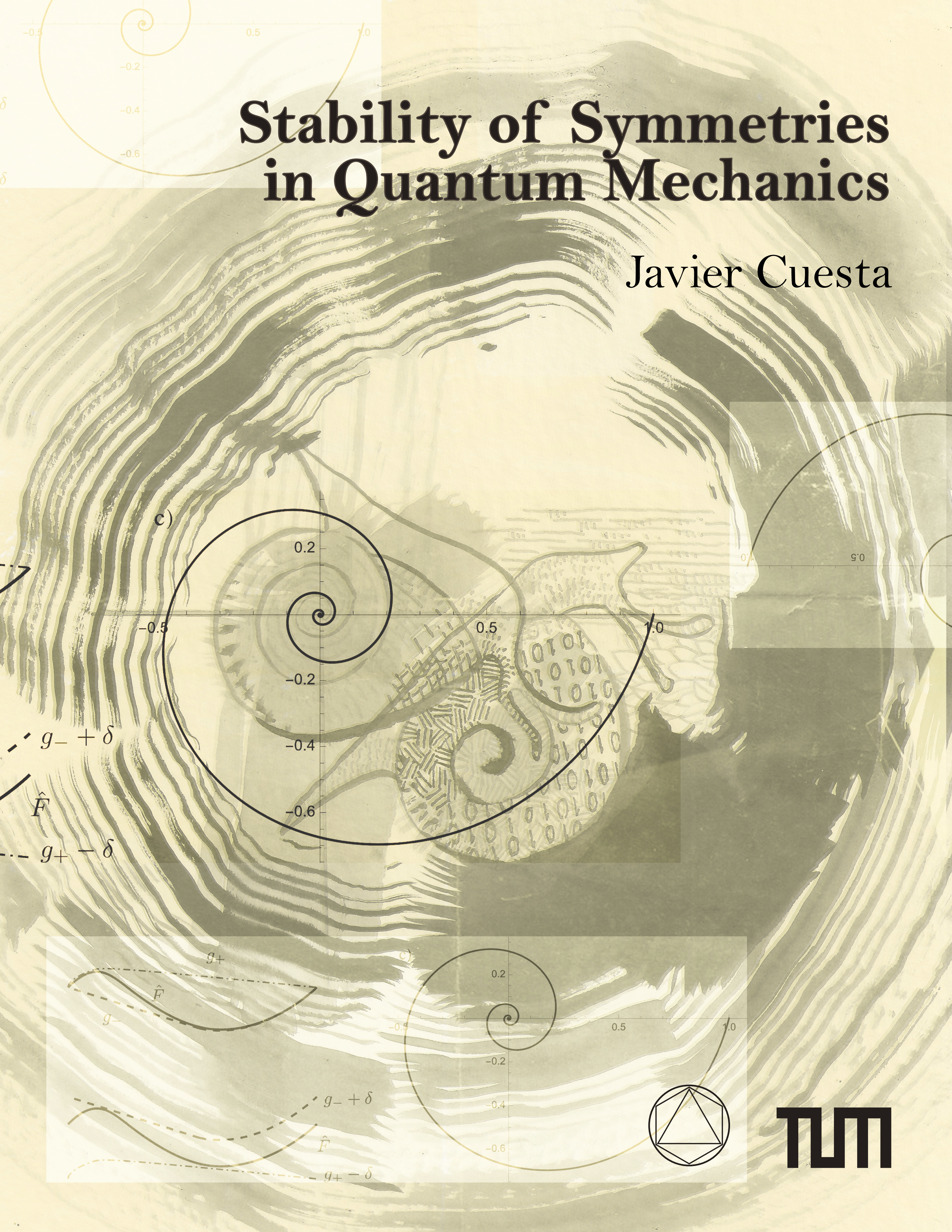
Stability of Symmetries in Quantum Mechanics

Javier Cuesta



Fakultät für Mathematik
Technische Universität München

Technische Universität München
Lehrstuhl für Mathematische Physik

Stability of Symmetries in Quantum Mechanics

Javier Enrique Cuesta Rueda

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende:
Prof. Dr. Simone Warzel

Prüfende der Dissertation:
1. Prof. Dr. Michael M. Wolf
2. Prof. Dr. David Pérez García,
   Universidad Complutense de Madrid and ICMAT

Die Dissertation wurde am 15.10.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 30.11.2019 angenommen.

Department of Mathematics
Technical Univeristy of Munich

Technical University of Munich
Chair of Mathematical Physics

Stability of Symmetries in Quantum Mechanics

Javier Enrique Cuesta Rueda

Full imprint of the dissertation approved by the Department of Mathematics of the Technical University of Munich to obtain the academic degree of

Doctor of Natural Sciences (Dr. rer. nat.)

Chairman:
Prof. Dr. Simone Warzel

Examiners of the dissertation:
1. Prof. Dr. Michael M. Wolf
2. Prof. Dr. David Pérez García,
   Universidad Complutense de Madrid and ICMAT

The dissertation was submitted to the Technical University of Munich on 15.10.2019 and was accepted by the Department of Mathematics on 30.11.2019.

Zusammenfassung

Diese Dissertation befasst sich mit der Stabilität der Symmetrien in der Quantenmechanik in zwei wichtigen Szenarien. Zunächst betrachten wir die lineare Stabilität von Wigners Symmetriesatz, welcher ein fundamentales Resultat in der theoretischen Darstellung von physikalischen Symmetrien ist. Danach widmen wir uns einer Symmetrie-Charakterisierung einer wichtigen Menge von quantenmechanischen Zuständen, nämlich gaußschen bosonischen Zuständen. Wir erforschen, wie stabil diese Charakterisierung ist, wenn wir leichte Änderungen in den Annahmen betrachten. Jedes unserer Stabilitätsresultate beinhaltet explizite Schranken inklusive exakter Angabe der Konstanten.

Abstract

This dissertation deals with the stability of symmetries in quantum mechanics in two important scenarios. First, we study the linear stability of Wigner's symmetry theorem, which is a foundational result on how physical symmetries are mathematically represented. Then, we turn our attention to a symmetry characterization of an important set of quantum states, namely, Gaussian bosonic states. We explore how stable this characterization is when we allow for small changes in the underlying assumptions. Throughout this work, we give explicit bounds on every stability result and provide exact constants.

One can learn a lot about a mathematical object by studying how it behaves under small perturbations.

