the Wigner-Rényi entropy of several pure states, and compare them with vacuum. The conjecture appears to be satisfied for all $\alpha \geq 2$. It is interesting to note that in the case $\alpha = 2$, the Wigner-Rényi entropy of any pure states gives $\ln 2\pi$, so that in that precise case the inequality of Conjecture 7.3 becomes an equality (and it is obviously proven).

We can also highlight an interesting observation from Figure 7.3. Indeed, in the case $1 < \alpha < 2$, it appears that the inequality of Conjecture 7.3 is reversed. This does not contradict the conjecture as Conjecture 7.3 is limited to the case $\alpha \geq 2$. We see thus that for $1 < \alpha < 2$, Gaussian pure states maximize the Wigner-Rényi entropy among pure states, whereas they minimize the Wigner-Rényi entropy for $\alpha \geq 2$. That observation can be understood as a manifestation of the fact that pure states are in general incomparable. Indeed, as we have said in the first section of this chapter, Wigner-Rényi entropies are consistent Schur-concave measures over the whole set $\mathcal{W}$ for any $\alpha > 1$. Figure 7.3 reveals that vacuum is incomparable to any non-Gaussian pure state, since the its Wigner-Rényi entropy is lower for $\alpha \geq 2$, but greater for $1 < \alpha < 2$.

Finally, let us conclude this chapter by a comparison of the Wigner-Rényi entropies of Fock states, which is pictured in Figure 7.4. Remember than at the end of Chapter 3, we conjectured a chain of square-majorization between Fock states, namely:

$$W_0 \succ_{\mathrm{s}} W_1 \succ_{\mathrm{s}} W_2 \succ_{\mathrm{s}} \cdots \succ_{\mathrm{s}} W_n \succ_{\mathrm{s}} W_{n+1} \tag{7.48}$$

where W_n is the Wigner function of the n^{th} Fock state. Following the same development carried earlier in this subsection, that square-majorization chain implies the following inequality:

$$h_\alpha(W_0) \leq h_\alpha(W_1) \leq \cdots \leq h_\alpha(W_n) \leq h_\alpha(W_{n+1}) \qquad \forall \alpha \geq 2. \tag{7.49}$$

That observation is confirmed by Figure 7.4, but the figure also reveals that the inequality in terms of Wigner-Rényi entropies is reversed when $1 < \alpha < 2$, in which case we observe:

$$h_\alpha(W_0) \geq h_\alpha(W_1) \geq \cdots \geq h_\alpha(W_n) \geq h_\alpha(W_{n+1}) \qquad \forall \alpha : \ 1 < \alpha < 2. \tag{7.50}$$

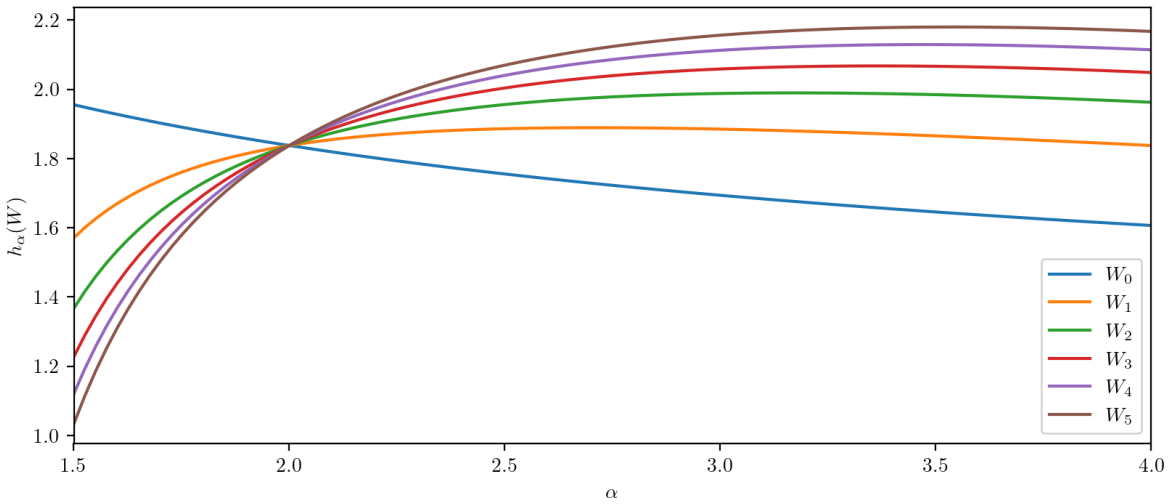

Figure 7.4: We plot here the Wigner-Rényi entropies $h_\alpha(W)$ for the first Fock states as a function of α. We restrict ourselves to Fock states from $n = 0$ to $n = 5$. We observe that for $\alpha > 2$, we have the relation $h_\alpha(W_n) \leq h_\alpha(W_{n+1})$. For α such that $1 < \alpha < 2$, we observe the reverse inequality and find $h_\alpha(W_n) \geq h_\alpha(W_{n+1})$. Since all lines cross each other, we can conclude that Fock states are incomparable with respect to each other.

Once again, this support the fact that pure states are in general incomparable. Indeed, Equations (7.49) and (7.50) illustrate that Fock states are all incomparable with respect to each other, since Wigner-Rényi are Schur-concave over the set $\mathcal{W}^*$ for all $\alpha > 1$.

Chapter 8

Majorization relations in bosonic channels

in phase space and in state space

The use we have made of the theory of majorization so far was to introduce general conjectures that applies indistinctly to large sets of quantum states. Indeed, in Chapter 6, we have introduced Conjecture 6.1 that applies to the whole set of Wigner-positive states, and Conjecture 6.2 that applies to the whole set of pure states. In this chapter, we will make a different use of majorization. Rather than being interested in a majorization relation between a precisely chosen state on one side and a whole set of states on the other side, we will establish pairs of states linked by a majorization relation, where the second state is obtained from the first by a particular transformation.

As a matter of fact, we are going to compare quantum states with their output through particular quantum channels. In quantum mechanics, a channel is generaly described by a global unitary acting on an input system and its environment, followed by a partial trace to discard the environment. The situation is described as follows:

$$\mathcal{M}\left(\hat{\rho}\right) = \mathrm{Tr}_E\left[\hat{U}\left(\hat{\rho}\otimes\hat{\rho}_E\right)\hat{U}^\dagger\right], \tag{8.1}$$

where the quantum channel $\mathcal{M}$ acts on the quantum states $\hat{\rho}$ as a global unitary $\hat{U}$ over the system $\hat{\rho}$ together with its environment $\hat{\rho}_E$, which is then discarded by a partial trace. In that context, Gaussian channels are channels that have the particularity of being described by a Gaussian environment and a Gaussian unitary. As a consequence, these channels associate a Gaussian output to any Gaussian input. It is to these channels that we turn in this chapter, and more particularly to Gaussian channels which are phase-invariant and that act on single-mode inputs.

In the first section of this chapter we will lay the theoretical foundations for describing Gaussian channels in simple terms. We will see that Gaussian single-mode phase-insensitive channels are in all generality defined by two real parameters. We will then also present how these channels can be implemented in an optical setup, making a link with Chapter 4. Following this, we will in the next section study how the Wigner function of a quantum state evolves in such a channel. This will lead us to observe that in some cases the output associated to the channel is always Wigner-positive for all input. We will also highlight a continuous majorization relation occurring in a two-mode squeezer.

In the last section of this chapter, we will take a step back from the phase-space perspective that we have used through this thesis. We will change our point of view to the state space picture and will consider the discrete distributions embodied by the probability vectors of eigenvalues of density operators. At that point, we will no longer use continuous majorization, but rather discrete majorization. Building on a previous work that revealed unexpected recurrence relations in Gaussian unitaries [35, 36], we will prove a discrete majorization chain for the outputs associated with Fock states in Gaussian bosonic channels. This work is an extension of a result previously highlighted in [20, 21]. This last section is the subject of a paper in preparation entitled *Majorization ladder in Gaussian bosonic channels*.

8.1 Single-mode phase-insensitive Gaussian bosonic channels

We have seen in Chapter 4 that simple optical transformations were often associated to convolution between Wigner functions. As a consequence, it appears that a spectral decomposition of Wigner functions can provide a simple formalism to deal with these operations. It is in this objective that we introduce hereafter the Fourier transform of a Wigner function:

$$\tilde{W}(u,v) = \mathcal{F}[W](u,v) = \frac{1}{2\pi}\iint W(x,p)\exp(-iux - ivp)\,\mathrm{d}x\mathrm{d}p, \tag{8.2}$$

$$W(x,p) = \mathcal{F}^{-1}\left[\tilde{W}\right](x,p) = \frac{1}{2\pi}\iint \tilde{W}(u,v)\exp(ixu + ipv)\,\mathrm{d}u\mathrm{d}v. \tag{8.3}$$

We associate the operator $\mathcal{F}$ to the Fourier transformation, and the operator $\mathcal{F}^{-1}$ as the inverse transformation. Surprisingly, the Fourier transform of the Wigner function $\tilde{W}$ is not often used in the literature. Instead, we mostly find the characteristic function $\chi(z)$, defined as $\chi(z) = \mathrm{Tr}\left[\hat{\rho}\hat{D}_z\right]$. The characteristic function is actually the Fourier transform of the Wigner function with a slight adjustment. Indeed, a small derivation gives us the following relationship:

$$\chi(x+ip) = 2\pi\tilde{W}\left(-\frac{p}{\sqrt{2}}, \frac{x}{\sqrt{2}}\right) \tag{8.4}$$

where $x, p \in \mathbb{R}$. The characteristic function χ can somehow be understood as a *symplectic* Fourier transform, since it also applies the symplectic form $\boldsymbol{\omega}$ over the vector $(x,p)^\intercal$.

At this point it is interesting to make two different observations. First, it is easily shown how the rescaling operator $\mathcal{L}_s$ acts with respect to the Fourier transformation:

$$\mathcal{F}[\mathcal{L}_s[W]] = \frac{1}{s^2}\,\mathcal{L}_{\frac{1}{s}}[\mathcal{F}[W]]. \tag{8.5}$$

Then, as we will deal with Gaussian states, and more particularly with thermal states, it is important for our further developments to compute the Fourier transform of thermal states. The Fourier transform of the Wigner function of vacuum W_0 can be found as the following:

$$\mathcal{F}[W_0](u,v) = \frac{1}{2\pi}\exp\left(-\frac{1}{4}(u^2+v^2)\right). \tag{8.6}$$

Notice that (8.6) can be equivalently stated as $\mathcal{F}[W_0] = \mathcal{L}_{\sqrt{2}}[W_0]$. Now, remember from Chapter 4 that we have the relation $W_{\hat{\tau}} = \mathcal{L}_{\sqrt{2\bar{n}+1}}[W_0]$ (see (4.91)). As a consequence, the Fourier transform of any thermal state can be easily computed from (8.5) and (8.6).

After this small theoretical development, we are ready to present in proper terms Gaussian phase-insensitive channels. Hereafter, we follow mostly reference [23]. A single-mode Gaussian phase-insensitive is defined from a gain/loss parameter $\kappa \in \mathbb{R}$ and a noise parameter $\mu \in \mathbb{R}_+$. Let us mention that κ and μ should satisfy the relation $\mu \geq |\kappa - 1|$ in order to be a physically realizable channel. In the case where the previous inequality is saturated, the quantum channel is said to be quantum-limited. These channels are divided in two categories: covariant channels for $\kappa \geq 0$ and contravariant channels for $\kappa < 0$. Let us now consider a single-mode quantum state $\hat{\rho}$ with characteristic function $\xi(z)$ and Wigner function $W(x,p)$. We consider that the state $\hat{\rho}$ evolves as follows:

$$\hat{\rho} \quad \rightarrow \quad \mathcal{M}_\kappa^{(\mu)}(\hat{\rho}), \tag{8.7}$$

where $\mathcal{M}_\kappa^{(\mu)}$ is a quantum channels with gain/loss parameter κ and noise parameter μ. The action of $\mathcal{M}_\kappa^{(\mu)}$ over $\hat{\rho}$ is generally described from the characteristic function of $\hat{\rho}$. Using the tools previously introduced in this section, we will then be able to translate that action onto $\tilde{W}$, and then onto W. In

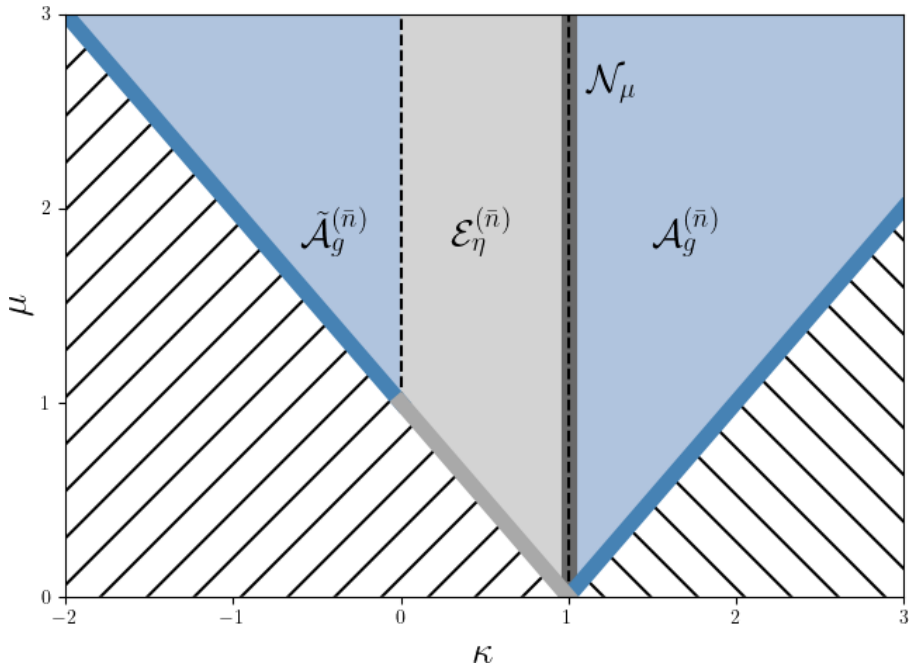

Figure 8.1: Illustration of phase-invariant single-mode Gaussian quantum channels. The hatched region corresponds to a forbidden region for physical channels ($\mu \geq |\kappa - 1|$). For $\kappa < 0$, each channel can be implemented as a conjugated amplifier $\tilde{\mathcal{A}}_g^{(\bar{n})}$, and the bold line corresponds to the quantum-limited conjugated amplifier $\mathcal{A}_g$. For $0 \leq \kappa < 1$, each channel can be implemented as a lossy channel $\mathcal{E}_\eta^{(\bar{n})}$, and the bold line corresponds to the pure-loss channel $\mathcal{E}_\eta$. For $\kappa > 1$, each channel can be implemented as an amplifier $\mathcal{A}_g^{(\bar{n})}$ and the bold line corresponds to the quantum-limited amplifier $\mathcal{A}_g$. For $\kappa = 1$, the channel corresponds to an added noise channel $\mathcal{N}_\mu$.

the covariant case ($\kappa \geq 0$), we obtain the following:

$$\chi(z) \quad \longrightarrow \quad \chi(\sqrt{\kappa}z) \exp\left(-\mu \frac{|z|^2}{2}\right) \tag{8.8}$$

$$\tilde{W}(u,v) \quad \longrightarrow \quad \tilde{W}\left(\sqrt{\kappa}u, \sqrt{\kappa}v\right) \exp\left(-\mu \frac{u^2+v^2}{4}\right) \tag{8.9}$$

$$W(x,p) \quad \longrightarrow \quad \mathcal{L}_{\sqrt{\kappa}}[W] * \mathcal{L}_{\sqrt{\mu}}[W_0] \tag{8.10}$$

Then, in the contravariant case ($\kappa < 0$), we obtain the following:

$$\chi(z) \quad \longrightarrow \quad \chi(-\sqrt{|\kappa|}z^*) \exp\left(-\mu \frac{|z|^2}{2}\right) \tag{8.11}$$

$$\tilde{W}(u,v) \quad \longrightarrow \quad \tilde{W}\left(\sqrt{|\kappa|}u, -\sqrt{|\kappa|}v\right) \exp\left(-\mu \frac{u^2+v^2}{4}\right) \tag{8.12}$$

$$W(x,p) \quad \longrightarrow \quad \Pi \circ \mathcal{L}_{\sqrt{|\kappa|}}[W] * \mathcal{L}_{\sqrt{\mu}}[W_0] \tag{8.13}$$

Notice that the concatenation of two Gaussian channels is a Gaussian channel, and find the following rule:

$$\mathcal{M}_{\kappa_2}^{(\mu_2)} \circ \mathcal{M}_{\kappa_1}^{(\mu_1)} = \mathcal{M}_{\kappa_1\kappa_2}^{(\mu_1\kappa_2+\mu_2)}. \tag{8.14}$$

As it appears from the above relation, such channels are not commutative.