– *Barry Mazur*

Acknowledgements

I would like to thank, first and foremost, my advisor Prof. Michael M. Wolf for his encouragement, patience and wisdom during these years. I cannot express enough how much I have learned during this time under his supervision. Next, I want to thank the group members of M5 (past and present) and specially Prof. Michael M. Wolf and Prof. Robert König for creating such an amazing work environment. I would like to specially thank Robert König for his professional advise as a mentor and for answering that E-mail in 2010; Silvia and Wilma for helping us with so many practical issues at the university and of course my fellow PhD students!

Special thanks go to David Peréz García for agreeing to be an examiner for this thesis and to Simone Warzel for accepting to be the chairman of the examination committee. Furthermore, I wish to thank Isabella Wiegand from the ISAM office especially for her supporting work to the PhD life at the TUM. The time spent in Garching would have not been the same, if I did not had my office mates Andreas Bluhm and Martin Idel. I am grateful for their open attitude and readiness for discussion. I am deeply grateful to Ion Nechita for sharing his passion about science and the friendly environment during discussions; I am obliged from these great interactions. I am very thankful with Cambyze Rouzé for being brave enough to hear me talk for long hours about type and cotype and accepting to co-organize a reading group about related topics. Likewise, I am very thankful with Alexander Kliesch for helping me out with some difficult German language questions.

During my PhD, I had the opportunity to visit the QMATH group in Copenhagen for a couple of months. I wish to thank Matthias Christandl and Jan Phillip Solovej for their immense hospitality and Péter Vrana, who was also visiting at that time, for many valuable discussions. I also want to thank all my friends who have supported me in various ways over the years. Specially Jens Grimm, Manuel Pusch, Manuel Trefftz, Fernando Abudinén, Juan C. Osorio, Mikel Rojo, Gustavo Lorgia and Gustavo Cipagauta; I would not be here without the support of Alonso Botero and Leonardo Pachón.

I would like to thank Nathalie Gerstner for her love and encouragement during the last stage of my PhD. Your genuine interest and your *Senf* ☺ is so appreciated. Thank you.

Finalmente quisiera agradecer a mi familia, en particular a mis padres, mi hermana Laura y hermano Andrés, que desde la distancia siempre han estado ahí para mí. Su apoyo incondicional, independiente del éxito y del anhelo a la perfección, significa todo.

List of contributed articles

This thesis is based on the following articles:

Core articles as principal author

I) Javier Cuesta and Michael Wolf
Are almost-symmetries almost linear?
J. Math. Phys. 60, 082101 (2019).
(see also article [1] in the bibliography)

II) Javier Cuesta
Type and cotype constants and the linear stability of Wigner's symmetry theorem.
Symmetry 11(9), 1107 (2019)
(see also article [2] in the bibliography)

Further articles

III) Javier Cuesta
A stable quantum Darmois-Skitovich theorem.
arXiv preprint 1902.05298 [math-ph] (2019).
Accepted for publication in *Journal of Mathematical Physics*.
(see also article [3] in the bibliography)

I, Javier Enrique Cuesta Rueda, am the principal author of articles I), II) and III).

Contents

1 Introduction

Symmetries permeate our understanding of nature and so it is difficult to conceive a theory without them. If a physical problem has an exact analytical solution, there is a great chance that there was an underlying symmetry making this possible. That is, there is a simplification of the problem manifested by the fact that "something" (e.g. a property or quantity) was preserved. These idealizations are sometimes very good approximations to the real world. But what happens when this "something" is not exactly preserved, but *almost*; can we still approximate the problem by the exact case and in that case, can we quantify this? This thesis deals with these sort of questions in relation with quantum mechanics by studying the stability of two important mathematical descriptions of symmetries in quantum theory.

1.1 Outline

In this Section we describe the "almost-symmetries" that we investigated in this thesis and give some guidelines to understand where this research question is located. In addition, an outline of the structure of the chapters is given.

Our research question deals with the stability of symmetries in quantum mechanics in two important scenarios. We first study the stability of the symmetry representation theorem of Wigner. This is a fundamental result in quantum theory as it tells us how we can mathematically represent physical symmetries in an abstract Hilbert space. This theorem set out the interest in group theoretical methods and Lie algebras in quantum physics. An exact statement of Wigner's theorem as well as a simple and self-contained proof can be found in Section 2.2. Certainly we have to clarify the notion of being stable and "almost-symmetry" used here. In Section 2.4 we explain in detail what we mean by *stability* and review how our research question stands in relation to other work.

In the field of quantum information theory (QIT) one is mainly interested in studying the advantages and limitations of information-processing tasks when we use quantum resources. In the so-called continuous variable setting one focuses in the case that the quantum states are necessarily represented in an infinite dimensional Hilbert space. This is the case of bosonic quantum systems which we describe in Section 3. Among the set of bosonic quantum states there is a particular subset of states, called Gaussian states, which play a prevailing role in continuous variable quantum information. It turns out that among the set of bosonic quantum states, Gaussian bosonic states are the only ones that satisfy a particular symmetry property. Besides the stability of Wigner's theorem, we study the stability of this symmetry that characterizes Gaussian bosonic states.

In the next Section we give a short presentation of the contributed articles and their scope. This is followed by an introduction to the main notions and formalism of quantum theory (see Chapter 2). After this, we move to an introduction of bosonic quantum systems in Chapter 3 where the symmetry property of Gaussian states is described. The key tool to describe general bosonic states will be the quantum characteristic function. We discuss the mathematical properties of the characteristic function and call attention to the similarities and differences

between the classical and quantum characteristic function. An essential tool in this thesis is the Hahn-Banach theorem which is used with great effect in two of the contributed articles. Since this theorem appears in different forms, we present them in Chapter 4 together with some modern tools of Banach space theory that are used in the contributed article A.2. In the latter article we make a connection between the problem of the stability of almost-symmetries and the geometry of Banach spaces. This will have repercusions on the quality of the approximation to an exact symmetry.
After this introduction and overview (Chapters 2 to 4), we present the contributed articles. Every article is preceded with a summary of the main results and a description of the individual contribution of the author of this thesis. The main core articles of this thesis have been accepted for publication and a permission to use them here is include it before each article.

1.2 Summary of Results

The contributed results take different approaches to the study of the stability of symmetries in quantum theory, being the common ground a mathematical analysis perspective. Core article I investigates the linear stability of Wigner's theorem. The focus here is whether an almost-symmetry operation in quantum theory can always be approximately represented by a linear map. In core article II, this problem is adressed from other perspective and a link between the stability of almost-linear maps on finite dimensional spaces and the geometry of Banach spaces is established. Article III complements the study of the stability of symmetries in quantum mechanics by studying the stability from a state point of view. In contrast with Core article I and II, here we consider an exact symmetry operation and quantum states which almost satisfy this exact symmetry. We ask then whether these states can be approximated by the ones which satisfy the exact symmetry.

Core articles as principal author

- *Article I [1]: Are almost-symmetries almost linear?*

  In this work, we investigate whether an almost-symmetry transformation can always be approximated by a linear map. Here a symmetry is understood in the most general sense in quantum physics: a transformation that preserves the statistics outcomes of an experiment. So an almost-symmetry is a transformation that almost preserves the statistics up to some small disagreement. We show that in infinite dimensional Hilbert spaces this is only possible in a weak sense and that in general the quality of the approximation has to depend on the dimension of the involved Hilbert space. The reason for the latter is that we can exhibit a non-linear map which almost preserves the statistics, but cannot be approximated by any linear map. This particular non-linear almost-symmetry will imply a lower bound on the quality of the approximation that depends logarithmically on the dimension of the involved Hilbert space. In addition, we obtain for finite dimensional systems an upper bound on the quality of the approximation which depends linearly on the dimension of the Hilbert space.

- *Article II [2]: Type and cotype constants and the linear stability of Wigner's symmetry theorem*

This work is a follow-up to Article I. Here, we improve the upper bound on the quality of the approximation of an almost-symmetry by a linear map. In order to achieve this, we develop a connection between the geometry of finite dimensional Banach spaces and the linear approximation of almost-linear maps. The quality of the approximation of an almost-linear map will depend on some geometric Banach space invariant of both the domain and codomain of the almost-linear map. This will allow to identify the quality of the approximation in a systematic way depending on which combination of domain and codomain are considered. Although this new method does not completely close the gap of the linear stability of Wigner's theorem, it provides an insight on the limitations that extending an almost-symmetry to an almost-linear map has. It turns out that there is a trade-off between the choice of domain and codomain of this extension and that the optimal choice is when both of them are Hilbert spaces. In the latter case, the order of approximation is logarithmic on the dimension of the Hilbert space. This provides a possible route to follow in order to close the gap.

Further articles

- *Article III [3]: A stable quantum Darmois-Skitovich theorem*

In the second core article, we investigate the stability of a symmetry characterization of Gaussian bosonic states. The latter set of states are important due to their extremal properties and their role in continuous variable quantum information can be put in the same level as the one of the normal distribution in classical probability. The symmetry considered here acts on a pair of quantum states and rotates the coordinates of their canonical observables. We show that if a pair of independent quantum states, i.e. whose statistic is uncorrelated, remains independent after the described symmetry transformation, then the states are necessarily Gaussian and with equal second moments. This corresponds to a quantum version of a classical theorem of Darmois and Skitovich from the 1950s. We show that this symmetry property of Gaussian states is stable in the quantum case. Namely, that states which are almost independent after the action of this symmetry can be approximated by Gaussian bosonic states. Furthermore, we give explicit bounds on the quality of this approximation which were not known even in the classical scenario.

2 Basic notions from Quantum Theory

We begin this chapter with a brief introduction to quantum mechanics from a mathematical perspective and with a focus on the representability of symmetries. The material presented here can be found in more detail in the textbooks [4, 5, 6, 7, 8]. The presentation that we choose is based on elementary concepts of quantum information theory. We will restrict sometimes our attention to finite dimensional systems. However, we will drop entirely the latter assumption in Chapter 3 and describe only quantum systems in an infinite dimensional Hilbert Space.

Let us first agree on a common notation. Consider a separable Hilbert space $\mathcal{H}$. The Adjoint of a linear operator $A : \mathcal{H} \to \mathcal{H}$ is written as A^* and the complex conjugate of $z \in \mathbb{C}$ as $\bar{z}$. The set of bounded linear operators on $\mathcal{H}$ is denoted by $\mathcal{B}(\mathcal{H})$. We follow sometimes the commonly used *bra-ket* notation: a vector ψ in $\mathcal{H}$ is written as $|\psi\rangle$. Due to the Riesz-Representation theorem (Theorem II.4 in [9]) we can identify the elements of the dual space $\mathcal{H}^*$ with vectors: we write these linear functionals as $\langle\psi|$. The inner product of two vectors $\psi, \varphi \in \mathcal{H}$ is then $\langle\psi|\phi\rangle$. The trace in $\mathcal{B}(\mathcal{H})$ is denoted by $\mathrm{Tr}[\cdot]$. For any positive operator $A \in \mathcal{B}(\mathcal{H})$, there exists a unique $B \in \mathcal{B}(\mathcal{H})$ such that $B^2 = A$. Therefore for any $A \in \mathcal{B}(\mathcal{H})$, $|A| := \sqrt{A^*A}$ is well-defined. The hermitian $p-$Schatten class is a real Banach space with norm

$$\|A\|_p := (\mathrm{Tr}\,|A|^p)^{1/p} .$$

If we consider a diagonal operator x with entries x_j (so that we can be identify it with a vector), we recover the l_p-norm

$$\|x\|_p = \left(\sum_j |x_j|^p \right)^{1/p} .$$

A rank-one operator in $\mathcal{B}(\mathcal{H})$ is written as $|x\rangle\langle z|$ and is the operator which takes the vector $|y\rangle$ to $\langle z|y\rangle|x\rangle$. We denote the unit ball of a Banach space Z by $\mathcal{B}_Z$.