8.1.1 Optical implementation

Let us now turn our interest towards a physical implementation of these channels, in the framework of quantum optics. We start by building an optical channel with a beam-splitter and a thermal state as follows:

$$\mathcal{E}_\eta^{(\bar{n})}(\hat{\rho}) = \mathrm{Tr}_2\left[\hat{B}_\theta\left(\hat{\rho}\otimes\hat{\tau}_{\bar{n}}\right)\hat{B}_\theta^\dagger\right] \tag{8.15}$$

where $\theta \in [0, \pi/2]$ is chosen such that $\cos^2\theta = \eta$. Remember that the operator $\hat{B}_\theta$ is defined at (4.55). Then, we can compute its effect on the Wigner function of $\hat{\rho}$ as follows:

$$\mathcal{E}_\eta^{(\bar{n})}[W] = \mathcal{L}_{\sqrt{\eta}}[W] * \mathcal{L}_{\sqrt{1-\eta}}[W_{\hat{\tau}}] \tag{8.16}$$

$$= \mathcal{L}_{\sqrt{\eta}}[W] * \mathcal{L}_{\sqrt{1-\eta}}\left[\mathcal{L}_{\sqrt{2\bar{n}+1}}[W_0]\right] \tag{8.17}$$

$$= \mathcal{L}_{\sqrt{\eta}}[W] * \mathcal{L}_{\sqrt{1-\eta}} \circ \mathcal{L}_{\sqrt{2\bar{n}+1}}[W_0] \tag{8.18}$$

$$= \mathcal{L}_{\sqrt{\eta}}[W] * \mathcal{L}_{\sqrt{(1-\eta)(2\bar{n}+1)}}[W_0] \tag{8.19}$$

so that we can identify $\mathcal{E}_\eta^{(\bar{n})} = \mathcal{M}_\kappa^{(\mu)}$ where $\kappa = \eta$ and $\mu = (1-\eta)(2\bar{n}+1)$. The channel is quantum limited when $\bar{n} = 0$ so that $\mu = 1-\eta$, and in that case we denote it as $\mathcal{E}_\eta$. We now construct a similar channel with a two-mode squeezer rather than a beam-splitter:

$$\mathcal{A}_g^{(\bar{n})}(\hat{\rho}) = \mathrm{Tr}_2\left[\hat{\Sigma}_r\left(\hat{\rho}\otimes\hat{\tau}_{\bar{n}}\right)\hat{\Sigma}_r^\dagger\right] \tag{8.20}$$

where r is chosen such that $g = \cosh^2 r$. Remember that the operator $\hat{\Sigma}_r$ is defined at (4.67). The resulting Wigner function can be computed as the following:

$$\mathcal{A}_g^{(\bar{n})}[W] = \mathcal{L}_{\sqrt{g}}[W] * \mathcal{L}_{\sqrt{g-1}}[W_{\hat{\tau}}] \tag{8.21}$$

$$= \mathcal{L}_{\sqrt{g}}[W] * \mathcal{L}_{\sqrt{g-1}}\left[\mathcal{L}_{\sqrt{2\bar{n}+1}}[W_0]\right] \tag{8.22}$$

$$= \mathcal{L}_{\sqrt{g}}[W] * \mathcal{L}_{\sqrt{g-1}} \circ \mathcal{L}_{\sqrt{2\bar{n}+1}}[W_0] \tag{8.23}$$

$$= \mathcal{L}_{\sqrt{g}}[W] * \mathcal{L}_{\sqrt{(g-1)(2\bar{n}+1)}}[W_0] \tag{8.24}$$

so that we can identify $\mathcal{A}_g^{(\bar{n})} = \mathcal{M}_\kappa^{(\mu)}$ where $\kappa = g$ and $\mu = (g-1)(2\bar{n}+1)$. The channel is quantum limited when $\bar{n} = 0$ so that $\mu = g-1$, and in that case we denote it as $\mathcal{A}_g$. A contravariant version of that channel can be built if we partial trace over the first mode rather than the second one:

$$\tilde{\mathcal{A}}_g^{(\bar{n})}(\hat{\rho}) = \mathrm{Tr}_1\left[\hat{\Sigma}_r\left(\hat{\rho}\otimes\hat{\tau}_{\bar{n}}\right)\hat{\Sigma}_r^\dagger\right] \tag{8.25}$$

where r is chosen such that $g = \cosh^2 r$. The Wigner function of the output can then be found as follows:

$$\tilde{\mathcal{A}}_g^{(\bar{n})}[W] = \Pi \circ \mathcal{L}_{\sqrt{g-1}}[W] * \mathcal{L}_{\sqrt{g}}[W_{\hat{\tau}}] \tag{8.26}$$

$$= \Pi \circ \mathcal{L}_{\sqrt{g-1}}[W] * \mathcal{L}_{\sqrt{g}}\left[\mathcal{L}_{\sqrt{2\bar{n}+1}}[W_0]\right] \tag{8.27}$$

$$= \Pi \circ \mathcal{L}_{\sqrt{g-1}}[W] * \mathcal{L}_{\sqrt{g}} \circ \mathcal{L}_{\sqrt{2\bar{n}+1}}[W_0] \tag{8.28}$$

$$= \Pi \circ \mathcal{L}_{\sqrt{g-1}}[W] * \mathcal{L}_{\sqrt{g(2\bar{n}+1)}}[W_0] \tag{8.29}$$

So that we can identify $\tilde{\mathcal{A}}_g^{(\bar{n})} = \mathcal{M}_\kappa^{(\mu)}$ where $\kappa = 1 - g$ and $\mu = g(2\bar{n} + 1)$. The channel is quantum limited when $\bar{n} = 0$ so that $\mu = g$, and in that case we simply denote the channel as $\tilde{\mathcal{A}}_g$.

Finally, let us notice that the optical channels that we have introduced until now do not provide us with the case $\kappa = 1$ and $\mu > 1$. Such a channel adds a Gaussian noise to the Wigner function, and we define the additive noise channel $\mathcal{N}_\mu$ as $\mathcal{M}_\kappa^{(\mu)}$ for $\kappa = 1$ and $\mu > 0$. We are going to show that it can be constructed from the successive application of two optical channels that we have introduce earlier. Let us consider a quantum channel built from a quantum-limited lossy channel $\mathcal{E}_\eta$ and a quantum-limited amplifier channel $\mathcal{A}_g$:

$$\mathcal{A}_g \circ \mathcal{E}_\eta [W] = \mathcal{A}_g \left[\mathcal{L}_{\sqrt{\eta}} [W] * \mathcal{L}_{\sqrt{1-\eta}} [W_0] \right] \tag{8.30}$$

$$= \mathcal{L}_{\sqrt{g}} \left[\mathcal{L}_{\sqrt{\eta}} [W] * \mathcal{L}_{\sqrt{1-\eta}} [W_0] \right] * \mathcal{L}_{\sqrt{g-1}} [W_0] \tag{8.31}$$

$$= \mathcal{L}_{\sqrt{g}} \circ \mathcal{L}_{\sqrt{\eta}} [W] * \mathcal{L}_{\sqrt{g}} \circ \mathcal{L}_{\sqrt{1-\eta}} [W_0] * \mathcal{L}_{\sqrt{g-1}} [W_0] \tag{8.32}$$

$$= \mathcal{L}_{\sqrt{g\eta}} [W] * \mathcal{L}_{\sqrt{g(1-\eta)}} [W_0] * \mathcal{L}_{\sqrt{g-1}} [W_0] \tag{8.33}$$

$$= \mathcal{L}_{\sqrt{g\eta}} [W] * \mathcal{L}_{\sqrt{g(1-\eta)+g-1}} [W_0] \tag{8.34}$$

Then, if we set $g\eta = 1$, the equation becomes more simply $\mathcal{A}_g \circ \mathcal{E}_{1/g} [W] = W * \mathcal{L}_{\sqrt{2(g-1)}} [W_0]$. We can thus identify $\mathcal{N}_\mu = \mathcal{A}_g \circ \mathcal{E}_{1/g}$ with $\mu = 2(g-1)$. The different channels are plotted in Figure 8.1.

8.2 Continuous majorization relations

At this point, we have introduced Gaussian single-mode phase-insensitive channels and we have defined them with the appropriate tools to work in phase space. In this section, we present two different observations. First, we show a simple condition for which such a channel always yields a Wigner-positive output. Then, we show that for special value of the gain/loss parameter g, the output of the channel is always majorized by its inuput. We also extend that property to two-mode squeezers.

8.2.1 Wigner-positiving channel

Gaussian distributions have the interesting property that the convolution of two Gaussian distributions gives another Gaussian distributions, as we have highlighted in Chapter 2. We are going to exploit that property in order to provide a different development of the quantum channel $\mathcal{M}_\kappa^{(\mu)}$. Let us first consider the covariant case, in which we can write:

$$\mathcal{M}_\kappa^{(\mu)} [W] = \mathcal{L}_{\sqrt{\kappa}} [W] * \mathcal{L}_{\sqrt{\mu}} [W_0] \tag{8.35}$$

$$= \mathcal{L}_{\sqrt{\kappa}} [W] * \mathcal{L}_{\sqrt{\kappa}} [W_0] * \mathcal{L}_{\sqrt{\mu-\kappa}} [W_0] \tag{8.36}$$

$$= \mathcal{L}_{\sqrt{\kappa}} [W * W_0] * \mathcal{L}_{\sqrt{\mu-\kappa}} [W_0] \tag{8.37}$$

Note that the above development only holds if $\mu \geq \kappa$. Remember now that the convolution of two Wigner functions is always non-negative, as we know from Chapter 5. From that property and the fact that the Wigner function W_0 is itself non-negative, it follows that the output described in (8.37) is non-negative. As a consequence, any covariant channel $\mathcal{M}_\kappa^{(\mu)}$ such that $\mu \geq \kappa$ always gives a Wigner-positive output. Note that for quantum-limited channels, this corresponds to a pure-loss channel $\mathcal{E}_\eta$ with transmittance $\eta \leq 1/2$.

Now, we turn to the contravariant case, so that we consider $\kappa < 0$. A very similar develoment

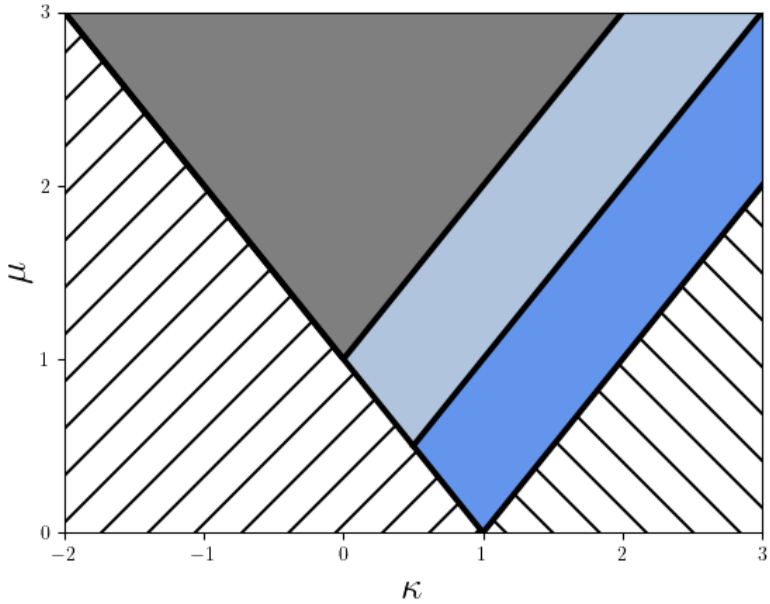

Figure 8.2: The dark-blue region corresponds to channels for which there exist outputs which are both Wigner-negative and entangled (with respect to another system entangled with the input). The light-blue region corresponds to channels for which every output is Wigner-positive, but there exist outputs which are entangled. The grey region corresponds to channels for which every output is Wigner-positive and separable.

gives:

$$\mathcal{M}_{\kappa}^{(\mu)}[W] = \Pi \circ \mathcal{L}_{\sqrt{|\kappa|}}[W] * \mathcal{L}_{\sqrt{\mu}}[W_0] \tag{8.38}$$

$$= \Pi \circ \mathcal{L}_{\sqrt{|\kappa|}}[W] * \mathcal{L}_{\sqrt{|\kappa|}}[W_0] * \mathcal{L}_{\sqrt{\mu-|\kappa|}}[W_0] \tag{8.39}$$

$$= \mathcal{L}_{\sqrt{|\kappa|}}[\Pi[W]] * \mathcal{L}_{\sqrt{|\kappa|}}[W_0] * \mathcal{L}_{\sqrt{\mu-|\kappa|}}[W_0] \tag{8.40}$$

$$= \mathcal{L}_{\sqrt{|\kappa|}}[\Pi[W] * W_0] * \mathcal{L}_{\sqrt{\mu-|\kappa|}}[W_0] \tag{8.41}$$

The above development only holds if $\mu \geq |\kappa|$. Note that this condition is always satisfied for contravariant channels, as can be seen from the physicality condition $\mu \geq |\kappa - 1|$. In that case, we can also conclude that the output is Wigner-positive from the same argument stated previously. Summarizing together the covariant and contravariant case, we can say that the channel $\mathcal{M}_{\kappa}^{(\mu)}$ always yields a Wigner-positive output if $\mu \geq |\kappa|$. We will qualify such a channel as *Wigner-positiving*.

It is instructive to compare the condition we have just derived with another condition that we are going to present. It is known that for values of κ and μ such that $\mu \geq |\kappa| + 1$, the action of the two-mode channel $\mathcal{M}_{\kappa}^{(\mu)} \otimes \mathcal{I}$ (where $\mathcal{I}$ operates the identity) over any two-mode state always yields a separable state [23]. Such a channel $\mathcal{M}_{\kappa}^{(\mu)}$ is said to be entanglement-breaking.

We plot on Figure 8.2 the different regions in the plane (κ, μ) associated to the conditions we have mention above. It is interesting to notice that whereas there exists channels which are Wigner-positiving and not entanglement-breaking, any entanglement-breaking channel is necessarily Wigner-positiving.

8.2.2 Continuous majorization relations in a two-mode squeezer

Let us now focus to continuous majorization relations as such. To that purpose, let us start by presenting a simple development about the rescaling operator $\mathcal{L}_s$. In particular, let us see how the cumulative integrals of a Wigner function evolve under the action of a rescaling transformation:

$$S_a^{\downarrow}\left(\mathcal{L}_s\left[W\right]\right) = \int_0^a \mathcal{L}_s\left[W\right]^{\downarrow}(r) 2\pi r \mathrm{d}r \tag{8.42}$$

$$= \int_0^a \mathcal{L}_s\left[W^{\downarrow}\right](r) 2\pi r \mathrm{d}r \tag{8.43}$$

$$= \int_0^a \frac{1}{s^2} W^{\downarrow}\left(\frac{r}{s}\right) 2\pi r \mathrm{d}r \tag{8.44}$$

$$= \int_0^{|s|a} W^{\downarrow}\left(r'\right) 2\pi r' \mathrm{d}r' \tag{8.45}$$

$$= S_{|s|a}^{\downarrow}(W) \tag{8.46}$$

A very similar development yields the relation $S_a^{\uparrow}\left(\mathcal{L}_s\left[W\right]\right) = S_{|s|a}^{\uparrow}(W)$. Remember now that $S_a^{\downarrow}(W)$ and $S_a^{\uparrow}(W)$ are respectively increasing and decreasing for increasing values of a. This enables to formulate the following majorization relations : $\mathcal{L}_s\left[W\right] \succ W$ when $|s| \leq 1$, and $\mathcal{L}_s\left[W\right] \prec W$ when $|s| \geq 1$ (which is the case for a two-mode squeezer). Notice that this holds for both Wigner-positive and Wigner-negative states. In addition to this, we find as a direct implication of Lemma 2, convolving a Wigner-function W with a non-negative distribution gives a new function W' which is majorized by W. Put together, these two observations allow us to write:

$$\mathcal{M}_{\kappa}^{(\mu)}\left[W\right] \prec W \qquad\qquad \forall \kappa : \ |\kappa| \geq 1 \tag{8.47}$$

where we have used the relation $\mathcal{M}_{\kappa}^{(\mu)}\left[W\right] = \mathcal{L}_{\sqrt{\kappa}}\left[W\right] * \mathcal{L}_{\sqrt{\mu}}\left[W_0\right]$ (with an extra Π for contravariant channels), together with the non-negativity of W_0. This is valid both for Wigner-positive and Wigner-negative states. It is possible to go one step further and generalize that result to any two-mode squeezer fed by a Wigner-positive state on one mode. Indeed, let us consider a two-mode squeezer fed by a state $\hat{\rho}$ with Wigner function W, and with another Wigner-positive state $\hat{\rho}_+$ with Wigner function W_+, as illustrated in Figure 8.3. The two outputs reads as

$$\begin{aligned} W_1' &= \mathcal{L}_{\sqrt{g}}\left[W\right] * \Pi \circ \mathcal{L}_{\sqrt{g-1}}\left[W_+\right] \\ W_2' &= \Pi \circ \mathcal{L}_{\sqrt{g-1}}\left[W\right] * \mathcal{L}_{\sqrt{g}}\left[W_+\right] \end{aligned} \tag{8.48}$$

Using the same arguments we have used previously, we can find that $W \succ W_1'$ for any value of the gain g (which is always greater than 1). Moreover, when the gain is greater than two so that $g \geq 2$, we also have $W \succ W_2'$.

Let us conclude by observing that this continuous majorization relations linking the input and output of a two-mode squeezer can not be extended to a beam-splitter fed by a Wigner-positive state. Indeed, in the case of the two-mode squeezer, we were successful because the rescaling operation and the convolution with a Wigner-positive state are both contributing to disordering the input Wigner function. In the case of the beam-splitter, however, the rescaling operation contributes to decrease the disorder of the input Wigner function, while the convolution with a Wigner-positive state increases its disorder. It is therefore not possible to derive a continuous majorization relation in that case.