2.1 Quantum states, Measurements and Symmetry

Quantum mechanics is a non-commutative probabilistic theory which, after a set of correspondence rules are established, describes the physics of the microscopic world. The way that the theoretical framework of quantum mechanics is built on implies that we cannot describe single events, but rather the result of a collection of statistical experiments. Therefore, with quantum mechanics we have only access to probabilities of the outcomes of many identically repeated experiments. It is a custom in quantum theory to divide a physical experiment in a preparation procedure and measurement part. Since the probability outcomes of a measurement must be independent of the specific experimental set up, we need to introduce a new concept which carries the information of the specific probability distribution associated to the observable quantity that is to be measured. The equivalence class of preparation procedures which generate the same set of probability outcomes for an specific observable measurement is called a *state preparation procedure*. If the probability outcomes of two different preparation procedures are the same, we say that the system being measured was in the same *state*. We describe quantum states by a positive trace-class operator $\rho \in \mathcal{B}(\mathcal{H})$ with trace $\operatorname{Tr}\rho = 1$; This operator is known as "density operator" and to each quantum state there corresponds a unique density operator. One can analogously describe the state of a physical system by a *ray* in the Hilbert space $\mathcal{H}$. That is two vectors $|\psi\rangle, |\varphi\rangle \in \mathcal{H}$ describe the same state if and only if $|\psi\rangle = \lambda|\varphi\rangle$ for some $\lambda \in \mathbb{C}$. A *pure state* $\psi \in \mathcal{H}$ is a ray of norm one $\|\psi\|_2 = 1$. They can be identified in terms of density operators as rank-one projections, i.e. $\rho = |\psi\rangle\langle\psi|$. This simple identification will be important for a transparent and simple proof of Wigner's theorem (See Theorem 2.2.2). The set of density operators in $\mathcal{B}(\mathcal{H})$ corresponding to pure states is denoted by $\mathbb{P}(\mathcal{H})$. A pure state can model, for instance, a non-degenerate energy level of an atom or more generally any isolated system. A density operator which is not pure is called *mixed*.

The set of quantum states is a convex set and its extremal points are the pure states. It is important to remark that in the quantum setting a general mixed state is not uniquely determined by the set of pure states as opposed to the set of classical states which are described by a simplex; For $\mathcal{H} = \mathbb{C}^d$ it suffices to consider the *maximally mixed state* $\rho = \mathbb{1}_d/d$ which has infinitely many decompositions. The smallest quantum system is a two-level system known as *qubit*. Qubits are mathematically represented on $\mathcal{H} = \mathbb{C}^2$ and their states can be neatly described by introducing the pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The pauli matrices together with the identity $\mathbb{1}_2$ form a basis in $\mathbb{C}^{2\times 2}$. By direct computation, one obtains that any qubit is described by the density operator $\rho = \frac{1}{2}(\mathbb{1}_2 + \eta \cdot \sigma)$ where $\eta \in \mathcal{B}_{\mathbb{R}^3}$ and $\eta \cdot \sigma = \sum_{j=1}^{3} \eta_j \sigma_j$. Therefore, for every qubit there corresponds a unique vector η in the unit Ball $\mathcal{B}_{\mathbb{R}^3}$. This unit ball is the so-called "Bloch-Ball". See Figure 2.1. Pure qubit states correspond then to vectors with $\|\eta\|_2 = 1$, as they are the extremal points of the unit ball, and the maximally mixed state to $\eta = 0$.

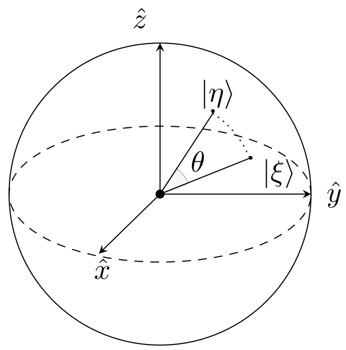


Figure 2.1: Bloch Ball. The state $|\eta\rangle$ of a qubit is represented as a vector $\eta \in \mathbb{R}^3$ on the euclidean unit ball. The density operator corresponding to the state $|\eta\rangle$ is $|\eta\rangle\langle\eta| = \frac{1}{2}(\mathbb{1}_2 + \eta \cdot \sigma)$. For two pure qubits $|\eta\rangle$ and $|\xi\rangle$ the transition probability is given by $\operatorname{Tr}[|\eta\rangle\langle\eta|\xi\rangle\langle\xi|] = \frac{1}{2}(1 + \eta \cdot \xi) = \cos^2(\theta/2)$ where θ is the angle between the vectors η and ξ.

We describe now the measurement procedure on finite dimensional systems, i.e. for $\mathcal{H} = \mathbb{C}^d$, $d \in \mathbb{N}$. Any observable is required to be represented by a self-adjoint operator $A \in \mathcal{B}(\mathcal{H})$ acting on $\mathcal{H}$. From the spectral decomposition, we write

$$A = \sum_{j=1}^{d} \lambda_j |a_j\rangle\langle a_j|, \qquad \lambda_j \in \mathbb{R}, \quad |a_j\rangle\langle a_j| \in \mathbb{P}(\mathcal{H}),$$

where $\langle a_j|a_l\rangle = \delta_{jl}$ and $\sum_{j=1}^{d} |a_j\rangle\langle a_j| = \mathbb{1}_d$. Abstractly, if $\Sigma = \{1, \ldots, n\}$ is the set that labels the measurement outcomes we define a *measurement* to be a mapping $M : \Sigma \to \mathcal{B}(\mathcal{H})$ such that *i)* $M(j) \geq 0$ for all $j \in \Sigma$ and *ii)* $\sum_{j=1}^{n} M(j) = \mathbb{1}_n$. This general type of measurement is known as a *Positive Operator-Valued Measurement* (POVM). If in addition, we impose that $M(j)^2 = M(j)$ for all $j \in \Sigma$, i.e. that $M(j)$ is a projection, then M is said to be a *sharp or projective measurement.* An important theorem of Naimark (see Theorem 2.42 in [10]) states that any POVM can be viewed as a Von-Neumann measurement on a larger system that includes the original system as a subsystem. In quantum mechanics it is postulated that the probability $p(j)$ of measuring an outcome $j \in \Sigma$ is

$$p(j|M, \rho) := \mathrm{Tr}[M(j)\rho],$$

if the measurement and preparation are described by a POVM M and the density operator $\rho \in \mathcal{B}(\mathcal{H})$, respectively. A relevant example of a sharp measurement is when $M(j)$ is a rank-one projection for every $j \in \Sigma$. In this case, $\{M(j)\}_{j=1}^{n}$ is called a *von Neumann measurement.* Thus an observable $A = \sum_{j=1}^{d} \lambda_j |a_j\rangle\langle a_j|$ gives rise to a von-Neumann measurement with measurement outcomes $\{\lambda_1, \ldots, \lambda_d\}$ and measurements $M(j) = |a_j\rangle\langle a_j|, j = 1, \ldots, d$. In this case the probability of obtaining the measurement outcome λ_j given the prepared state ρ is $p(j|M, \rho) = \mathrm{Tr}[|a_j\rangle\langle a_j|\rho] = \langle a_j|\rho|a_j\rangle$. Given two pure states $|\psi\rangle\langle\psi|, |\varphi\rangle\langle\varphi| \in \mathbb{P}(\mathcal{H})$ it is customary to call the overlap

$$\mathrm{Tr}[|\psi\rangle\langle\psi|\varphi\rangle\langle\varphi|] = |\langle\psi|\varphi\rangle|^2$$

the *transition probability* between the state $|\psi\rangle$ and $|\varphi\rangle$. The distance between two pure states in terms of the Schatten-norms depends directly on the transition probability. In particular (see Lemma 2.62 in [5]),

$$\begin{aligned} \| |\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi| \|_1 &= 2\sqrt{1 - |\langle\psi|\varphi\rangle|^2}, \\ \| |\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi| \|_2 &= \sqrt{2(1 - |\langle\psi|\varphi\rangle|^2)}, \\ \| |\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi| \|_\infty &= \sqrt{1 - |\langle\psi|\varphi\rangle|^2}. \end{aligned} \tag{2.1}$$

2.2 Wigner's symmetry theorem

It is well known and accepted that the laws of physics are invariant in every inertial reference frame. In particular, the outcomes of a sequence of experiments might be different depending on the observer, but the transition probabilities of the outcomes (possibly with different labels) must be the same between observers. This principle can be stated as: "the statistics of any experiment is the same in any inertial frame"[6]. This section aims to explain how we mathematically represent a symmetry transformation that preserves the probability outcomes of an experiment.

Let us consider an isolated quantum system whose state $\rho \in \mathcal{B}(\mathbb{C}^d)$ may be described by the pure state $\rho = |\varphi\rangle\langle\varphi|$. The state of the system is ought to be measured so a physicist describes the measurement procedure by a set of von-Neumann measurements $\{M(j) = |\psi_j\rangle\langle\psi_j|\}_{j=1}^d \subset \mathbb{P}(\mathbb{C}^d), \sum_{j=1}^d |\psi_j\rangle\langle\psi_j| = \mathbb{1}_d$ (see Section 2.1). The probability that the state ρ is in the state $|\psi_j\rangle\langle\psi_j|$ is then $p(j|M,\rho) = |\langle\psi_j|\varphi\rangle|^2$. Now, another physicist sees the same experiment, but describes the state and measurement differently as $\rho' = |\psi_j'\rangle\langle\psi_j'| \in \mathcal{B}(\mathbb{C}^d)$, $\{M'(j) = |\psi_j'\rangle\langle\psi_j'|\}_{j=1}^d \subset \mathbb{P}(\mathbb{C}^d)$, respectively. Although he represents the pure state of the system and measurements differently, he obtains the same probability outcomes

$$p(j|M,\rho) = |\langle\psi_j|\varphi\rangle|^2 = |\langle\psi_j'|\varphi'\rangle|^2 = p(j|M',\rho').$$

So what is the relation between $|\varphi\rangle$ and $|\varphi'\rangle$ as well as between $|\varphi_j\rangle$ and $|\varphi_j'\rangle$? More generally, what sort of transformations $g : \mathbb{C}^d \to \mathbb{C}^d$ are allowed such that

$$|\langle\psi|\varphi\rangle|^2 = |\langle g(\psi)|g(\varphi)\rangle|^2,$$

holds for all $\psi, \varphi \in \mathbb{C}^d$ with $\|\psi\|_2 = \|\varphi\|_2 = 1$. This question is answered by a celebrated result of Wigner. Before stating the theorem in its full generality, we recall that a mapping $U : \mathcal{H} \to \mathcal{H}$ is *antilinear* if

$$U(\psi + z\varphi) = U\psi + \bar{z}U\varphi,$$

for all $\psi, \varphi \in \mathcal{H}$ and $z \in \mathbb{C}$. A linear *isometry* $U : \mathcal{H} \to \mathcal{H}$ is a map such that $\langle U\psi, U\psi\rangle = \langle\psi,\psi\rangle$ for all $\psi \in \mathcal{H}$. If U is an invertible isometry, as is the case $\mathcal{H} = \mathbb{C}^d$, then U is called a unitary operator. Thus an antiunitary operator U has the property that $\langle U\psi, U\varphi\rangle = \langle\varphi,\psi\rangle$ for all $\psi, \varphi \in \mathcal{H}$. See section 2.3.1 in [5] and section 3.3 in [8] for a deeper discussion of unitary and antiunitary operations in quantum mechanics.

Theorem 2.2.1 (General Wigner)**.** *Let $\mathcal{H}$ be any separable or non-separable complex Hilbert space and $f : \mathbb{P}(\mathcal{H}) \to \mathbb{P}(\mathcal{H})$ a map that preserves transition probabilities, i.e. such that*

$$\operatorname{Tr} f(X)f(Y) = \operatorname{Tr} XY \qquad \textit{for all } X, Y \in \mathbb{P}(\mathcal{H}). \tag{2.2}$$

Then there exists a linear or antilinear isometry $U : \mathcal{H} \to \mathcal{H}$ such that $f(X) = UXU^*$.

To find a complete rigorous, but *simple*, proof of theorem 2.2.1 has been for a long time of great interest to mathematicians. Although Wigner itself did not give a full proof of this result (see pp. 233-236 in [11]), he inspired subsequent work on this topic. Bargmann [12] gave 30 years after Wigner's original idea a proof which works for the case where $\mathcal{H}$ was separable and f was not necessarily bijective. It is remarkable, that not until very recently an elementary proof has been obtained for the general case [13]. We provide a simple self-contained proof of Wigner's theorem in the next section for the case $\mathcal{H} = \mathbb{C}^d$.

Wigner's theorem has a straightforward interpretation (and proof) in the qubit case. It states that the only possible symmetries that preserve the angle formed by two pure states on the bloch sphere are either the group of rotations or the discrete group of reflections with respect to the xz-plane. See Figure 2.1.

It is clear that unitaries preserve transition probabilities. However what makes theorem 2.2.1 remarkable is that *apriori* there is no reason to believe that from the preservation of transition probabilities the map should be even linear. Furthermore, it is worth to note that the symmetry condition on the transition probabilities is only required to hold on the set of pure states. The result extends to the full algebra of $\mathcal{B}(\mathcal{H})$.