8.3 Discrete majorization ladder in state space

As announced, we will adopt in this section a vision in state space, contrary to what we have been used to do until now. More precisely, we are interested in the density operators corresponding to the

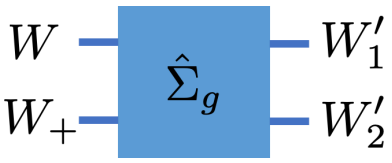

Figure 8.3: A two-mode squeezer of gain g is fed on mode 1 by a state with Wigner function W and on mode 2 by a Wigner-positive state with Wigner function W_+. The first output is W_1' and the second output is W_2', as defined in (8.48). For any value of the gain $g \geq 1$, we have the relation $W \succ W_1'$. For any value of the gain $g \geq 2$, we have the relation $W \succ W_2'$. Notice it is only required that W_+ is non-negative. The Wigner functions W, W_1', W_2' can be partly negative, and the majorization relations still hold.

outputs of Gaussian channels which are fed by Fock states. The present work echoes reference [20, 21], in which a discrete majorization relation between the output of Fock states through a pure-loss channel $\mathcal{E}_\eta$ or a quantum-limited amplifier $\mathcal{A}_g$ is highlighted. The result reads as follows:

$$\mathcal{E}_\eta\left(|n\rangle\langle n|\right) \succ \mathcal{E}_\eta\left(|n+1\rangle\langle n+1|\right), \tag{8.49}$$

$$\mathcal{A}_g\left(|n\rangle\langle n|\right) \succ \mathcal{A}_g\left(|n+1\rangle\langle n+1|\right), \tag{8.50}$$

where the majorization relation has to be understood as follows: the vector of eigenvalues of $\mathcal{E}_\eta(|n\rangle\langle n|)$ majorizes the vetor of eigenvalues of $\mathcal{E}_\eta(|n+1\rangle\langle n+1|)$ (and similalry for the channel $\mathcal{A}_g$). The proof of the result is based on the derivation of a column-stochastic matrix linking the components of the two outputs. Indeed, a sufficient condition for the majorization relation $\mathbf{x} \succ \mathbf{y}$ where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{\mathbb{N}}$ is the existence of a column-stochastic matrix $\mathbf{D} \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ such that $\mathbf{y} = \mathbf{D}\mathbf{x}$ [21]. Note that the matrix $\mathbf{D}$ is said to be column-stochastic if and only if each element is non-negative ($D_{ij} \geq 0$), each column sums up to 1 ($\sum_i D_{ij} = 1$) and each rows sums up to less or equal to 1 ($\sum_j D_{ij} \leq 1$).

In the rest of this section, we will present an extension of result (8.50) to the channels $\mathcal{A}_g^{(\bar{n})}$ and $\mathcal{E}_\eta^{(\bar{n})}$. It is an extension sbeyond the quantum-limited case. The proof of the result is based on the decomposition of $\mathcal{A}_g^{(\bar{n})}$ and $\mathcal{E}_\eta^{(\bar{n})}$ as Gaussian unitaries coupled with thermal states. Then, we succeed in the derivation of a column-stochastic matrix for the two different channels. This was made possible thanks to the use of recurrence relations applying to the Gaussian unitaries present in these transformations [35, 36].

In order to come to this result, we will first define the notion of transition probabilies in the general case of a unitary fed by Fock states. We will then define the generating function of the sequence determined by these transition probabilities. After that, we will particularize our development to the Gaussian unitaries we are interested in, namely the beam-splitter and the two-mode squeezer. The next step is to introduce the recurrence relations acting on these unitaries, which will allow us to build the column-stochastic matrix and conclude our proof.

8.3.1 Transition probabilities in Fock space

As we explained in Chapter 4, the set of Fock states forms a complete orthogonormal basis, so that any two-mode unitary operator $\hat{U}$ can be completely described by the transition amplitudes $\langle n, m|\,\hat{U}\,|i, j\rangle$, where we use the notation $|i, j\rangle = |i\rangle \otimes |j\rangle$ where $|i\rangle$ and $|j\rangle$ are respectively the i^{th} and j^{th} Fock state. We then define the transition probability $U_{n,m}^{(i,j)}$ as:

$$U_{n,m}^{(i,j)} = \left|\langle n, m|\,\hat{U}\,|i, j\rangle\right|^2 \tag{8.51}$$

$$= \mathrm{Tr}\left[\hat{U}\left(|i\rangle\langle i| \otimes |j\rangle\langle j|\right)\hat{U}^\dagger\left(|n\rangle\langle n| \otimes |m\rangle\langle m|\right)\right] \tag{8.52}$$

The channels that we are studying discard the second output of the unitary. Therefore, we will give a particular interest to the reduced transition probability $U_n^{(i,j)}$, which corresponds to the probability

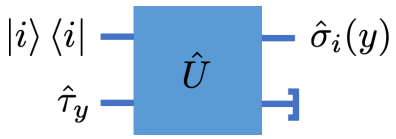

Figure 8.4: The state $\hat{\sigma}_i(y)$ is obtained as described in (8.56). In the particular case where $\hat{U}$ is chosen to be the unitary of the beam-splitter or the two-mode squeezer, the state $\hat{\sigma}_i(y)$ corresponds to respectively $\mathcal{E}_\eta^{(\bar{n})}(|i\rangle\langle i|)$ or $\mathcal{A}_g^{(\bar{n})}(|i\rangle\langle i|)$ where $\bar{n} = y/(1-y)$.

to measure n photons at the first output, whatever the second output. It is defined as follows:

$$U_n^{(i,j)} = \sum_{m=0}^{\infty} U_{n,m}^{(i,j)} \tag{8.53}$$

$$= \mathrm{Tr}\left[\hat{U}\left(|i\rangle\langle i| \otimes |j\rangle\langle j|\right)\hat{U}^\dagger\left(|n\rangle\langle n| \otimes \hat{\mathbb{1}}\right)\right] \tag{8.54}$$

In this chapter, we are interested in channels that can be described as a Gaussian unitary coupled with a thermal environment. Recall that a thermal $\hat{\tau}_y$ state is a mixture of Fock states determined by a thermal parameter y (or equivalently by its mean number of photons $\bar{n}$), as defined in Chapter 4. We now define the reduced thermal transition probability $T_n^{(i)}$ as the probability to measure n photons at the output of a unitary fed by i photons and a thermal state of parameter y:

$$T_n^{(i)}(y) = \mathrm{Tr}\left[\hat{U}\left(|i\rangle\langle i| \otimes \hat{\tau}_y\right)\hat{U}^\dagger\left(|n\rangle\langle n| \otimes \hat{\mathbb{1}}\right)\right] \tag{8.55}$$

Finally, let us define a particular output which is built from the thermal transition probabilities. This enables us to define a particular set of states that we will write $\hat{\sigma}_i(y)$:

$$\hat{\sigma}_i(y) = \mathrm{Tr}_2\left[\hat{U}\left(|i\rangle\langle i| \otimes \hat{\tau}_y\right)\hat{U}^\dagger\right] \tag{8.56}$$

$$= \sum_{n=0}^{\infty} T_n^{(i)}(y)\,|n\rangle\langle n| \tag{8.57}$$

For the sake of brevity we will sometimes simply denote that state as $\hat{\sigma}_i$. Figure 8.4 represents how to construct the state. When $\hat{U}$ is chosen to the unitary of the beam-splitter or the two-mode squeezer, the states $\hat{\sigma}_i$ corresponds respectively to the states $\mathcal{E}_\eta^{(\bar{n})}(|i\rangle\langle i|)$ and $\mathcal{A}_g^{(\bar{n})}(|i\rangle\langle i|)$.

8.3.2 Generating function

The generating function of a sequence $\{c_n\}$ with $n \in \mathbb{N}_0$ is defined as:

$$g(x) = \sum_{n=0}^{\infty} c_n x^n \tag{8.58}$$

The generating function $g(x)$ encapsulates all the information of the sequence $\{c_n\}$ since the value each element of the sequence can be computed as the n^{th} evaluated in $x = 0$:

$$c_n = \frac{1}{n!}\frac{\mathrm{d}^n}{\mathrm{d}x^n} g(x)\bigg|_{x=0} \tag{8.59}$$

The generating function naturally generalizes to multi-indices sequences. If we have a N-indices sequence $\{c_{n_1,\ldots,n_N}\}$ with $n_i \in \mathbb{N}_0\ \forall i$, the generating function of that sequence is the multi-parameters function $g(x_1,\ldots,x_N)$:

$$g(x_1,\ldots,x_N) = \sum_{n_1=0}^{\infty} x_1^{n_1} \cdots \sum_{n_N=0}^{\infty} x_N^{n_N}\ \ c_{n_1,\ldots,n_N} \tag{8.60}$$

Similarly to the 1-parameter case, the multi-parameters generating function encapsulates all the information of the multi-indices sequence.

$$c_{n_1,...,n_N} = \prod_{i=1}^{N} \frac{1}{n_i!} \frac{\partial^{n_i}}{\partial x_i^{n_i}} g(x_1, ..., x_N)\Big|_{x_1,...,x_N=0} \tag{8.61}$$

We then naturally define the generating function of transition probabilities as:

$$f(x, y, z, w) = \sum_{i=0}^{\infty} x^i \sum_{j=0}^{\infty} y^j \sum_{n=0}^{\infty} z^n \sum_{m=0}^{\infty} w^m \; U_{n,m}^{(i,j)} \tag{8.62}$$

where we have associated the parameters x, y, z, w respectively to the indices i, j, n, m. We define the reduced generating function as the generating function of the reduced transition probabilities, which is easily derived from the generating function:

$$f(x, y, z) = \sum_{i=0}^{\infty} x^i \sum_{j=0}^{\infty} y^j \sum_{n=0}^{\infty} z^n \; U_n^{(i,j)} \tag{8.63}$$

$$= \sum_{i=0}^{\infty} x^i \sum_{j=0}^{\infty} y^j \sum_{n=0}^{\infty} z^n \sum_{m=0}^{\infty} \; U_{n,m}^{(i,j)} \tag{8.64}$$

$$= f(x, y, z, 1) \tag{8.65}$$

Note that we use the same writing for the generating function and the reduced generating function since they are unequivocally distinguished by their number of arguments. Let us now focus our attention on the generating function of thermal transition probabilities, that we define as:

$$h_y(x, z) = \sum_{i=0}^{\infty} x^i \sum_{n=0}^{\infty} z^n \; T_n^{(i)}(y). \tag{8.66}$$

By a simple development, it is possible to link the reduced generating function to the thermal generating function.

$$\begin{aligned}
h_y(x, z) &= \sum_{i=0}^{\infty} x^i \sum_{n=0}^{\infty} z^n \; T_n^{(i)}(y) \\
&= \sum_{i=0}^{\infty} x^i \sum_{n=0}^{\infty} z^n \, \mathrm{Tr}\left[\hat{U} \left(|i\rangle \langle i| \otimes \hat{\tau}_y\right) \hat{U}^\dagger \left(|n\rangle \langle n| \otimes \hat{\mathbb{1}}\right)\right] \\
&= \sum_{i=0}^{\infty} x^i \sum_{n=0}^{\infty} z^n \, \mathrm{Tr}\left[\hat{U} \left(|i\rangle \langle i| \otimes (1-y) \sum_{j=0}^{\infty} y^j \, |j\rangle \langle j|\right) \hat{U}^\dagger \left(|n\rangle \langle n| \otimes \hat{\mathbb{1}}\right)\right] \\
&= (1-y) \sum_{i=0}^{\infty} x^i \sum_{j=0}^{\infty} y^j \sum_{n=0}^{\infty} z^n \, \mathrm{Tr}\left[\hat{U} \left(|i\rangle \langle i| \otimes |j\rangle \langle j|\right) \hat{U}^\dagger \left(|n\rangle \langle n| \otimes \hat{\mathbb{1}}\right)\right] \\
&= (1-y) \sum_{i=0}^{\infty} x^i \sum_{j=0}^{\infty} y^j \sum_{n=0}^{\infty} z^n \; U_n^{(i,j)} \\
&= (1-y) f(x, y, z)
\end{aligned} \tag{8.67}$$

We have now properly defined the generating function $f(x, y, z, w)$, the reduced generating function $f(x, y, z)$ and the thermal generating function $h_y(x, z)$. Their relation can be summarized as follows:

$$f(x, y, z, 1) = f(x, y, z) = \frac{1}{1-y} h_y(x, z) \tag{8.68}$$

8.3.3 Application to Gaussian unitaries

Let us now illustrate our previous findings by considering the two Gaussian unitaries we are interested in.

Beam-splitter

In Chapter 4, we introduced the unitary of the beam-splitter as $\hat{B}_\theta$ where θ is a real parameter. In what follows, we will refer to that unitary by the notation $\hat{B}_\eta$ where η is the transmittance of the beam-splitter and is related to θ as $\eta = \cos^2\theta$. The beam-splitter transition probabilities are then defined as:

$$\left(U_\eta^{\mathrm{BS}}\right)_{n,m}^{(i,j)} = \left|\langle n,m|\,\hat{B}_\eta\,|i,j\rangle\right|^2 \tag{8.69}$$

Note that the beam-splitter conserve the total photon number, so that these transitions probabilities are non-zero only if $n+m = i+j$. The generating function of the beam-splitter transition probabilities is derived in [35, 36]:

$$f_\eta^{\mathrm{BS}}(x,y,z,w) = \frac{1}{1-\eta(xz+yw)-(1-\eta)(xw+yz)+xyzw} \tag{8.70}$$

$$f_\eta^{\mathrm{BS}}(x,y,z) = \frac{1}{1-\eta(xz+y)-(1-\eta)(x+yz)+xyz} \tag{8.71}$$

$$h_y^{\mathrm{BS}}(x,z) = \frac{1-y}{1-\eta(xz+y)-(1-\eta)(x+yz)+xyz} \tag{8.72}$$

Two-mode squeezer

Let us now turn our attention to the two-mode squeezer. We have introduced its unitary in Chapter 4 as $\hat{\Sigma}_r$ where r is a real parameter. In what follows, we will use the parameter $\lambda = \tanh^2 r$ to define the two-mode squeezer, and we will denote its unitary as $\hat{\Sigma}_\lambda$. Note that the parameter λ is related to the gain g as $\lambda = (g-1)/g$. We then define the transition probabilites of a two-mode squeezer as follows:

$$\left(U_\lambda^{\mathrm{TMS}}\right)_{n,m}^{(i,j)} = \left|\langle n,m|\,\hat{\Sigma}_\lambda\,|i,j\rangle\right|^2. \tag{8.73}$$

The two-mode squeezer has the property to conserve the photon-difference between each mode, so that the transition probabilities are non-zero only if $n-m = i-j$. The generating function of the two-mode squeezer is derived in [35, 36]:

$$f_\lambda^{\mathrm{TMS}}(x,y,z,w) = \frac{1-\lambda}{1-\lambda(xy+zw)-(1-\lambda)(xz+yw)+xyzw} \tag{8.74}$$

$$f_\lambda^{\mathrm{TMS}}(x,y,z) = \frac{1-\lambda}{1-\lambda(xy+z)-(1-\lambda)(xz+y)+xyz} \tag{8.75}$$

$$h_y^{\mathrm{TMS}}(x,z) = \frac{(1-\lambda)(1-y)}{1-\lambda(xy+z)-(1-\lambda)(xz+y)+xyz} \tag{8.76}$$

Notice the interesting link between f_λ and f_η:

$$f_\lambda^{\mathrm{TMS}}(x,y,z,w) = (1-\lambda) f_{1-\lambda}^{\mathrm{BS}}(x,w,z,y) \tag{8.77}$$

That relation is at the origin of a partial-time reversal symmetry, recently exploited in [13]. We now have introduced all the theory we need to proceed to the proof of our main result.

8.3.4 From recurrence relations to a column-stochastic matrix

Remember that we want to prove the two following relations:

$$\mathcal{A}_g^{(\bar{n})}\left(|i\rangle\langle i|\right) \succ \mathcal{A}_g^{(\bar{n})}\left(|i+1\rangle\langle i+1|\right) \tag{8.78}$$

$$\mathcal{E}_\eta^{(\bar{n})}\left(|i\rangle\langle i|\right) \succ \mathcal{E}_\eta^{(\bar{n})}\left(|i+1\rangle\langle i+1|\right) \tag{8.79}$$

In the light of the previous subsection, that statement can be formulated equivalently as $\hat{\sigma}_i \succ \hat{\sigma}_{i+1}$ and should be proven for all y (which is related to $\bar{n}$ as $\bar{n} = y/(1-y)$). Remember that $\hat{\sigma}_i$ is a Fock-diagonal state, so that its eigenvalues are simply its diagonal components in the Fock basis. For that reason, we define the infinite-dimensional vector $\mathbf{t}^{(i)}$ as follows:

$$\mathbf{t}^{(i)} = \left(T_0^{(i)}, T_1^{(i)}, T_2^{(i)}, ...\right)^\intercal, \tag{8.80}$$

where we omit to write the y dependence for clarity. The vector $\mathbf{t}^{(i)}$ is a vector made of the eigenvalues of $\hat{\sigma}^{(i)}$. To show our result, we are going to show that there exist a column-stochastic matrix $\mathbf{D}$ such that:

$$\mathbf{t}^{(i+1)} = \mathbf{D}\ \mathbf{t}^{(i)} \tag{8.81}$$

In general, the matrix $\mathbf{D}$ can depend on multiple parameters, such as the parameters of the channel (y, η or λ), but also on the index i.

Our proof is constructed in two parts. First, we are going to derive a recurrence relation from the generating function of the thermal transition probabilities. That recurrence relation links the components of $\mathbf{t}^{(i+1)}$ to the components of $\mathbf{t}^{(i)}$. Then, we are going to build a matrix linking $\mathbf{t}^{(i+1)}$ to $\mathbf{t}^{(i)}$. We will finally verify that the matrix is column-stochastic.

Recurrence relation

One of the powerful uses of the generating function is that it can be used to derive recurrence relations straightforwardly. This has been initiated in the context of quantum optics by Jabbour in [35, 36]. To illustrate this, let us consider the thermal transition probabilities $T_n^{(i)}$ and their generating function $h(x,z)$. The parameters x, z are respectively associated to the indices i, n. If the generating function is multiplied by some parameter x or z, the sequence will be shifted along the associated index:

$$\begin{aligned} h(x,z) &\longleftrightarrow T_n^{(i)} \\ x\ h(x,z) &\longleftrightarrow T_n^{(i-1)} \\ z\ h(x,z) &\longleftrightarrow T_{n-1}^{(i)} \\ xz\ h(x,z) &\longleftrightarrow T_{n-1}^{(i-1)} \end{aligned} \tag{8.82}$$

Note that we use the convention that transition probabilities are zero for negative indices, so that $T_n^{(i)} = 0$ as soon as $n < 0$ or $i < 0$. We are going to use this corresponding rules to derive recurrence relations for the thermal transition probabilities. Firstly for the beam-splitter, and then for the two-mode squeezer.