Proof of Wigner's theorem

In this section we provide a self-cointained proof of Wigner's theorem based on the ideas of [13, 14] and [15]. This proof highlights two important features of a Wigner symmetry: (i) it is a linear operation at the level of self-adjoint operators and (ii) it is a Jordan *-homomorphism on $\mathcal{B}(\mathcal{H})$. We choose to present a proof for a finite dimensional Hilbert space, but the proof could be in principle extended to any separable Hilbert space by considering more carefully some limits. We do not dwell with this here as we will have enough reasons to remain in the finite dimensional case (see Section A.1).

Theorem 2.2.2 (Wigner)**.** *Let $f : \mathbb{P}(\mathbb{C}^d) \to \mathbb{P}(\mathbb{C}^d)$ be a map that preserves transition probabilities, i.e. such that*

$$\operatorname{Tr} f(X)f(Y) = \operatorname{Tr} XY \qquad \text{for all} \quad X, Y \in \mathbb{P}(\mathbb{C}^d). \tag{$\mathbf{S}$}$$

Then there exists a linear or antilinear unitary $U : \mathbb{C}^d \to \mathbb{C}^d$ such that $f(X) = UXU^*$.

Proof. **Step 1 (Linear extension):** As in lemma 1 of [15], we extend f uniquely to $\mathcal{H}_d$ and then to $\mathbb{C}^{d\times d}$. For $X \in \mathcal{H}_d$ write the spectral decomposition $X = \sum_k \lambda_k X_k$ where $X_k \in \mathbb{P}(\mathbb{C}^d)$ and define $F : \mathcal{H}_d \to \mathcal{H}_d$ by $F(X) := \sum_k \lambda_k f(X_k)$. Every $M \in \mathbb{C}^{d\times d}$ can be uniquely decomposed as $M = X + iY$ with $X, Y \in \mathcal{H}_d$. Thus we define $\tilde{F} : \mathbb{C}^{d\times d} \to \mathbb{C}^{d\times d}$ by $\tilde{F}(M) = F(X) + iF(Y)$. For two hermitian matrices with spectral decomposition $X = \sum_k \lambda_k X_k$ and $Y = \sum_k \mu_l Y_l$, Wigner's symmetry condition is extended to

$$\operatorname{Tr} F(X)F(Y) = \sum_{k,l} \lambda_k \mu_l \operatorname{Tr} f(X_k)f(Y_l) = \sum_{k,l} \lambda_k \mu_l \operatorname{Tr} X_k Y_l = \operatorname{Tr} XY.$$

Likewise for $M, N \in \mathbb{C}^{d\times d}$,

$$\operatorname{Tr} \tilde{F}(N)\tilde{F}(M) = \operatorname{Tr} NM.$$

The maps F and $\tilde{F}$ are unique linear extensions of f. Indeed, let us assume that $X = \sum_k \lambda_k X_k$ has another spectral decomposition $Y = \sum_k \lambda_k Y_k$, $\{Y_k\} \subset \mathbb{P}(\mathbb{C}^d)$ such that $X = Y$. Then for any $Z \in \mathbb{P}(\mathbb{C}^d)$ we compute using linearity of the trace and Eq. ($\mathbf{S}$)

$$\begin{aligned}\operatorname{Tr} \sum_k \lambda_k f(X_k)f(Z) = \operatorname{Tr} \sum_k \lambda_k X_k Z &= \operatorname{Tr} \sum_k \lambda_k Y_k Z. \\ &= \operatorname{Tr} \sum_k \lambda_k F(Y_k)F(Z)\end{aligned}$$

Since $Z \in \mathbb{P}(\mathbb{C}^d)$ was arbitrary, again from linearity of the trace we obtain $\|\sum_k \lambda_k f(X_k) - \sum_k \lambda_k f(Y_k)\|_2^2 = 0$ and therefore $\sum_k \lambda_k f(X_k) = \sum_k \lambda_k f(Y_k)$ as claimed.

Step 2 (Jordan *-homomorphism) A Jordan $*-$homomorphism is a linear map T such that i) $T(A^*) = T(A)^*$ and ii) $T(A^2) = T(A)^2$. Property (i) is immediately fulfilled by $\tilde{F}$ by definition, so we focus on the second property of Jordan homomorphisms. For a unit vector $|\psi\rangle \in \mathcal{H}$ we use the short-hand notation $\mathbf{P}[|\psi\rangle] := |\psi\rangle\langle\psi|$. Since f maps rank-1 projections into rank-1 projections denote by $\tilde{f} : \mathbb{C}^d \to \mathbb{C}^d$ the associated map to f such that $f(|\psi\rangle\langle\psi|) = |\tilde{f}(\psi)\rangle\langle\tilde{f}(\psi)|$. Then it follows from Eq. ($\mathbf{S}$) that if $\{|e_j\rangle\}_{j=1}^d \subset \mathbb{C}^d$ is an orthonormal set, so is $\{|\tilde{f}(e_j)\rangle\}_{j=1}^d \subset \mathbb{C}^d$. Therefore with the spectral decomposition of X as above

$$\left(F\left(\sum_k \lambda_k X_k\right)\right)^2 = \sum_{k,l} \lambda_k \lambda_l f(X_k)f(X_l) = \sum_{k,l} \lambda_k^2 f(X_k) = F\left(\sum_k \lambda_k^2 X_k\right),$$

which shows property ii) for F. Using this property for $A = X + Y$ one obtains that $F(XY + YX) = F(X)F(Y) + F(Y)F(X)$. This is already enough to show that $\tilde{F}$ is a Jordan $*-$homomorphism as for $M = X + iY$, $M^2 = X^2 - Y^2 + i(XY + YX)$.

At this point, one could follow different routes: show that F is either a homomorphism or antihomomorphism (Herstein's theorem [16]) or show directly that $\tilde{F}$ preserves rank and then use the respective linear preserver characterization (Hou's theorem [17] or the fundamental theorem of projective geometry). However, we plan to present an elementary construction as in [13] by constructing the sought unitaries.

Step 3 (W and Fixed points) As noted in Step 2 the map f takes an orthonormal system $\{|e_j\rangle\}_{j=1}^d \subset \mathbb{C}^d$ to another orthonormal system $\{|\tilde{f}(e_j)\rangle\}_{j=1}^d \subset \mathbb{C}^d$. We define a unitary $W : \mathbb{C}^d \to \mathbb{C}^d$ such that $W|e_j\rangle = |\tilde{f}(e_j)\rangle$, $j = 1, \ldots, d$ and use for any map g the short-hand notation $g_W(\cdot) := W^* g(\cdot) W$. From now on, we fix an arbitrary basis $\{|n\rangle\}_{n=1}^d \in \mathbb{C}^d$ and consider the respective unitary W such that $f_W(\mathbf{P}[|n\rangle]) = \mathbf{P}[|n\rangle]$.

Let $|\psi\rangle = \sum_{n=1}^d c_n|n\rangle$ be an arbitrary vector in $\mathbb{C}^d$ and let $f_W(\mathbf{P}[|\psi\rangle]) = \mathbf{P}[|\phi\rangle]$. Thus, if $|\phi\rangle = \sum_{n=1}^d c'_n|n\rangle$ we obtain from Eq. ($\mathbf{S}$)

$$\begin{aligned} |c_n|^2 = \operatorname{Tr}\mathbf{P}[|\psi\rangle]\,\mathbf{P}[|n\rangle] &= \operatorname{Tr} f_W(\mathbf{P}[|\psi\rangle]) f_W(\mathbf{P}[|n\rangle]), \\ &= \operatorname{Tr}\mathbf{P}[|\phi\rangle]\,\mathbf{P}[|n\rangle] = |c'_n|^2. \end{aligned}$$

This implies together with Eq. ($\mathbf{S}$) that for all different $n, m \in \{1, \ldots, d\}$

$$\begin{aligned} f_W\left(\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle \pm |m\rangle)\right]\right) &= \mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle \pm \lambda_{n,m}|m\rangle)\right], \\ f_W\left(\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle \pm i|m\rangle)\right]\right) &= \mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle \pm \omega_{n,m}|m\rangle)\right], \end{aligned}$$

where $\lambda_{nm}, \omega_{nm} \in S_d$ are unit complex vectors.

Step 4 (Unitarity vs Anti-unitarity)

Now we show the existing relation between λ_{nm} and ω_{nm}. For that matters, compute using Eq. ($\mathbf{S}$)

$$\begin{aligned} 1 &= \operatorname{Tr}\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle + |m\rangle)\right]\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle + i|m\rangle)\right] \\ &= \operatorname{Tr} f_W\left(\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle + |m\rangle)\right]\right) f_W\left(\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle + i|m\rangle)\right]\right), \\ &= \operatorname{Tr}\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle + \lambda_{n,m}|m\rangle)\right]\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle + \omega_{n,m}|m\rangle)\right], \\ &= \frac{1}{2}\,|1 + \lambda_{n,m}\overline{\omega_{n,m}}|^2. \end{aligned}$$

The equation $|1 + \lambda_{n,m}\overline{\omega_{n,m}}|^2 = 2$ has only two solutions: either $\omega_{n,m} = i\lambda_{n.m}$ or $\omega_{n,m} = -i\lambda_{n,m}$.

Step 5 (Preservation of Rank-2 and its consequences) For any $z \in \mathbb{C}$, $|z| = 1$ we have the following spectral decomposition

$$\frac{\overline{z}|n\rangle\langle m| + z|m\rangle\langle n|}{2} = \frac{1}{2}\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle + z|m\rangle)\right] - \frac{1}{2}\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle - z|m\rangle)\right].$$

Accordingly, using the definition of F

$$\begin{aligned} F_W\left(\frac{|n\rangle\langle m| + |m\rangle\langle n|}{2}\right) &= \frac{1}{2}f_W\left(\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle + |m\rangle)\right]\right) - \frac{1}{2}f_W\left(\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle - |m\rangle)\right]\right), \\ &= \frac{1}{2}\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle + \lambda_{n,m}|m\rangle)\right] - \frac{1}{2}\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle - \lambda_{n,m}|m\rangle)\right], \\ &= \frac{\overline{\lambda_{n,m}}|n\rangle\langle m| + \lambda_{n,m}|m\rangle\langle n|}{2}. \end{aligned}$$

Likewise,

$$F_W\left(\frac{|n\rangle\langle m| - |m\rangle\langle n|}{2i}\right) = \frac{\overline{\omega_{n,m}}|n\rangle\langle m| + \omega_{n,m}|m\rangle\langle n|}{2}.$$

Moreover, depending if $\omega_{n,m} = \pm i\lambda_{n,m}$

$$F_W\left(\frac{|n\rangle\langle m| - |m\rangle\langle n|}{2i}\right) = \pm\frac{\overline{\lambda_{n,m}}|n\rangle\langle m| - \lambda_{n,m}|m\rangle\langle n|}{2i}.$$

As a consequence, if $\omega_{n,m} = i\lambda_{n,m}$

$$\begin{aligned} \tilde{F}_W(|n\rangle\langle m|) &= F_W\left(\frac{|n\rangle\langle m| + |m\rangle\langle n|}{2}\right) + iF_W\left(\frac{|n\rangle\langle m| - |m\rangle\langle n|}{2i}\right), \\ &= \frac{\overline{\lambda_{n,m}}|n\rangle\langle m| + \lambda_{n,m}|m\rangle\langle n|}{2} + \frac{\overline{\lambda_{n,m}}|n\rangle\langle m| - \lambda_{n,m}|m\rangle\langle n|}{2}, \\ &= \overline{\lambda_{n,m}}|n\rangle\langle m|. \end{aligned}$$

Similarly, if $\omega_{n,m} = -i\lambda_{n,m}$

$$\tilde{F}_W(|n\rangle\langle m|) = \lambda_{n,m}|m\rangle\langle n|.$$

Now we prove that if $\omega_{n,m} = i\lambda_{n,m}$ or $\omega_{n,m} = -i\lambda_{n,m}$ is satisfied for some pair $n \neq m$ then it is true for all different pairs $n, m \in \{1, \dots, d\}$. To see this, we show first that if one of the phase equations is satisfied for some $n \neq m$, then that equation is also satisfied for the same n and all $n \neq m \in \{1, \dots, d\}$. Indeed, assume that $\omega_{n,k} = i\lambda_{n,k}$ and $\omega_{n,m} = -i\lambda_{n,m}$ for arbitrary different n, k, m. Then with $A := |n\rangle\langle k| + |n\rangle\langle m|$, $A^2 = 0$, but $F(A)^2 = \left(\overline{\lambda_{n,k}}|n\rangle\langle k| + \lambda_{n,m}|m\rangle\langle n|\right)^2 = \lambda_{n,m}\overline{\lambda_{n,k}}|m\rangle\langle k|$ which contradicts that $F(A^2) = F(A)^2$. Analogously, if $\omega_{n,m} = \pm i\lambda_{n,m}$ for some n, m, then $\omega_{k,m} = \pm i\lambda_{k,m}$ for the same m and all $m \neq k \in \{1, \dots, d\}$. Combining these last observations we have that if $\omega_{n,m} = \pm i\lambda_{n,m}$ for some $n \neq m$, then $\omega_{k,l} = \pm i\lambda_{k,l}$ for all $l \neq k$.