Beam-splitter

Starting from the expression of $h_y^{\mathrm{BS}}(x,z)$, we obtain the following relation:

$$(1-\eta y)\ h(x,z) + (\eta-1)x\ h(x,z) + y(\eta-1)z\ h(x,z) + (y-\eta)xz\ h(x,z) = 1-y, \tag{8.83}$$

where we simply note $h_y^{\mathrm{BS}}(x,z)$ as $h(x,z)$ for readability. We can rewrite that expression as:

$$h(x,z) = \frac{1-\eta}{1-\eta y}\ x\ h(x,z) + \frac{y(1-\eta)}{1-\eta y}\ z\ h(x,z) + \frac{\eta-y}{1-\eta y}\ xz\ h(x,z) + \frac{1-y}{1-\eta y}. \tag{8.84}$$

Identifying the sequence to their generating function in the latter relation gives us the following recurrence relation:

$$T_n^{(i)} = \frac{1-\eta}{1-\eta y}\, T_n^{(i-1)} + \frac{y(1-\eta)}{1-\eta y}\, T_{n-1}^{(i)} + \frac{\eta - y}{1-\eta y}\, T_{n-1}^{(i-1)} + \frac{1-y}{1-\eta y}\, \delta_n^{(i)} \tag{8.85}$$

where $\delta_n^{(i)}$ is the Kronecker delta, which is equal to 1 if $i = n = 0$ and is zero otherwise. The generating function of $\delta_n^{(i)}$ is precisely 1.

Two-mode squeezer

We proceed in a similar fashion for the two-mode squeezer. Starting from the expression of $h_y^{\mathrm{TMS}}(x,z)$, we derive the following relation:

$$(1-y+\lambda y)\; h(x,z) - \lambda y x\; h(x,z) - \lambda z\; h(x,z) + (\lambda + y - 1) xz\; h(x,z) = (1-y)(1-\lambda) \tag{8.86}$$

where we note $h_y^{\mathrm{TMS}}(x,z)$ as $h(x,z)$ for readability, while it is of course a different function than in Eq. (8.83). That expression is then rewritten as:

$$h(x,z) = \frac{\lambda y}{1-y+\lambda y}\; x\; h(x,z) + \frac{\lambda}{1-y+\lambda y}\; z\; h(x,z) \frac{1-y-\lambda}{1-y+\lambda y}\; xz\; h(x,z) + \frac{(1-y)(1-\lambda)}{1-y+\lambda y} \tag{8.87}$$

which then yields the recurrence relation:

$$T_n^{(i)} = \frac{\lambda y}{1-y+\lambda y}\, T_n^{(i-1)} + \frac{\lambda}{1-y+\lambda y}\, T_{n-1}^{(i)} + \frac{1-y-\lambda}{1-y+\lambda y}\, T_{n-1}^{(i-1)} + \frac{(1-y)(1-\lambda)}{1-y+\lambda y}\, \delta_n^{(i)}. \tag{8.88}$$

We have derived recurrence relations in the case for the beam-splitter and in the case of the two-mode squeezer. We will see that both cases can be treated similarly.

8.3.5 Column-stochastic matrix

It is interesting to note that the recurrence relation of the beam-splitter and the two-mode squeezer have the same form, only the coefficients vary. Indeed, both Eqs. (8.85) and (8.88) can be written as:

$$T_n^{(i)} = \alpha T_n^{(i-1)} + \beta T_{n-1}^{(i)} + \gamma T_{n-1}^{(i-1)} + \chi \delta_n^{(i)}, \tag{8.89}$$

where the coefficient $\alpha, \beta, \gamma, \chi$ are different whether we consider a beam-splitter or a two-mode squeezer. Interestingly, these coefficient do not depend on n nor i.

The term $\delta_n^{(i)}$ is only non-zero when $n = i = 0$. When $n = i = 0$, that term is actually the only non-zero term in the RHS of Eq. (8.89), and χ gives the value of $T_0^{(0)} = \chi$ which is the probability to measure zero photon at the output when no photon is sent at the input. As soon as $i \geq 1$, that term disappears and we are left with:

$$T_n^{(i)} = \alpha T_n^{(i-1)} + \beta T_{n-1}^{(i)} + \gamma T_{n-1}^{(i-1)}. \tag{8.90}$$

Our goal is to express the components of $\mathbf{t}^{(i+1)}$ from the components of $\mathbf{t}^{(i)}$ (or equivalently the components of $\mathbf{t}^{(i)}$ from the components of $\mathbf{t}^{(i-1)}$). This means that we are looking for a relation that expresses $T_n^{(i)}$ only from $T_{n'}^{(i-1)}$ where n' can take different values but we should only have terms in $(i-1)$. We are thus going to reuse the recurrence relation on the term $T_{n-1}^{(i)}$. Proceeding iteratively

gives:

$$\begin{aligned} T_n^{(i)} &= \alpha T_n^{(i-1)} + \gamma T_{n-1}^{(i-1)} + \beta T_{n-1}^{(i)} \\ &= \alpha T_n^{(i-1)} + \gamma T_{n-1}^{(i-1)} + \beta \left(\alpha T_{n-1}^{(i-1)} + \gamma T_{n-2}^{(i-1)} + \beta T_{n-2}^{(i)} \right) \\ &= \alpha T_n^{(i-1)} + (\gamma + \beta\alpha) T_{n-1}^{(i-1)} + \beta\gamma T_{n-2}^{(i-1)} + \beta^2 T_{n-2}^{(i)} \\ &= \alpha T_n^{(i-1)} + (\gamma + \beta\alpha) T_{n-1}^{(i-1)} + \beta\gamma T_{n-2}^{(i-1)} + \beta^2 \left(\alpha T_{n-2}^{(i-1)} + \gamma T_{n-3}^{(i-1)} + \beta T_{n-3}^{(i)} \right) \\ &= \alpha T_n^{(i-1)} + (\gamma + \beta\alpha) T_{n-1}^{(i-1)} + \beta(\gamma + \beta\alpha) T_{n-2}^{(i-1)} + \beta^2\gamma T_{n-3}^{(i-1)} + \beta^3 T_{n-3}^{(i)} \\ &= \cdots \end{aligned} \tag{8.91}$$

Summed up, this comes up to:

$$T_n^{(i)} = \alpha T_n^{(i-1)} + \nu \sum_{k=1}^{n} \beta^{k-1} T_{n-k}^{(i-1)}, \tag{8.92}$$

where we have defined the parameter $\nu = \gamma + \beta\alpha$. That relation links each component of the vector $\mathbf{t}^{(i)}$ to components of the vector $\mathbf{t}^{(i-1)}$, which is precisely what we wanted. We can now construct the following matrix (after shifting i to $i+1$):

$$\underbrace{\begin{pmatrix} T_0^{(i+1)} \\ T_1^{(i+1)} \\ T_2^{(i+1)} \\ T_3^{(i+1)} \\ T_4^{(i+1)} \\ T_5^{(i+1)} \\ \vdots \end{pmatrix}}_{\mathbf{t}^{(i+1)}} = \underbrace{\begin{pmatrix} \alpha & 0 & 0 & 0 & 0 & 0 & \cdots \\ \nu & \alpha & 0 & 0 & 0 & 0 & \cdots \\ \nu\beta & \nu & \alpha & 0 & 0 & 0 & \cdots \\ \nu\beta^2 & \nu\beta & \nu & \alpha & 0 & 0 & \cdots \\ \nu\beta^3 & \nu\beta^2 & \nu\beta & \nu & \alpha & 0 & \cdots \\ \nu\beta^4 & \nu\beta^3 & \nu\beta^2 & \nu\beta & \nu & \alpha & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}}_{\mathbf{D}} \underbrace{\begin{pmatrix} T_0^{(i)} \\ T_1^{(i)} \\ T_2^{(i)} \\ T_3^{(i)} \\ T_4^{(i)} \\ T_5^{(i)} \\ \vdots \end{pmatrix}}_{\mathbf{t}^{(i)}} \tag{8.93}$$

Now that we have built the matrix $\mathbf{D}$, we need to show that it is indeed column-stochastic. Such a matrix only has non-negative entries and its columns sum up to 1. The non-negativity conditions are the following three conditions:

$$\alpha \geq 0, \qquad \nu \geq 0, \qquad \beta \geq 0. \tag{8.94}$$

These conditions have to be verified separately for the beam-splitter and for the two-mode squeezer. The calculations are straightforward given that $0 \leq y, \eta, \lambda \leq 1$. Checking that each column sums up

to 1 comes to the following condition:

$$\begin{aligned} & \alpha + \nu \sum_{k=0}^{\infty} \beta^k = 1 \\ \Leftrightarrow \quad & \alpha + (\gamma + \beta\alpha)\frac{1}{1-\beta} = 1 \\ \Leftrightarrow \quad & (1-\beta)\alpha + \gamma + \beta\alpha = 1 - \beta \\ \Leftrightarrow \quad & \alpha - \beta\alpha + \gamma + \beta\alpha = 1 - \beta \\ \Leftrightarrow \quad & \alpha + \beta + \gamma = 1 \end{aligned} \tag{8.95}$$

That relation can be verified directly by looking at Eqs. (8.85) and (8.88). From (8.94) it is also obvious that the sum of each row is lower or equal to 1.

8.3.6 Extension to mixtures of Fock states

That relation is proven through the construction of a column-stochastic matrix $\mathbf{D}$ which links the two outputs $\mathbf{t}^{(i)}$ and $\mathbf{t}^{(i+1)}$. What is very remarkable about the matrix $\mathbf{D}$ is that it does not depend on i. It is only a function of y (noise of the environment) and η (transmittance of the beam-splitter) or λ (parameter of the two-mode squeezer). This implies the following relation:

$$\mathbf{t}^{(i)} = \mathbf{D}\mathbf{t}^{(i-1)} = \mathbf{D}^2\mathbf{t}^{(i-2)} = \mathbf{D}^k\mathbf{t}^{(i-k)} = \mathbf{D}^i\mathbf{t}^{(0)} \tag{8.96}$$

where k can be chosen from 0 to i.

Let us now notice that any power of column-stochastic $\mathbf{D}^n$ is itself column-stochastic. The same fact can be noticed for mixtures of column-stochastic matrix. Indeed, it is easily shown that if $\mathbf{D}_1$ and $\mathbf{D}_2$ are column-stochastic, then $p_1\mathbf{D}_1 + p_2\mathbf{D}_2$ is column-stochastic for any $p_1, p_2 \geq 0$ such that $p_1 + p_2 = 1$.

We can take advantage of these observations to derive a more general majorization ladder involving mixtures of Fock states. To that purpose, we define a couple of states $\hat{\rho}_k$ and $\hat{\rho}_{k-1}$ defined as mixtures of Fock states with photon-number greater or equal to respectively k and $k-1$ according to a probability distribution $\{p_n\}$:

$$\hat{\rho}_k = \sum_{n=0}^{\infty} p_n \left|k+n\right\rangle \left\langle k+n\right|, \tag{8.97}$$

$$\hat{\rho}_{k-1} = \sum_{n=0}^{\infty} p_n \left|k-1+n\right\rangle \left\langle k-1+n\right|. \tag{8.98}$$

In what follows, we use the symbol $\mathcal{M}$ to represent either of the two channels $\mathcal{A}_g^{(\bar{n})}$ or $\mathcal{E}_\eta^{(\bar{n})}$. When $\hat{\rho}_k$ is sent through $\mathcal{M}$, its output $\mathcal{M}(\hat{\rho}_k)$ is another mixture of Fock states which can be characterized by a vector $\mathbf{q}^{(k)}$, such that $\mathcal{M}(\hat{\rho}_k) = \sum_n q_n^{(k)} \left|n\right\rangle \left\langle n\right|$. Since $\hat{\rho}_k$ is a mixture of Fock states, the vector $\mathbf{q}^{(k)}$ is the corresponding mixture of the vectors $\mathbf{t}^{(i)}$:

$$\mathbf{q}^{(k)} = \sum_{n=0}^{\infty} p_n \mathbf{t}^{(k+n)} \tag{8.99}$$

$$= \sum_{n=0}^{\infty} p_n \mathbf{D}^n \mathbf{t}^{(k)} \tag{8.100}$$

Since a mixture of column-stochastic matrices is column-stochastic, this implies that $\mathbf{q}^{(k)} \prec \mathbf{t}^{(k)}$, which in turn implies the following relation:

$$\mathcal{M}(\hat{\rho}_k) \prec \mathcal{M}(\left|k\right\rangle \left\langle k\right|). \tag{8.101}$$

Now, let us consider the two states $\hat{\rho}_k$ and $\hat{\rho}_{k-1}$ introduced in Eqs. (8.97) and (8.98), where we choose $k \geq 1$. Let $\mathbf{q}^{(k)}$ and $\mathbf{q}^{(k-1)}$ be the vectors associated to the outputs of respectively $\hat{\rho}_k$ and $\hat{\rho}_{k-1}$ through the channel $\mathcal{M}$. We make the following simple development:

$$\mathbf{q}^{(k)} = \sum_{n=0}^{\infty} p_n \mathbf{t}^{(k+n)} \tag{8.102}$$

$$= \sum_{n=0}^{\infty} p_n \mathbf{D}\mathbf{t}^{(k-1+n)} \tag{8.103}$$

$$= \mathbf{D}\mathbf{q}^{(k-1)} \tag{8.104}$$

This directly shows that $\mathbf{q}^{(k)} \prec \mathbf{q}^{(k-1)}$, which in turns implies the following:

$$\mathcal{M}\left(\hat{\rho}_k\right) \prec \mathcal{M}\left(\hat{\rho}_{k-1}\right). \tag{8.105}$$

Chapter 9

Resource theory of local Gaussian work extraction

In this last chapter, we take a step back from the main thread of our thesis, and focus on the study of multipartite quantum systems from a thermodynamic point of view. We remain in the context of quantum optics, with a particular attention to Gaussian states, and therefore we will use many notions introduced in Chapter 4. The content of this chapter has been published in the paper entitled *Quantum thermodynamics in a multipartite setting: A resource theory of local Gaussian work extraction for multimode bosonic systems* [61]. This chapter is not meant to be exhaustive since we leave out some mathematical derivations that the interested reader will find in Reference [61]. Here, we try to convey the general idea of the paper as well as the main results that have been reached.

Thermodynamics is a macroscopic theory applicable in the limit where the number of particles and volume tend to infinity [12]. However, with our increasing ability to control or manipulate small systems and the realization of molecular motors [56, 33, 25] and nanomachines [58, 22, 2, 59, 19], the scope of applicability of thermodynamics is starting to stretch beyond the macroscopic region. One of the main goals of the thermodynamics of small systems—quantum thermodynamics—is the extraction of work by means of cyclic Hamiltonian transformations of a quantum state. Evidently, it is of great importance to know which states do not allow for any work extraction under unitary transformations. Such states are known as passive states [54, 41], and we will give more details about them in Chapter 5. A central result in quantum thermodynamics is that the only completely passive states are the thermal states, which we have introduced in Chapter 4.

A resource theory of thermodynamics can be developed to systematically describe work extraction from a quantum system and, in general, the allowed state transformations are those where the system interacts via an energy-preserving unitary together with an ancilla chosen to be in a thermal Gibbs state (with an arbitrary Hamiltonian) at some fixed temperature [32, 9, 8, 24]. In this resource-theoretic treatment of quantum thermodynamics, the thermal Gibbs state of the system at the same temperature as that of the ancilla is the only free state [32, 9, 8]. Although considering arbitrary Hamiltonians and arbitrary energy-preserving unitaries is satisfying in the context of establishing a general framework for quantum thermodynamics, it may also be interesting to focus on states and unitaries of higher practical relevance. For bosonic systems, for example, restricting to Gaussian states and Gaussian operations has proven to be very fruitful, particularly in the field of quantum information theory with continuous variables [66, 52, 1]. Similarly, exploring quantum thermodynamics with Gaussian bosonic systems is a promising avenue [11], which we investigate here.

In this chapter, we explore a multipartite quantum thermodynamical scenario as illustrated in Figure 9.1, where each party can extract work locally by applying a local unitary and this process is being assisted with a global energy-preserving unitary (hence, allowing no global work extraction as such). This is not a trivial extension of work extraction for a single party because there exist situations where an energy-preserving coupling allows the parties to extract work locally even though their local (reduced) states are initially passive, as we are going to see later in this chapter. Given the definition of passive states, a natural choice may be to consider them (instead of Gibbs states) as free states in a resource theory for extractable work. However, considering passive states as free states

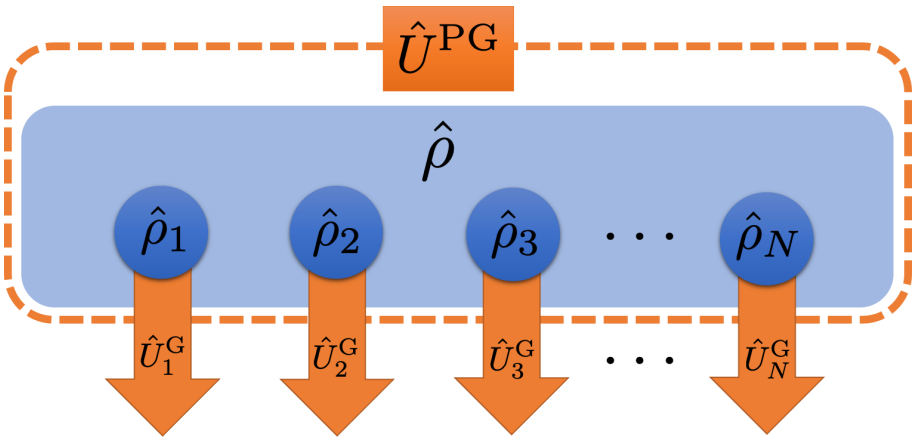

Figure 9.1: Illustration of the setup for our resource theory of local Gaussian work extraction. We consider a multimode state $\hat{\rho}$ which with we can interact globally through the passive Gaussian unitary $\hat{U}^{\mathrm{PG}}$. We then proceed to extract energy locally from each mode through single-mode Gaussian unitaries $\hat{U}_i^{\mathrm{G}}$.

defies a plausible criterion for any reasonable resource theory, namely that if a state $\hat{\rho}$ is free, then $\hat{\rho}^{\otimes n}$ should also be free for any integer n. For this reason, we rather take thermal states as building blocks of our free states for each party, which allows us to develop a multipartite resource theory for local extractable work within this restriction.

This will thus be our starting point in order to build a resource theory for local work extraction under Gaussian unitaries. In the following of this chapter, we will first set up the tools of the symplectic formalism, which will constitute the mathematical backgound of our further developments. After that, we will move on to the rigourous definition of our set of free states and set of free operations. Then, in the last section, we will highlight two different monotones to quantify the resourceful states. As we will show, these two monotones have very different properties, and one of them will lead us to the definition to an extended set of free states.

9.1 Symplectic formalism

In Chapter 4, we have introduced two important objects that are the mean displacement $\bar{\mathbf{q}}$ and covariance matrix $\mathbf{V}$ of a quantum state. The definition of these two objects is not restricted to the framework of Gaussian states, since they are defined in all generality for any quantum state. In that context, we have shown that the evolution of $\bar{\mathbf{q}}$ and $\mathbf{V}$ under a symplectic transformation was easily described through the symplectic matrix $\mathbf{U}$ associated to the symplectic transformation, as follows:

$$\bar{\mathbf{q}} \to \mathbf{U}\bar{\mathbf{q}}, \qquad \mathbf{V} \to \mathbf{U}\mathbf{V}\mathbf{U}^{\intercal}. \tag{9.1}$$

In what follows, we will present several powerful tools associated with the symplectic formalism, namely the Williamson theorem which reduces any covariance matrix to a standardized form, and the Bloch Messiah decomposition which applies to any sympletic matrix.

9.1.1 Williamson theorem

Let us now introduce a theorem which plays a central role within the symplectic formalism. The Williamson theorem states that any covariance matrix $\mathbf{V}$ can be brought to its Williamson form $\mathbf{V}_{\mathrm{s}}$ with the appropriate symplectic transformation. We use the notation $\mathbf{V}_{\mathrm{s}}$ to denote the Williamson form of the covariance matrix $\mathbf{V}$ and will define it a bit later in this subsection. Williamson theorem thus states that for any covariance matrix $\mathbf{V}$ there exists a symplectic matrix $\mathbf{U}_{\mathrm{s}}$ such that:

$$\mathbf{V}_{\mathrm{s}} = \mathbf{U}_{\mathrm{s}}\mathbf{V}\mathbf{U}_{\mathrm{s}}^{\intercal}. \tag{9.2}$$

Let us now properly define the Williamson form of a covariance matrix $\mathbf{V}$, that we denote as $\mathbf{V}_{\mathrm{s}}$:

$$\mathbf{V}_{\mathrm{s}} = \bigoplus_{k=1}^{N} \nu_k \mathbf{I}_2. \tag{9.3}$$

The Williamson form of a covariance matrix is diagonal, and its diagonal is composed of a sequence of eigenvalues following one another in pairs. The eigenvalues ν_k are called the symplectic eigenvalues of $\mathbf{V}$, and differ in general from the eigenvalues of $\mathbf{V}$. They can alternatively be obtained as the absolute values of the complex eigenvalues of the matrix $i\mathbf{\Omega V}$. It is interesting to note that condition (4.83) implies that all the symplectic eigenvalues ν_k are greater or equal to $1/2$, so that we can write the relation $\mathbf{V}_{\mathrm{s}} \geq (1/2)\mathbf{I}_{2N}$.

It should be noted that the set of symplectic eigenvalues $\{\nu_k\}$ is invariant under symplectic transformation, so that the matrix $\mathbf{V}$ and the matrix $\mathbf{UVU}^\intercal$ have the same eigenvalues when $\mathbf{U}$ is symplectic. Let us also define the symplectic trace of a covariance matrix which corresponds to the sum of its symplectic eigenvalues:

$$\mathrm{Str}\left[\mathbf{V}\right] = 2\sum_{i=1}^{N} \nu_k. \tag{9.4}$$

As a consequence to what we have explained, the symplectic trace is invariant under symplectic transformations. Also, we obviously have the equality $\mathrm{Str}\left[\mathbf{V}\right] = \mathrm{Tr}\left[\mathbf{V}_{\mathrm{s}}\right]$. Moreover, the symplectic trace is in general a lower-bound on the trace, so that for any covariance matrix $\mathbf{V}$ we have the relation $\mathrm{Str}\left[\mathbf{V}\right] \leq \mathrm{Tr}\left[\mathbf{V}\right]$.

In the case of Gaussian states, the Williamson form corresponds to the covariance matrix of a product of thermal states. It is then possible to associate each symplectic eigenvalue with a mean number of photons as $\nu = \bar{n} + 1/2$. Also, since the von Neumann entropy is additive under tensoring, we can compute the von Neumann entropy of the product of thermal states $\hat{\gamma}$ as:

$$S(\hat{\gamma}) = \sum_{k=1}^{N} g(\bar{n}_k) \tag{9.5}$$

where the function $g(\bar{n})$ gives the von Neumann entropy of a thermal state with mean number of photon $\bar{n}$ and is defined in (4.89). Also, since von Neumann entropy is invariant under unitary operations, (9.5) can be computed for any Gaussian state $\hat{\gamma}$ (not necessarily a product of thermal states). Note that in a similar fashion, the purity of a multimode Gaussian state can be computed as

$$\mu(\hat{\gamma}) = \prod_{k=1}^{N} \frac{1}{2\nu_k} \tag{9.6}$$

where $\hat{\gamma}$ is any Gaussian state with symplectic eigenvalues $\{\nu_k\}$.

9.1.2 Bloch-Messiah decomposition

In that context, the Bloch-Messiah decomposition is particularly useful as it provides a decomposition of any symplectic matrix with orthogonal matrices and squeezing matrices. An orthogonal matrix is a matrix $\mathbf{O}$ such that $\mathbf{OO}^\intercal = \mathbf{O}^\intercal\mathbf{O} = \mathbf{I}$, and is associated to a passive transformation. Indeed, orthogonal matrices preserve the spectrum of eigenvalues, so that the matrix $\mathbf{V}$ and $\mathbf{OVO}^\intercal$ of the same set of eigenvalues $\{\lambda_k\}$ when $\mathbf{O}$ is orthogonal. As a consequence, the trace is invariant under orthogonal operation and we have the relation $\mathrm{Tr}\left[\mathbf{OVO}^\intercal\right] = \mathrm{Tr}\left[\mathbf{V}\right]$ (which translates into the fact the energy is conserved). Using notation from Chapter 4, orthogonal matrices can be obtained by combining rotations matrices $\mathbf{R}_\theta$ and beam-splitter matrices $\mathbf{B}_\theta$ (which can be understood as a rotation between different modes).

The Bloch-Messiah decomposition of a symplectic matrix $\mathbf{U}$ reads as follows:

$$\mathbf{U} = \mathbf{O}_1 \left(\bigoplus_{k=1}^{N} \mathbf{S}_{r_k} \right) \mathbf{O}_2, \tag{9.7}$$

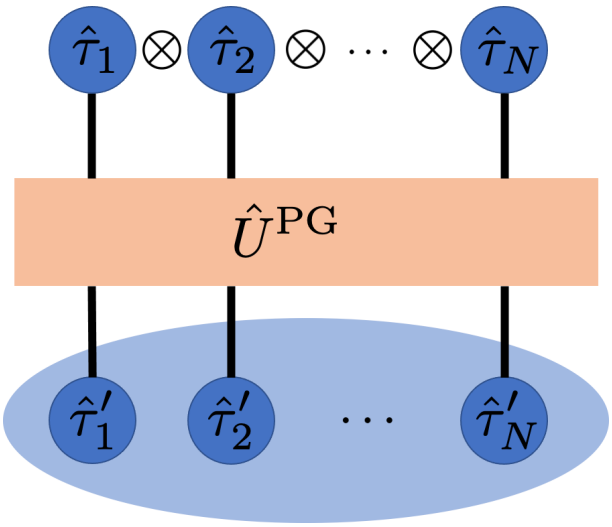

Figure 9.2: Illustration of the setup to build our set of free states. A global passive Gaussian unitary $\hat{U}^{\mathrm{PG}}$ is applied onto a tensor product of thermal states $\hat{\tau}_i$, possibly with different temperatures.

where $\mathbf{O}_1$, $\mathbf{O_2}$ are two orthogonal matrices and $\mathbf{S}_r$ is a single-mode squeezing matrix as defined in Chapter 4. The Bloch-Messiah provides us with a systematic decomposition of any symplectic transformation. As we are going to see, it will play a central role in the following of this chapter.

9.2 Free states ans free operations

A general resource theory comprises two basic elements: the set of free states and the set of free operations. Based on these two elements, the resource states can be identified and the amount of the resource is then quantified with the help of resource monotones which satisfy certain *bona fide* criteria. In the present section we define the set of free states and the set of free operations. We will turn to the definition of monotones in the next section.

9.2.1 Free states

The free states that we want to define should be such that no energy could be extracted from them through the device pictured in Figure 9.1. As discussed in the introduction, our starting point for the set of free state will be to consider products of thermal states, possibly with different temperatures. In addition to this, remember that we are allowed to interact with the system through a global passive unitary. That operation shouldn't create any resource. With this in mind, we define our set of free states as $\mathcal{I}_f$:

$$\mathcal{I}_f = \left\{ \hat{U}^{\mathrm{PG}} \left(\bigotimes_{i=1}^{N} \hat{\tau}_i \right) \hat{U}^{\mathrm{PG}\dagger} \right\}, \tag{9.8}$$

where $\hat{U}^{\mathrm{PG}}$ is a N-mode passive Gaussian unitary and each thermal state $\hat{\tau}_i$ may have different temperatures. Figure 9.2 illustrates how to build such states.

The presence of a global passive unitary is not insignificant. Let us illustrate this by the following example. The output of a two-mode squeezer fed by a two-mode vacuum reads as follows:

$$\hat{\Sigma}_r \left|0,0\right\rangle = \frac{1}{\cosh r} \sum_{n=0}^{\infty} (\tanh r)^n \left|n,n\right\rangle . \tag{9.9}$$

A two-mode squeezed vaccum is a pure state which locally looks like a thermal state. Indeed, if we partial trace it over one mode, we obtain a thermal state of parameter $y = \tanh^2 r$, which is a passive state. From that observation, we see that no energy can be extracted locally from a two-mode squeezed

vacuum. Now, if we apply a beam-splitter with parameter $\theta = \pi/4$, we obtain the following state:

$$\hat{B}_{\frac{\pi}{4}} \hat{\Sigma}_r \left|0,0\right\rangle = \hat{S}_r \left|0\right\rangle \otimes \hat{S}_{-r} \left|0\right\rangle, \tag{9.10}$$

which is actually a tensor product of two squeezed states with orthogonal squeezing. In that situation, it is possible to extract energy locally, since a squeezed state can be unsqueezed with a Gaussian unitary in order to produce a vacuum state, with lower energy. We understand here that the global unitary plays a decisive role, even if as a passive unitary it does not allow to extract energy. By allowing ourselves to act with a global unitary on the multipartite state, we allow ourselves to extract energy from states for which it seemed, at first glance, impossible to do so.

Now, notice that since our free states are Gaussian, it is sufficient to consider their covariance matrices. It is shown in [61] that the covariance matrices of free states take the following form:

$$\mathbf{V}_f \quad = \quad \mathbf{O} \left(\bigoplus_{i=1}^{N} \nu_i \mathbf{I}_2 \right) \mathbf{O}^\intercal \quad = \quad \begin{pmatrix} a_{11}\mathbf{I}_2 & a_{12}\mathbf{R}_{12} & \cdots & a_{1N}\mathbf{R}_{1N} \\ a_{12}\mathbf{R}_{12}^\intercal & a_{22}\mathbf{I}_2 & \cdots & a_{2N}\mathbf{R}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1N}\mathbf{R}_{1N}^\intercal & a_{2N}\mathbf{R}_{2N}^\intercal & \cdots & a_{NN}\mathbf{I}_2 \end{pmatrix} \tag{9.11}$$

where a_{ij} are some coefficients and $\mathbf{R}_{ij}$ are rotation matrices. Free states are Gaussian and are locally thermal. In the following, we will denote the set of all covariance matrices associated to some free state as $\mathcal{V}_f$. An important observation that can be made about this set is that it is a convex set. Indeed, for any $\mathbf{V}_1, \mathbf{V}_2 \in \mathcal{V}_f$, it can be shown that $p_1\mathbf{V}_1 + p_2\mathbf{V}_2 \in \mathcal{V}_f$ where (p_1, p_2) is a probability vector.

9.2.2 Free operations

Now that we have defined our set of free states $\mathcal{I}_f$, let us define our set of free operation Λ_f. Free operations are operations which shouldn't create any resource. As a consequence, when a free states undergoes a free operation, the resulting state should remain a free state. In what follows, we are going to present the free operations that we consider, and we will verify that they are consistent with our set of free states $\mathcal{I}_f$.

Linear interferometer

The first free operation we consider is the action of a global passive Gaussian unitary $\hat{U}^{\mathrm{PG}}$, also called linear interferometer. From the construction of the set of our free states, it is straightforward to see that such a unitary applies a free state on a free state. We prove that a linear interferometer acting on a free state produce another free state in Figure 9.3.

Tensoring of free states

Then, the second free operation we consider is the tensoring of free states. This operation seems natural to consider as free, since it simply means that it is not possible to create a resource by gathering states that do not contain any. And indeed, it can easily be shown that tensoring free states does not produce a resourceful state, as we illustrate in Figure 9.4.

Partial tracing

The last free operation we consider is partial tracing. Again, this is a natural operation, since it means that a multipartite system that does not contain resource cannot give rise to a resourceful subsystem. We can show that partial tracing a free state gives a free state by making two observations. First, partial tracing a Gaussian state yields a Gaussian state, so that we know that the state remain Gaussian after partial tracing. Then, by looking at the form of the covariance matrices of free state (see (9.11)), we can deduce that the resulting covariance matrix will be associated to the one of a free

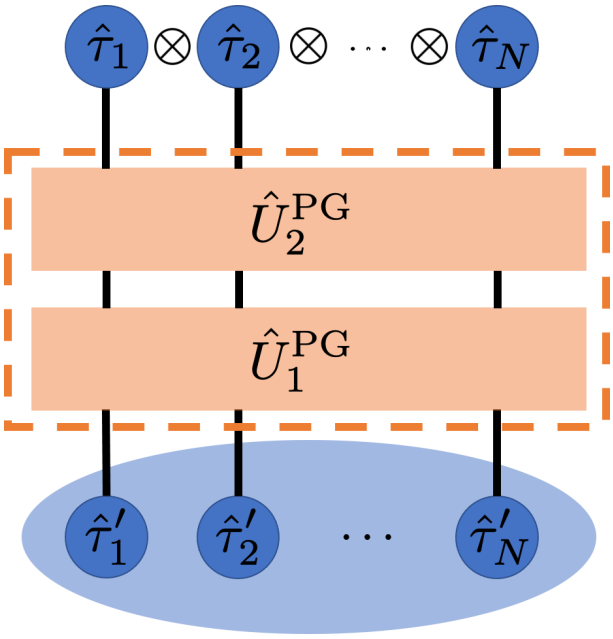

Figure 9.3: Illustration of a linear interferometer acting on a free state, read from top to bottom. We can consider the composition of the two global passive Gaussian unitaries $\hat{U}_1^{\mathrm{PG}}$ and $\hat{U}_2^{\mathrm{PG}}$ as a new global passive unitary $\hat{U}^{\mathrm{PG}} = \hat{U}_1^{\mathrm{PG}}\hat{U}_2^{\mathrm{PG}}$.

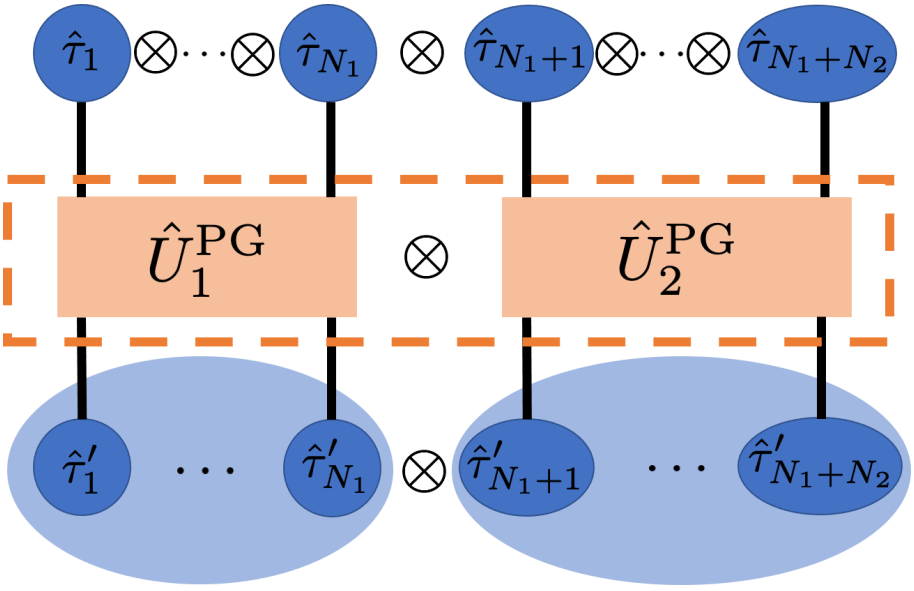

Figure 9.4: Illustration of the tensoring of free states, read from top to bottom. By decomposing each free state as a linear interferometer acting on a product of thermal states, we show that the state obtained by tensoring can itself be considered a product of thermal states on which acts a linear interferometer $\hat{U}_1^{\mathrm{PG}} \otimes \hat{U}_2^{\mathrm{PG}}$.

state. Indeed, the partial tracing of a $(N+1)$-mode free states over its last mode gives the following covariance matrix:

$$\left(\begin{array}{cccc|c} a_{11}\mathbf{I}_2 & a_{12}\mathbf{R}_{12} & \cdots & a_{1N}\mathbf{R}_{1N} & a_{1,N+1}\mathbf{R}_{1,N+1} \\ a_{12}\mathbf{R}_{12}^{\intercal} & a_{22}\mathbf{I}_2 & \cdots & a_{2N}\mathbf{R}_{2N} & a_{2,N+1}\mathbf{R}_{2,N+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{1N}\mathbf{R}_{1N}^{\intercal} & a_{2N}\mathbf{R}_{2N}^{\intercal} & \cdots & a_{NN}\mathbf{I}_2 & a_{N,N+1}\mathbf{R}_{N,N+1} \\ \hline a_{1,N+1}\mathbf{R}_{1,N+1}^{\intercal} & a_{2,N+1}\mathbf{R}_{2,N+1}^{\intercal} & \cdots & a_{N,N+1}\mathbf{R}_{N,N+1}^{\intercal} & a_{N+1,N+1}\mathbf{I}_2 \end{array}\right). \tag{9.12}$$

9.3 Monotones

Now that our set of free state $\mathcal{I}_f$ and set of free operations Λ_f are propertly defined, let us turn our attention to defining monotones. A monotone must be consistent with the set of free states so that it takes the value zero for free states. It should also be consistent with free operations so that the value of a monotone monotonically decreases when any state evolves under the action of a monotone. In what follows, we propose two different monotones, which have different properties as we are going to see. The first one is based on the relative entropy of two quantum states, and we call it the relative entropy of local activity. The second has a more intuitive interpretation since it simply corresponds to the maximum work that can be extracted from Figure 9.1, and we call it the locally extractable work.

9.3.1 Relative entropy of local activity

When defining monotones in a resource theory, it is common to use a contractive distance. Such a distance is a measure between two states which can only decrease when a completely positive map acts on both system. In other words, the distance D is contractive if for any completely positive map ϕ acting on the states $\hat{\rho}$ and $\hat{\sigma}$ we have the following relation:

$$D(\phi(\hat{\rho})\|\phi(\hat{\sigma})) \leq D(\hat{\rho}\|\hat{\sigma}). \tag{9.13}$$

In this context, it is particularly interesting to use the relative entropy. The relative entropy between a state $\hat{\rho}$ and a state $\hat{\sigma}$ is denoted as $S(\hat{\rho}\|\hat{\sigma})$ and is defined as $S(\hat{\rho}\|\hat{\sigma}) = \mathrm{Tr}\left[\hat{\rho}(\ln\hat{\rho} - \ln\hat{\sigma})\right]$. It is a contractive distance, so that it satisfies Equation (9.13). We define the relative entropy of local activity of a state $\hat{\rho}$ as the relative entropy between $\hat{\rho}$ and the closest free state. It is defined as follows:

$$A_l(\hat{\rho}) = \min_{\hat{\sigma}\in\mathcal{I}_f} S(\hat{\rho}\|\hat{\sigma}) \tag{9.14}$$

Let us now show that thanks to the fact that relative entropy is a contractive distance, the relative entropy of local activity can only decrease under a free operation. To that purpose, we consider a state $\hat{\rho}$. Let $\hat{\sigma}_f \in \mathcal{I}_f$ be the closest free state with respect to $\hat{\rho}$, i.e. the free state which minimizes the relative entropy with respect to $\hat{\rho}$ over the set of free states. We can then write:

$$A_l(\hat{\rho}) = \min_{\hat{\sigma}\in\mathcal{I}_f} S\left(\hat{\rho}\|\hat{\sigma}\right) \tag{9.15}$$

$$= S\left(\hat{\rho}\|\hat{\sigma}_f\right) \tag{9.16}$$

$$\geq S\left(\phi_f(\hat{\rho})\|\phi_f(\hat{\sigma}_f)\right) \tag{9.17}$$

$$= S\left(\phi_f(\hat{\rho})\|\hat{\sigma}'_f\right) \tag{9.18}$$

$$\geq \min_{\hat{\sigma}\in\mathcal{I}_f} S\left(\phi_f(\hat{\rho})\|\hat{\sigma}\right) \tag{9.19}$$

$$= A_l(\phi_f(\hat{\rho})) \tag{9.20}$$

We have used the fact that a free operation ϕ_f applies a free state $\hat{\sigma}_f$ onto another free state, so that $\phi_f(\hat{\sigma}_f) = \hat{\sigma}'_f$ is also a free state.

The relative entropy of local activity is thus a consistent monotone with our set of free states. It takes the value zero for any free state, and some positive value for any state which do note belong to the set $\mathcal{I}_f$. As a consequence, the relative entropy of local activity is non-zero for any non-Gaussian state, since our free states are Gaussian. We should note that this property has a slight inconvenient. Indeed, it implies that a mixture of free states yields in general a state with a non-zero relative entropy of local activity. This can be understood from the simple fact that a mixture of Gaussian states is in general non-Gaussian. Of course, this does not contradict our resource theory since we have not considered the mixing of free states as a free operation.

Let us now highlight an interesting property of the relative entropy of local activity which is derived in [61]. When considering a multimode state $\hat{\rho}_{AB}$ and its two reduced subsytems $\hat{\rho}_A$ and $\hat{\rho}_B$ (which are possibly multimode too), the following relation holds:

$$A_l(\hat{\rho}_{AB}) \leq A_l(\hat{\rho}_A) + A_l(\hat{\rho}_B) + S(\hat{\rho}_{AB} \| \hat{\rho}_A \otimes \hat{\rho}_B). \tag{9.21}$$

This relation can be understood as a relaxed form of subbaditivity. Note that this is a rather surprising property, since we would have expected the inequality to be in the other direction. Indeed, in general, we expect that a global state contains at least as much resource as the sum of the resource contained in each of its subsystems. However, that relation is not a strict relation of subbaditivty because of the extra term, and there is thus no contradiction. We will see that the other monotone that we will introduce in the following of this section has properties which are more intuitive than the relative entropy of local activity.

Finally, let us mention that in the particular case of single-mode states, the relative entropy of local activity can be computed explicitly. Indeed, in that context, the free states are thermal states and we find the following relation:

$$A_l(\hat{\rho}) = g(\bar{n}) - S(\hat{\rho}), \tag{9.22}$$

where $\bar{n}$ is the mean number of photon of $\hat{\rho}$, and $g(\bar{n})$ is the von Neumann entropy of a thermal state with mean number of photon $\bar{n}$ (as defined in 4.89). Since thermal states maximize the von Neumann entropy at a given energy, it is clear from relation (9.22) that $A_l(\hat{\rho})$ is non-negative and is zero only in the case of thermal states, as expected.

9.3.2 Locally extractable work

Let us now define a monotone from a different approach. As we know, energy is a quantity that is conserved. Therefore, it is natural that the work extracted from our system cannot grow through the free operations that we have defined. In order to determine what is the maximum energy that can be extracted from a state $\hat{\rho}$, we need to determine what is the lowest energy state that can be constructed from $\hat{\rho}$ according to our setup pictured in Figure 9.1.

Now, remember that the energy of a system is only a function of its covariance matrix $\mathbf{V}$ and its displacement vector $\bar{\mathbf{q}}$, as it appears from Equation (4.82). The first thing to do in order to reduce the energy of a state is therefore to bring its mean displacement vector $\bar{\mathbf{q}}$ onto zero. This can be done with the help of the displacement operator $\hat{D}_\alpha$, which is local and Gaussian and thus allowed by our setup. Also, since all the operations that we may perform are symplectic, the spectrum of symplectic eigenvalues of the covariance matrix is conserved. From what we have explained in Section 9.1, we know that for a given set of symplectic eigenvalues associated to a covariance matrix $\mathbf{V}$, it is the Williamson form $\mathbf{V}_\mathrm{s}$ that will minimize the trace of the covariance matrix.

We show with the help of Figure 9.5 that it is possible thanks to our setup to bring any state to this state of least energy that we have presented. This allows us to define the following monotone, which we call the locally extractable work and which we denote by W_l:

$$W_l(\hat{\rho}) = \frac{1}{2}\bar{\mathbf{q}}^2 + \frac{1}{2}\,\mathrm{Tr}\,[\mathbf{V}] - \frac{1}{2}\mathrm{Str}\,[\mathbf{V}]\,. \tag{9.23}$$

The monotone W_l has an intuitive interpretation as it corresponds to the maximum extractable work with the use of our setup. It is interesting to note that it only depends on the mean displacement

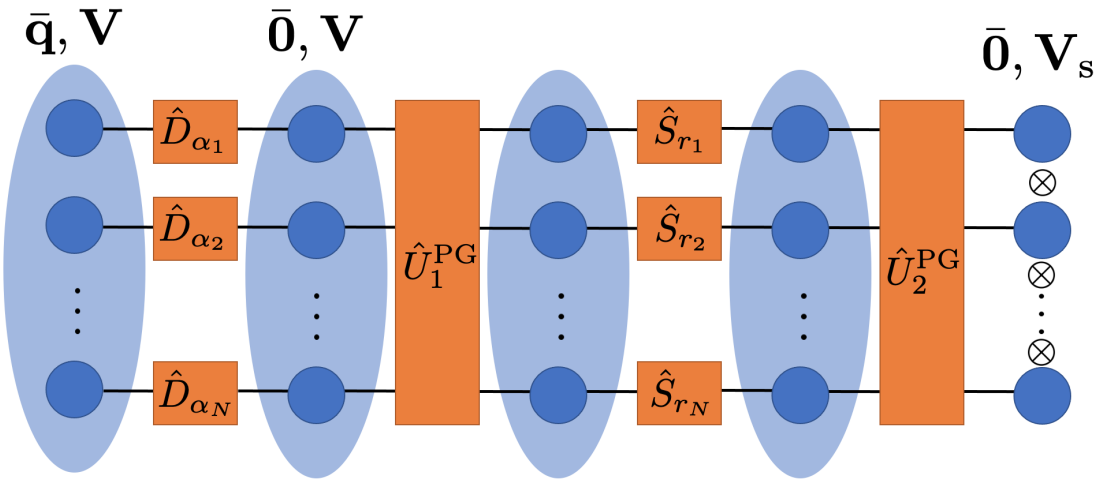

Figure 9.5: This scheme presents the optimal way to extract energy using the setup described in Figure 9.1. We start from a quantum state $\hat{\rho}$ with mean displacement $\bar{\mathbf{q}}$ and covariance matrix $\mathbf{V}$. We then apply a displacement operator $\hat{D}_{\alpha_i}$ on every mode in order to bring the mean displacement onto zero. Then, taking advantage of the Bloch-Messiah decomposition, we bring the covariance matrix $\mathbf{V}$ onto its Williamson form $\mathbf{V}_{\mathrm{s}}$ with the help of a global Gaussian passive unitary $\hat{U}_1^{\mathrm{PG}}$, followed by a set of single-mode squeezer $\hat{S}_{r_i}$ and followed by another global Gaussian unitary $\hat{U}_2^{\mathrm{PG}}$. The state we obtain at the end has a mean displacement $\mathbf{0}$ and a covariance matrix $\mathbf{V}_{\mathrm{s}}$. The difference of energy between the input state and the output state corresponds to the locally extractable work $W_l(\hat{\rho})$.

$\bar{\mathbf{q}}$ and covariance matrix $\mathbf{V}$ of the state. As a consequence, it takes the value zero for our set of free state $\mathcal{I}_f$, but also for any state which has the same mean displacement and covariance matrix as a free state. This leads us to define an extended set $\mathcal{I}_w$ which is the set of states, possibly non-Gaussian, that have the covariance matrix of a free state and a zero mean displacement:

$$\mathcal{I}_w = \{\hat{\rho} : \bar{\mathbf{q}}_{\hat{\rho}} = \mathbf{0}, \mathbf{V}_{\hat{\rho}} \in \mathcal{V}_f\}\,. \tag{9.24}$$

From its definition, it is straightforward to see that the set $\mathcal{I}_w$ contains the set $\mathcal{I}_f$. Moreover, since the set of free covariance matrices $\mathcal{V}_f$ is convex, it directly follows that $\mathcal{I}_w$ is also convex. The set $\mathcal{I}_w$ provides then a natural convex extension of our set of free states. We plot in Figure 9.6 a representation of $\mathcal{I}_f$ and $\mathcal{I}_w$.

Finally, let us mention two properties of the locally extractable work W_l which makes it a montone more convenient than the relative entropy of local activity A_l. First, it can be shown that W_l is convex, and we have the following relation:

$$W_l\left(\sum p_i\hat{\rho}_i\right) \leq \sum p_i W_l(\hat{\rho}_i). \tag{9.25}$$

This can be understood as the fact that extracting energy from a mixture result in less energy than if we had extracted the energy from each state separately. A second interesting property of the locally extractable work is that it is superadditive. If we consider a multimode state $\hat{\rho}_{AB}$ with its two possibly multimode subsystems $\hat{\rho}_A$ and $\hat{\rho}_B$, the following relation holds:

$$W_l(\hat{\rho}_{AB}) \geq W_l(\hat{\rho}_A) + W_l(\hat{\rho}_B) \tag{9.26}$$

Notice the strinking difference between property (9.26) and property (9.21). The property of superadditivity is consistent with the example of the two-mode squeezed vacuum that we have provided earlier (see (9.9)). Indeed, since a two-mode squeezed vacuum is locally thermal, each subsytem has a zero locally extractable work. However, if we have access to the whole system, it is possible to extract a certain amount of energy.

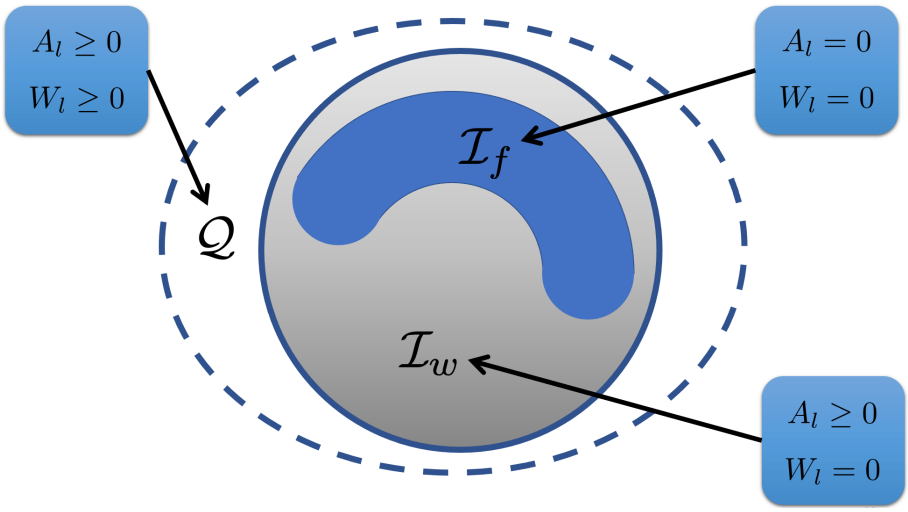

Figure 9.6: Representation of the sets $\mathcal{I}_f$ and $\mathcal{I}_w$. The set $\mathcal{Q}$ is the set of all possible quantum states, which have in general a non-negative relative entropy of local activity and a non-negative locally extractable work. The set $\mathcal{I}_w$ is a convex set, which contains states such that the locally extractable work is zero. States in $\mathcal{I}_w$ are not necessarily Gaussian and can have a non-negative relative entropy of local activity. Finally, the set $\mathcal{I}_f$ is our set of free states and is included in $\mathcal{I}_w$. Every state in $\mathcal{I}_f$ is Gaussian and has a zero relative entropy of local activity and a zero locally extractable work. Note that any state in $\mathcal{I}_w$ has the same covariance matrix as some state in $\mathcal{I}_f$.

Conclusion

The initial ambition of this work is to confront the mathematical theory of majorization to quantum distributions as embodied by Wigner functions. Addressing the notion of disorder in phase space is not a novelty in itself. Such an approach has for instance already been carried out based on the Husimi Q-distribution and has led to the Wehrl conjecture and its generalized form, which is in fact nothing else but a majorization relation. The reason why the Wigner function has not yet been subject to a characterization of its disorder is certainly that it is not a genuine probability distribution in itself. At first sight, it seems impossible to apply the usual tools of probability theory to a distribution that may be partly negative. And indeed, several obstacles needed to be overcome in order to accomplish this goal.

First, majorization theory was not appropriately defined for continuous distributions taking both positive and negative values on a domain of infinite size. We have circumvented this problem by proposing in Chapter 3 a formulation of majorization that can be extended to arbitrary distributions defined over an infinite domain. With this improved tool, we were ready to tackle the characterization of disorder in phase space.

Second, the set of Wigner functions that constitute *de facto* a true probability distribution is not well understood. The corresponding states, which are Wigner-positive, are of particular interest in our work because they are precisely the states for which all the usual measures of disorder are well defined. The Wigner-positive set lacks a comprehensive and clean characterization as this task seems difficult to accomplish [10, 46]. In Chapter 5, we have presented an overview of the Wigner-positive set and highlighted a general technique to build a large set of Wigner-positive states using a balanced beam-splitter. The beam-splitter states are expected to play a crucial role in the structure of the Wigner-positive set as they coincide with some if its extremal states. However, things are more complicated and we have also shown the existence of extremal Wigner-positive that cannot be constructed within the beam-splitter setup.

Having addressed these two obstacles and partially resolved them, we have considered the central question of finding the state of least disorder in phase space in Chapter 6. A first calculation led us to the observation that pure states are in general incomparable (or then simply equivalent) in terms of majorization. This result seems consistent with our intuition of a pure state, which minimizes the statistical uncertainty. It therefore seems natural that a pure state cannot be objectively deemed more disordered than another. We then formulated a majorization conjecture restricted to the set of Wigner-positive states. This conjecture, which can be viewed as an extension of Hudson's theorem, states that Gaussian pure states are the states with least disorder among the set of Wigner-positive states. We then partially proved this conjecture in a special case, restricting to the subset of phase-invariant Wigner-positive states containing up to 2 photons.

Then, in Chapter 7, we have first been investigating disorder measures that are consistent with the set of positive and negative Wigner functions. It appeared to us that Rényi entropies of parameter $\alpha > 1$ are possible candidates since for $\alpha \leq 1$ they are not Schur-concave anymore over the set of Wigner-negative states. Then, we have focused on the set of Wigner-positive states and have defined the Shannon differential entropy of a Wigner-positive state as its Wigner entropy, and have highlighted several of its properties. We have in particular conjectured a lower bound, which is somehow analogous to (but stronger that) the Wehrl conjecture, and have proven it for the set of passive states.

Finally, in Chapter 8, we have started to expand our findings to quantum Gaussian phase-insensitive bosonic channels. These channels, which are ubiquitous in quantum information, are completely characterized by a gain/loss parameter κ and a noise parameter μ. Such channels can in

general be constructed with a beam-splitter or two-mode squeezer coupled with a thermal state. We have shown the existence of a continuous majorization relation between the input and output Wigner functions in channels characterized by a gain parameter κ such $|\kappa| \geq 1$. This property actually extends to any two-mode squeezer coupled with a Wigner-positive state. We have also shown that the output of a Gaussian phase-insensitive channel is always Wigner-positive under the simple condition that $\mu \geq |\kappa|$, which echoes the condition for entanglement-breaking that is $\mu \geq |\kappa| + 1$. We then studied a discrete majorization chain for Fock states in these bosonic channels, and showed that the vector of eigenvalues associated to the output of the nth Fock state majorizes the vector of eigenvalues associated to the output of the $(n+1)$th Fock state. This is a non-trivial generalization of a result for quantum-limited bosonic channels.

Let us also mention some additional work that we carried in the present thesis. In Chapter 9, we have taken advantage of the symplectic formalism to develop a resource theory applicable to bosonic systems. In this context, we have defined the free states as the states from which no work could be extracted locally, assisted by a Gaussian global unitary. We have then defined the free operations consistently with the free states. We have highlighted a monotone which was expressed in simple terms as the difference between the trace and the symplectic trace, and which can be associated to the locally extractable work.

This summarizes the main advances we have achieved throughout this thesis. However, several of the research directions discussed here remain partly open. First, let us mention the role of beam-splitter states within the set of Wigner-positive states. As we have seen in Chapter 5, the convex hull of beam-splitter states seems to constitute a major portion of the Wigner-positive set. Hence, a further investigation of the Wigner-positive states that cannot be expressed as a convex mixture of beam-splitter states could give precious insights into the full set of Wigner-positive states. We already know that these states correspond precisely to the states requiring an entangled state if produced with the beam-splitter setup, so the link between Wigner-negativity and entanglement deserves to be studied in detail.

Regarding continuous majorization, we have focused on our central conjecture over the Wigner-positive set. This conjecture expresses that Gaussian pure states majorize any Wigner-positive state, and is non-trivial from the fact that Wigner-positive states go beyond the convex hull of Gaussian states. We can in fact associate to any pure state $|\psi\rangle$ a set of states whose Wigner function are majorized by the one of $|\psi\rangle$. That set is non-trivial as it is presumably larger than the convex mixtures of symplectic transformations applied on $|\psi\rangle$. Another promising direction to be able to express that any give pure state is more disordered that Gaussian pure states relies on square majorization.

With respect to Wigner entropy, an obvious future direction of research is the extension of this quantity to Wigner-negative states. Indeed, we know that the current definition of entropy does not converge to a usable value for such states. Let us simply list the properties that we expect from such a *pseudo* Wigner entropy, which we label as $\tilde{h}(W)$ in the following. First, $\tilde{h}(W)$ should of course reduce to the usual Wigner-entropy for Wigner-positive states. Second, it should be a Schur-concave quantity such that $W_1 \succ W_2$ implies $\tilde{h}(W_1) \leq \tilde{h}(W_2)$, even when W_1 and W_2 are not Wigner-positive. Finally, an important property of $\tilde{h}(W)$ relates to the subadditivity of entropy: it should be the case that $\tilde{h}(W)$ constitute a lower-bound on the sum of the marginal entropies, so that $\tilde{h}(W) \leq h(\rho_x) + h(\rho_p)$.

Finally, we would like to present a last open problem. So far, our most general technique to build Wigner-positive states is the use of a balanced beam-splitter. Starting from any quantum state, we have with this technique a way to construct a Wigner-positive quantum state. To do this, we need to mix the initial state $|\psi_1\rangle$ together with another quantum state $|\psi_2\rangle$ at the beam splitter. A natural question is thus to ask what is the state $|\psi_2\rangle$ which minimizes the Wigner entropy of the output state for a given state $|\psi_1\rangle$. The situation is illustrated on Figure 9.7. When $|\psi_1\rangle$ is a pure Gaussian states, $|\psi_2\rangle$ can easily be chosen to the same Gaussian pure states, so that the output is also a pure Gaussian states and its Wigner entropy is $\ln \pi + 1$, which is our conjectured lowest possible Wigner entropy. However, when considering non-Gaussian states as input, the situation becomes more complex.

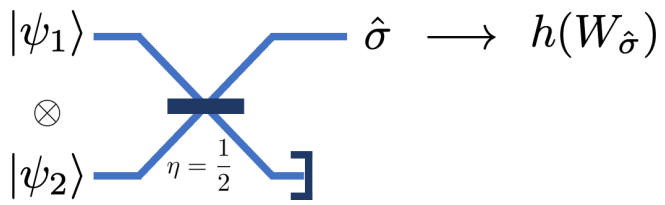

Figure 9.7: Given a pure state $|\psi_1\rangle$, what is the other state $|\psi_2\rangle$ that minimizes the Wigner entropy of the output $h(W_{\hat{\sigma}})$? The question can somehow be understood as a generalization of the entropy power inequality, since the Wigner functions of $|\psi_1\rangle$ and $|\psi_2\rangle$ are in general partly negative. The Wigner function of $\hat{\sigma}$, however, is always non-negative and possesses a well-defined Shannon differential entropy. If $|\psi_1\rangle$ is Gaussian, the choice corresponds simply to the same Gaussian state $|\psi_2\rangle = |\psi_1\rangle$, so that the Wigner entropy of the output is $\ln \pi + 1$ and reaches our conjectured lower-bound. When $|\psi_1\rangle$ is not Gaussian, the question becomes more difficult to answer. A reasonable guess is to choose $|\psi_2\rangle$ to be the Gaussian pure state with covariance matrix proportional to the one of $|\psi_1\rangle$.

Appendix A

Proofs of majorization lemmas

A.1 Proof of lemma 1

Remember that lemma 1 reads as follows:
If f and g are two n-dimensional radial distributions defined on $\mathbb{R}^n$ such that $f(\mathbf{r}) = f_{\mathrm{R}}(\|\mathbf{r}\|)$ and $g(\mathbf{r}) = g_{\mathrm{R}}(\|\mathbf{r}\|)$ with f_{R} and g_{R} defined on $\mathbb{R}_+$, then $f \succ g$ is equivalent to $\tilde{f} \succ \tilde{g}$, where $\tilde{f}$ and $\tilde{g}$ are 1-dimensional distributions defined on $\mathbb{R}_+$ as $\tilde{f}(x) = f_{\mathrm{R}}(\sqrt[n]{x})$ and $\tilde{g}(x) = g_{\mathrm{R}}(\sqrt[n]{x})$.

Using polar coordinates, we have that $f \succ g$ is equivalent to the following conditions:

$$\begin{cases} \int_0^\infty f_{\mathrm{R}}(r) r^{n-1} \mathrm{d}r = \int_0^\infty g_{\mathrm{R}}(r) r^{n-1} \mathrm{d}r \\ \int_0^\infty \left[f_{\mathrm{R}}(r) - t\right]^+ r^{n-1}\, \mathrm{d}r \geq \int_0^\infty \left[g_{\mathrm{R}}(r) - t\right]^+ r^{n-1}\, \mathrm{d}r \qquad \forall t \geq 0 \\ \int_0^\infty \left[f_{\mathrm{R}}(r) - t\right]^- r^{n-1}\, \mathrm{d}r \leq \int_0^\infty \left[g_{\mathrm{R}}(r) - t\right]^- r^{n-1}\, \mathrm{d}r \qquad \forall t \leq 0 \end{cases} \tag{A.1}$$

Introducing the simple change of variables $x = r^n$, we rewrite the above conditions as

$$\begin{cases} \int_0^\infty f_{\mathrm{R}}(\sqrt[n]{x}) \mathrm{d}x = \int_0^\infty g_{\mathrm{R}}(\sqrt[n]{x}) \mathrm{d}x \\ \int_0^\infty \left[f_{\mathrm{R}}\left(\sqrt[n]{x}\right) - t\right]^+ \mathrm{d}x \geq \int_0^\infty \left[g_{\mathrm{R}}\left(\sqrt[n]{x}\right) - t\right]^+ \mathrm{d}x \qquad \forall t \geq 0 \\ \int_0^\infty \left[f_{\mathrm{R}}\left(\sqrt[n]{x}\right) - t\right]^- \mathrm{d}x \geq \int_0^\infty \left[g_{\mathrm{R}}\left(\sqrt[n]{x}\right) - t\right]^- \mathrm{d}x \qquad \forall t \leq 0 \end{cases} \tag{A.2}$$

Defining $\tilde{f}(x) = f_{\mathrm{R}}(\sqrt[n]{x})$ and $\tilde{g}(x) = g_{\mathrm{R}}(\sqrt[n]{x})$, it follows that $\tilde{f} \succ \tilde{g}$ is equivalent to $f \succ g$.

A.2 Proof of lemma 2

Remember that lemma 2 reads as follows:
Consider two probability distributions f and g defined on the same domain $\mathbb{A}$. If there exists a collection of level-equivalent distributions $f^{(\alpha)}$ on $\mathbb{A}$ depending on the parameter α with $f^{(\alpha)} \equiv f$ for all α such that

$$g(\mathbf{r}) = \int k(\alpha)\, f^{(\alpha)}(\mathbf{r})\, \mathrm{d}\alpha, \quad \forall \mathbf{r} \in \mathbb{A}, \tag{A.3}$$

where k is a probability density distribution, then $f \succ g$.

First, notice that since $f^{(\alpha)} \equiv f$, it follows that they have the same normalization. As a consequence, we can show that g and f have the same normalization:

$$\int g(\mathbf{r})\mathrm{d}\mathbf{r} = \int \mathrm{d}\mathbf{r} \int \mathrm{d}\alpha k(\alpha) f^{(\alpha)}(\mathbf{r}) \tag{A.4}$$

$$= \int \mathrm{d}\alpha k(\alpha) \int f^{(\alpha)}(\mathbf{r})\mathrm{d}\mathbf{r} \tag{A.5}$$

$$= \int f(\mathbf{r})\mathrm{d}\mathbf{r} \tag{A.6}$$

Since the function $\gamma_t(z) = [z-t]^+$ is convex, we can exploit Jensen's inequality to get

$$\begin{aligned} \gamma_t\left(g\left(\mathbf{r}\right)\right) &= \gamma_t\left(\int_0^\infty k(\alpha)\, f^{(\alpha)}\left(\mathbf{r}\right) \mathrm{d}\alpha\right) \\ &\leq \int_0^\infty k(\alpha)\, \gamma_t\left(f^{(\alpha)}\left(\mathbf{r}\right)\right) \mathrm{d}\alpha. \end{aligned} \tag{A.7}$$

Integrating both terms over the domain $\mathbb{A}$ then leads to

$$\int_{\mathbb{A}} \gamma_t\left(g\left(\mathbf{r}\right)\right) \mathrm{d}\mathbf{r} \leq \int_0^\infty k(\alpha) \int_{\mathbb{A}} \gamma_t\left(f^{(\alpha)}\left(\mathbf{r}\right)\right) \mathrm{d}\mathbf{r}\, \mathrm{d}\alpha. \tag{A.8}$$

Since f and $f^{(\alpha)}$ are level-equivalent, we have that

$$\int \gamma_t\left(f^{(\alpha)}\left(\mathbf{r}\right)\right) \mathrm{d}\mathbf{r} = \int \gamma_t\left(f\left(\mathbf{r}\right)\right) \mathrm{d}\mathbf{r}. \tag{A.9}$$

The integral over $k(\alpha)$ reduces to 1 since it is a probability distribution. Writing γ_t explicitly, we end up with

$$\int [f(\mathbf{r}) - t]_+ \, \mathrm{d}\mathbf{r} \geq \int [g(\mathbf{r}) - t]_+ \, \mathrm{d}\mathbf{r} \tag{A.10}$$

which can equivalently be written as $\Phi_t^+(f) \geq \Phi_t^+(g)$.

As similar development using the concavity of γ_t^- yields that $\Phi_t^-(f) \leq \Phi_t^-(g)$. We can thus write

$$\begin{cases} \int f(\mathbf{r})\mathrm{d}\mathbf{r} = \int g(\mathbf{r})\mathrm{d}\mathbf{r} \\ \Phi_t^+(f) \geq \Phi_t^+(g) \qquad \forall t \geq 0 \\ \Phi_t^-(f) \leq \Phi_t^-(g) \qquad \forall t \leq 0 \end{cases} \tag{A.11}$$

which is equivalent to $f \succ g$.

A.3 Proof of lemma 3

Remember that lemma 3 reads as follows:
Consider four functions f_1, f_2, g_1, and g_2 defined on the same domain $\mathbb{A}$ and such that f_1 and f_2 do not both take non-zero values in the same element of $\mathbb{A}$, and similarly g_1 and g_2 do not both take non-zero values in the same element of $\mathbb{A}$. If the functions satisfy $f_1 \succ g_1$ and $f_2 \succ g_2$, then $(f_1 + f_2) \succ (g_1 + g_2)$.

Notice first that since $f_1 \succ g_1$, f_1 and g_1 are normalized to the same value. Also, since $f_2 \succ g_2$, f_2 and g_2 are normalized to the same value. As a consequence, we can write:

$$\begin{aligned}
\int (f_1 + f_2)(\mathbf{r})\mathrm{d}\mathbf{r} &= \int \left(f_1(\mathbf{r}) + f_2(\mathbf{r})\right) \mathrm{d}\mathbf{r} && \text{(A.12)}\\
&= \int f_1(\mathbf{r})\mathrm{d}\mathbf{r} + \int f_2(\mathbf{r})\mathrm{d}\mathbf{r} && \text{(A.13)}\\
&= \int g_1(\mathbf{r})\mathrm{d}\mathbf{r} + \int g_2(\mathbf{r})\mathrm{d}\mathbf{r} && \text{(A.14)}\\
&= \int \left(g_1(\mathbf{r} + g_2(\mathbf{r}))\right) \mathrm{d}\mathbf{r} && \text{(A.15)}\\
&= \int (g_1 + g_2)(\mathbf{r})\mathrm{d}\mathbf{r} && \text{(A.16)}
\end{aligned}$$

From the fact that for each $\mathbf{r}$, either $f_1(\mathbf{r})$ or $f_2(\mathbf{r})$ is equal to zero, it follows that :

$$\forall t \geq 0 : \qquad [f_1(\mathbf{r}) - t]^+ + [f_2(\mathbf{r}) - t]^+ = [f_1(\mathbf{r}) + f_2(\mathbf{r}) - t]^+ , \tag{A.17}$$

$$\forall t \leq 0 : \qquad [f_1(\mathbf{r}) - t]^- + [f_2(\mathbf{r}) - t]^- = [f_1(\mathbf{r}) + f_2(\mathbf{r}) - t]^- . \tag{A.18}$$

The same applies to g_1 and g_2. Let $f = f_1 + f_2$ and $g = g_1 + g_2$. For all $t \geq 0$, we have:

$$\begin{aligned}
\int [f(\mathbf{r}) - t]^+ \mathrm{d}\mathbf{r} &= \int [f_1(\mathbf{r}) + f_2(\mathbf{r}) - t]^+ \mathrm{d}\mathbf{r}\\
&= \int [f_1(\mathbf{r}) - t]^+ \mathrm{d}\mathbf{r} + \int [f_2(\mathbf{r}) - t]^+ \mathrm{d}\mathbf{r}\\
&\geq \int [g_1(\mathbf{r}) - t]^+ \mathrm{d}\mathbf{r} + \int [g_2(\mathbf{r}) - t]^+ \mathrm{d}\mathbf{r} && \text{(A.19)}\\
&= \int [g_1(\mathbf{r}) + g_2(\mathbf{r}) - t]^+ \mathrm{d}\mathbf{r}\\
&= \int [g(\mathbf{r}) - t]^+ \mathrm{d}\mathbf{r},
\end{aligned}$$

where the inequality follows from $f_1 \succ g_1$ and $f_2 \succ g_2$. Similarly, we can write for all $t \leq 0$:

$$\begin{aligned}
\int [f(\mathbf{r}) - t]^- \mathrm{d}\mathbf{r} &= \int [f_1(\mathbf{r}) + f_2(\mathbf{r}) - t]^- \mathrm{d}\mathbf{r}\\
&= \int [f_1(\mathbf{r}) - t]^- \mathrm{d}\mathbf{r} + \int [f_2(\mathbf{r}) - t]^- \mathrm{d}\mathbf{r}\\
&\leq \int [g_1(\mathbf{r}) - t]^- \mathrm{d}\mathbf{r} + \int [g_2(\mathbf{r}) - t]^- \mathrm{d}\mathbf{r} && \text{(A.20)}\\
&= \int [g_1(\mathbf{r}) + g_2(\mathbf{r}) - t]^- \mathrm{d}\mathbf{r}\\
&= \int [g(\mathbf{r}) - t]^- \mathrm{d}\mathbf{r},
\end{aligned}$$

where the inequality follows from $f_1 \succ g_1$ and $f_2 \succ g_2$. In conclusion, we thus have shown that:

$$\begin{cases}
\displaystyle\int (f_1 + f_2)(\mathbf{r})\mathrm{d}\mathbf{r} = \int (g_1 + g_2)(\mathbf{r})\mathrm{d}\mathbf{r} & \\
\Phi_t^+(f_1 + f_2) \geq \Phi_t^+(g_1 + g_2) \qquad \forall t \geq 0 & \\
\Phi_t^-(f_1 + f_2) \leq \Phi_t^-(g_1 + g_2) \qquad \forall t \leq 0 &
\end{cases} \tag{A.21}$$

which is equivalent to $(f_1 + f_2) \succ (g_1 + g_2)$.

Appendix B

Proof of the majorization relation for the states located on the ellipse

In this appendix, we prove that the Wigner functions of the extremal Wigner-positive states located on the ellipse represented in Fig. 6.3 are majorized by the Wigner function of the vacuum state, by showing that $f_0 \succ g_t$ for all t such that $0 \leq t \leq 1$. Note that the proof is very similar to the proof of $f_0 \succ f_c$. The function $g_t(x)$ defined in Eq. (6.48) has one zero at $x = a_t$, where

$$a_t = 1 - \sqrt{\frac{1-t}{1+t}}. \tag{B.1}$$

We "split" g_t in two different functions g_t^- and g_t^+ from either sides of a_t:

$$g_t^-(x) = \begin{cases} g_t(x), & \text{for } 0 \leq x \leq a_t, \\ 0, & \text{else}, \end{cases} \tag{B.2}$$

$$g_t^+(x) = \begin{cases} 0, & \text{for } 0 \leq x \leq a_t, \\ g_t(x), & \text{else}. \end{cases} \tag{B.3}$$

If we shift g_t^+ from a_t towards the origin, we have a distribution that is proportional to $\exp(-x)x^2/2$, which we know is majorized by $f_0(x) = \exp(-x)$ from our previous result. The idea now is to split f_0 in two different functions with the same normalization as g_t^- and g_t^+. Define b_t as

$$b_t = 1 - \ln(1+t) - \sqrt{\frac{1-t}{1+t}}. \tag{B.4}$$

It satisfies

$$\begin{aligned} \int_0^{a_t} g_t(x)\mathrm{d}x &= \int_0^{b_t} f_0(x)\mathrm{d}x, \\ \int_{a_t}^{\infty} g_t(x)\mathrm{d}x &= \int_{b_t}^{\infty} f_0(x)\mathrm{d}x. \end{aligned} \tag{B.5}$$

We now "split" f_0 from either side of $x = b_t$, and define f_0^- and f_0^+ as:

$$f_0^-(x) = \begin{cases} f_0(x), & \text{for } 0 \leq x \leq b_t, \\ 0, & \text{else}, \end{cases} \tag{B.6}$$

$$f_0^+(x) = \begin{cases} 0, & \text{for } 0 \leq x \leq b_t, \\ f_0(x), & \text{else}. \end{cases} \tag{B.7}$$

We are now going to prove that $f_0^- \succ g_t^-$ and $f_0^+ \succ g_t^+$. Let us shift g_t^+ and f_0^+ towards the origin respectively from a_t and b_t. We end up with the functions:

$$\begin{aligned}\tilde{g}_t^+(x) &= g_t^+(x + a_t) \\ &= (t+1)\exp\left(\sqrt{\frac{1-t}{1+t}} - 1\right)\exp(-x)\frac{1}{2}x^2,\end{aligned} \tag{B.8}$$

and

$$\begin{aligned}\tilde{f}_0^+(x) &= f_0^+(x + b_t) \\ &= (t+1)\exp\left(\sqrt{\frac{1-t}{1+t}} - 1\right)\exp(-x).\end{aligned} \tag{B.9}$$

$\tilde{f}_0^+(x)$ and $\tilde{g}_t^+(x)$ are respectively proportional to $\exp(-x)$ and $\exp(-x)x^2/2$, with the same proportionality factor. Since we have already shown that $\exp(-x) \succ \exp(-x)x^2/2$, it follows that $\tilde{f}_0^+ \succ \tilde{g}_t^+$. Since $\tilde{f}_0^+$ and $\tilde{g}_t^+$ are level-equivalent to respectively f_0^+ and g_t^+, we then have $f_0^+ \succ g_t^+$.

We now turn to $f_0^- \succ g_t^-$. They are both monotonically decreasing, making them decreasing rearrangements. Therefore their cumulative integrals are given by:

$$S_s(f_0^-) = \int_0^s f_0^-(x)\mathrm{d}x \quad \text{and} \quad S_s(g_t^-) = \int_0^s g_t^-(x)\mathrm{d}x. \tag{B.10}$$

We will now show that $S_s(f_0^-) \geq S_s(g_t^-)$ for all $s \in \mathbb{R}_+$. Since f_0^- and g_t^- are both monotonically decreasing and $f_0^-(x) = 0$ for all $x > b_t$ coupled with the fact that $b_t < a_t$, it is sufficient to show that $f_0^-(x) \geq g_t^-(x)$ for all $x \in [0, b_t]$. To prove this, note that the ratio $g_t(x)/f_0(x)$ is less than 1 for $x \in [0, b_t]$:

$$\frac{g_t(x)}{f_0(x)} = \frac{t+1}{2}\left(x - 1 + \sqrt{\frac{1-t}{1+t}}\right)^2 \leq 1 \quad \text{for } 0 \leq x \leq b_t, \tag{B.11}$$

which follows from the fact that $x \leq b_t \leq 1$ and $0 \leq t \leq 1$. Since f_0^- and g_t^- are non-negative and normalized to the same value, we conclude that $f_0^- \succ g_t^-$. From Lemma 3, we finally end up with $f_0 \succ g_t$.

Bibliography

[1] Gerardo Adesso, Sammy Ragy, and Antony R Lee. Continuous variable quantum information: Gaussian states and beyond. *Open Systems & Information Dynamics*, 21(01n02):1440001, 2014.

[2] Robert Alicki. The quantum open system as a model of the heat engine. *Journal of Physics A: Mathematical and General*, 12(5):L103, 1979.

[3] Konstantin Ivanovich Babenko. An inequality in the theory of Fourier integrals. *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, 25(4):531–542, 1961.

[4] Simon Becker, Nilanjana Datta, Ludovico Lami, and Cambyse Rouzé. Convergence rates for the quantum central limit theorem. *Communications in Mathematical Physics*, 383(1):223–279, 2021.

[5] William Beckner. Inequalities in Fourier analysis. *Annals of Mathematics*, 102(1):159–182, 1975.

[6] Iwo Białynicki-Birula. Formulation of the uncertainty relations in terms of the Rényi entropies. *Physical Review A*, 74(5):052101, 2006.

[7] Iwo Białynicki-Birula and Jerzy Mycielski. Uncertainty relations for information entropy in wave mechanics. *Communications in Mathematical Physics*, 44(2):129–132, 1975.

[8] Fernando Brandao, Michał Horodecki, Nelly Ng, Jonathan Oppenheim, and Stephanie Wehner. The second laws of quantum thermodynamics. *Proceedings of the National Academy of Sciences*, 112(11):3275–3279, 2015.

[9] Fernando GSL Brandao, Michał Horodecki, Jonathan Oppenheim, Joseph M Renes, and Robert W Spekkens. Resource theory of quantum states out of thermal equilibrium. *Physical Review Letters*, 111(25):250404, 2013.

[10] T Bröcker and RF Werner. Mixed states with positive Wigner functions. *Journal of mathematical physics*, 36(1):62–75, 1995.

[11] Eric G Brown, Nicolai Friis, and Marcus Huber. Passivity and practical work extraction using Gaussian operations. *New Journal of Physics*, 18(11):113028, 2016.

[12] Herbert B Callen. Thermodynamics and an introduction to thermostatistics, 1998.

[13] Nicolas J Cerf and Michael G Jabbour. Two-boson quantum interference in time. *Proceedings of the National Academy of Sciences*, 117(52):33107–33116, 2020.

[14] NJ Cerf and Sofyan Iblisdir. Phase conjugation of continuous quantum variables. *Physical Review A*, 64(3):032307, 2001.

[15] Patrick J Coles, Mario Berta, Marco Tomamichel, and Stephanie Wehner. Entropic uncertainty relations and their applications. *Reviews of Modern Physics*, 89(1):015002, 2017.

[16] Thomas M Cover and Joy A Thomas. Information theory and statistics. *Elements of Information Theory*, 1(1):279–335, 1991.

[17] Stephan De Bièvre, Dmitri B Horoshko, Giuseppe Patera, and Mikhail I Kolobov. Measuring nonclassicality of bosonic field quantum states via operator ordering sensitivity. *Physical Review Letters*, 122(8):080402, 2019.

[18] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.

[19] Luc P Faucheux, LS Bourdieu, PD Kaplan, and Albert J Libchaber. Optical thermal ratchet. *Physical Review Letters*, 74(9):1504, 1995.

[20] CN Gagatsos, Ognyan Oreshkov, and NJ Cerf. Majorization relations and entanglement generation in a beam splitter. *Physical Review A*, 87(4):042307, 2013.

[21] Raul Garcia-Patron, Carlos Navarrete-Benlloch, Seth Lloyd, Jeffrey H Shapiro, and Nicolas J Cerf. Majorization theory approach to the Gaussian channel minimum entropy conjecture. *Physical Review Letters*, 108(11):110505, 2012.

[22] JE Geusic, EO Schulz-DuBios, and HED Scovil. Quantum equivalent of the carnot cycle. *Physical Review*, 156(2):343, 1967.

[23] Vittorio Giovannetti, Raul Garcia-Patron, Nicolas J Cerf, and Alexander S Holevo. Ultimate classical communication rates of quantum optical channels. *Nature Photonics*, 8(10):796–800, 2014.

[24] John Goold, Marcus Huber, Arnau Riera, Lídia Del Rio, and Paul Skrzypczyk. The role of quantum information in thermodynamics—a topical review. *Journal of Physics A: Mathematical and Theoretical*, 49(14):143001, 2016.

[25] Peter Hänggi and Fabio Marchesoni. Artificial brownian motors: Controlling transport on the nanoscale. *Reviews of Modern Physics*, 81(1):387, 2009.

[26] Anaelle Hertz. *Exploring continuous-variable entropic uncertainty relations and separability criteria in quantum phase space*. PhD thesis, Université libre de Bruxelles, 2018.

[27] Anaelle Hertz, Nicolas J Cerf, and Stephan De Bièvre. Relating the entanglement and optical nonclassicality of multimode states of a bosonic quantum field. *Physical Review A*, 102(3):032413, 2020.

[28] Anaelle Hertz and Stephan De Bièvre. Quadrature coherence scale driven fast decoherence of bosonic quantum field states. *Physical Review Letters*, 124(9):090402, 2020.

[29] Anaelle Hertz, Michael G Jabbour, and Nicolas J Cerf. Entropy-power uncertainty relations: towards a tight inequality for all Gaussian pure states. *Journal of Physics A: Mathematical and Theoretical*, 50(38):385301, 2017.

[30] Tohya Hiroshima. Majorization criterion for distillability of a bipartite quantum state. *Physical review letters*, 91(5):057902, 2003.

[31] Otto Hölder. Ueber einen mittelwerthabsatz. *Nachrichten von der Königl. Gesellschaft der Wissenschaften und der Georg-Augusts-Universität zu Göttingen*, 1889:38–47, 1889.

[32] Michał Horodecki and Jonathan Oppenheim. Fundamental limitations for quantum and nanoscale thermodynamics. *Nature communications*, 4(1):1–6, 2013.

[33] Joe Howard. Molecular motors: structural adaptations to cellular functions. *Nature*, 389(6651):561–567, 1997.

[34] Robin L Hudson. When is the Wigner quasi-probability density non-negative? *Reports on Mathematical Physics*, 6(2):249–252, 1974.

[35] Michael G. Jabbour. *Bosonic systems in quantum information theory : Gaussian-dilatable channels, passive states, and beyond*. PhD thesis, Université libre de Bruxelles, 2018.

[36] Michael G. Jabbour and Nicolas J. Cerf. Multiparticle quantum interference in Bogoliubov bosonic transformations. *Physical Review Research*, 3:043065, Oct 2021.

[37] Michael G Jabbour, Raúl García-Patrón, and Nicolas J Cerf. Majorization preservation of Gaussian bosonic channels. *New Journal of Physics*, 18(7):073047, 2016.

[38] J Robert Johansson, Paul D Nation, and Franco Nori. Qutip: An open-source python framework for the dynamics of open quantum systems. *Computer Physics Communications*, 183(8):1760–1772, 2012.

[39] Anatole Kenfack and Karol Życzkowski. Negativity of the wigner function as an indicator of non-classicality. *Journal of Optics B: Quantum and Semiclassical Optics*, 6(10):396, 2004.

[40] Rolf Landauer. Information is physical. *Physics Today*, 44(5):23–29, 1991.

[41] Andrew Lenard. Thermodynamical proof of the gibbs formula for elementary quantum systems. *Journal of Statistical Physics*, 19(6):575–586, 1978.

[42] Ulf Leonhardt. *Essential quantum optics: from quantum measurements to black holes*. Cambridge University Press, 2010.

[43] Elliott H. Lieb. Proof of an entropy conjecture of Wehrl. *Communications in Mathematical Physics*, 62:35–41, 1978.

[44] Elliott H Lieb. Proof of an entropy conjecture of Wehrl. In *Inequalities*, pages 359–365. Springer, 2002.

[45] Rodney Loudon. *The quantum theory of light*. OUP Oxford, 2000.

[46] A. Mandilara, E. Karpov, and N. J. Cerf. Extending hudson's theorem to mixed quantum states. *Physical Review A*, 79:062302, Jun 2009.

[47] Andrea Mari and Jens Eisert. Positive wigner functions render classical simulation of quantum computation efficient. *Physical review letters*, 109(23):230503, 2012.

[48] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. 2011.

[49] Michael A Nielsen. Conditions for a class of entanglement transformations. *Physical Review Letters*, 83(2):436, 1999.

[50] Michael A Nielsen and Julia Kempe. Separable states are more disordered globally than locally. *Physical Review Letters*, 86(22):5184, 2001.

[51] Michael A Nielsen and Guifré Vidal. Majorization and the interconversion of bipartite states. *Quantum Inf. Comput.*, 1(1):76–93, 2001.

[52] Stefano Olivares. Quantum optics in the phase space. *The European Physical Journal Special Topics*, 203(1):3–24, 2012.

[53] Zbigniew Puchała, Łukasz Rudnicki, and Karol Życzkowski. Majorization entropic uncertainty relations. *Journal of Physics A: Mathematical and Theoretical*, 46(27):272002, 2013.

[54] Wiesław Pusz and Stanisław L Woronowicz. Passive states and kms states for general quantum systems. *Communications in Mathematical Physics*, 58(3):273–290, 1978.

[55] Olivier Rioul. Information theoretic proofs of entropy power inequalities. *IEEE Transactions on Information Theory*, 57(1):33–55, 2010.

[56] Juliette Rousselet, Laurence Salome, Armand Ajdari, and Jacques Prostt. Directional motion of brownian particles induced by a periodic asymmetric potential. *Nature*, 370(6489):446–447, 1994.

[57] Łukasz Rudnicki, Zbigniew Puchała, and Karol Życzkowski. Strong majorization entropic uncertainty relations. *Physical Review A*, 89(5):052115, 2014.

[58] Henry ED Scovil and Erich O Schulz-DuBois. Three-level masers as heat engines. *Physical Review Letters*, 2(6):262, 1959.

[59] Marlan O Scully. Quantum afterburner: Improving the efficiency of an ideal heat engine. *Physical Review Letters*, 88(5):050602, 2002.

[60] Claude Elwood Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[61] Uttam Singh, Michael G. Jabbour, Zacharie Van Herstraeten, and Nicolas J. Cerf. Quantum thermodynamics in a multipartite setting: A resource theory of local Gaussian work extraction for multimode bosonic systems. *Physical Review A*, 100:042104, Oct 2019.

[62] Gabor Szegő. *Orthogonal polynomials*, volume 23. American Mathematical Soc., 1939.

[63] Zacharie Van Herstraeten. Differential entropies in phase space for quantum photonic systems. Master's thesis, Université libre de Bruxelles, 2017.

[64] Zacharie Van Herstraeten and Nicolas J Cerf. Quantum Wigner entropy. *Physical Review A*, 2021.

[65] Zacharie Van Herstraeten, Michael G Jabbour, and Nicolas J Cerf. Continuous majorization in quantum phase space. *arXiv preprint arXiv:2108.09167*, 2021.

[66] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J Cerf, Timothy C Ralph, Jeffrey H Shapiro, and Seth Lloyd. Gaussian quantum information. *Reviews of Modern Physics*, 84(2):621, 2012.

[67] Alfred Wehrl. On the relation between classical and quantum-mechanical entropy. *Reports on Mathematical Physics*, 16(3):353–358, 1979.

[68] William Henry Young. On the multiplication of successions of Fourier constants. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 87(596):331–339, 1912.