Step 6 (Multiplicativity of phases) We show in this step that $\lambda_{n,m} = \prod_{k>n}^{m} \lambda_{k-1,k}$. Without lost of generality assume that $\omega_{n,m} = i\lambda_{n,m}$ as the other case is analogous. First, note that for different $n, k, m \in \{1, \ldots, d\}$

$$\begin{aligned}
|\lambda_{n,m} - \lambda_{n,k}\lambda_{k,m}| = \left\|\overline{\lambda_{n,m}}|n\rangle\langle m| - \overline{\lambda_{n,k}\lambda_{k,m}}|n\rangle\langle m|\right\| &= \left\|\overline{\lambda_{n,m}}|n\rangle\langle m| - \left(\overline{\lambda_{n,k}}|n\rangle\langle k| + \overline{\lambda_{k,m}}|k\rangle\langle m|\right)^2\right\|, \\
&= \left\|\tilde{F}(|n\rangle\langle m|) - \left(\tilde{F}(|n\rangle\langle k| + |k\rangle\langle m|)\right)^2\right\|, \\
&= \left\|\tilde{F}((|n\rangle\langle k| + |k\rangle\langle m|)^2) - \left(\tilde{F}(|n\rangle\langle k| + |k\rangle\langle m|)\right)^2\right\|, \\
&= 0,
\end{aligned}$$

as $\tilde{F}$ is a Jordan $*-$homomorphism. Then $\lambda_{n,m} = \lambda_{n,k}\lambda_{k,m}$ and applying this recursively gives $\lambda_{n,m} = \lambda_{n,n+1}\lambda_{n+1,n+2}\ldots\lambda_{m-1,m}$.

Step 7 (Conclusion) Define the unitary transformation $V : \mathbb{C}^d \to \mathbb{C}^d$ by

$$V|n\rangle = \left(\prod_{k=1}^{n} \lambda_{k-1,k}\right)|n\rangle,$$

where $\lambda_{01} := 1$ and

$$\lambda_{n,m} := \operatorname{Tr} f_W\left(\mathbf{P}\left[\frac{1}{\sqrt{2}}(|n\rangle + |m\rangle)\right]\right)|n\rangle\langle m| \ \in S_1.$$

From Step 6 we have $V|n\rangle\langle m|V^* = \prod_{k>n}^{m} \lambda_{k-1,k}|n\rangle\langle m| = \lambda_{nm}|n\rangle\langle m|$ and so

$$\begin{aligned}
F_{WV}\left(\frac{|n\rangle\langle m| + |m\rangle\langle n|}{2}\right) &= \frac{|n\rangle\langle m| + |m\rangle\langle n|}{2}, \\
F_{WV}\left(\frac{|n\rangle\langle m| - |m\rangle\langle n|}{2i}\right) &= \pm\frac{|n\rangle\langle m| - |m\rangle\langle n|}{2i}.
\end{aligned}$$

Then since F (and $\tilde{F}$) is an homogeneous function, the map F_{WV} leaves invariant the standard Hilbert-Schmidt orthonormal basis of $\mathcal{H}_d$, that is

$$\left\{|n\rangle\langle n|, \frac{|n\rangle\langle m| + |m\rangle\langle n|}{\sqrt{2}}, \frac{|n\rangle\langle m| - |m\rangle\langle n|}{\sqrt{2}i} : n, m = 1, \ldots, d, n < m\right\}.$$

Consequently, the transformation F_{WV} acts as either the identity map or the transpose map on $\mathcal{H}_d$ and thus with $U := WV$

$$f(X) = UXU^* \qquad \text{or} \qquad f(X) = UX^TU^*.$$

□

2.3 Composite systems and Quantum channels

A system composed of more than one part is described in quantum mechanics by a tensor product. If $\mathcal{H}_A$ and $\mathcal{H}_B$ are two Hilbert spaces describing two subsystems A and B, respectively, then $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ is the Hilbert space describing the joint system AB. Moreover, $\dim(\mathcal{H}) = \dim(\mathcal{H}_A) \cdot \dim(\mathcal{H}_B)$. A state $\rho_{AB} \in \mathcal{B}(\mathcal{H}_{AB})$ is a *product state* if it can be written as $\rho_{AB} = \rho_A \otimes \rho_B$ with $\rho_A \in \mathcal{H}_A$ and $\rho_B \in \mathcal{H}_B$. A *separable state* is a state that can be written as convex combination of product states. If $\rho_{AB} \in \mathcal{B}(\mathcal{H}_{AB})$ is not a separable state, then we say that ρ_{AB} is *entangled*. The *maximally entangled state* in $\mathbb{C}^d \otimes \mathbb{C}^d$ is defined as the state

$$|\Omega\rangle := \frac{1}{\sqrt{d}} \sum_{k=1}^{d} |kk\rangle$$

where $|kk\rangle$ is short for $|k\rangle \otimes |k\rangle$ and $\{|k\rangle\}_{k=1}^{d}$ is an orthonormal basis in $\mathbb{C}^d$. The *reduced density operator* $\rho_A \in \mathcal{B}(\mathcal{H}_A)$ or marginal of $\rho \in \mathcal{B}(\mathcal{H}_{AB})$ with respect to the subsystem A is defined via

$$\mathrm{Tr}[\rho_A X] = \mathrm{Tr}[\rho(X \otimes \mathbb{1})] \qquad \text{for all } X \in \mathcal{B}(\mathcal{H}_A).$$

A linear map $T \in \mathfrak{B}(\mathbb{C}^{d\times d})$ is positive if $T(X^*X) \geq 0$ for all $X \in \mathbb{C}^{d\times d}$. It turns out that one needs a stronger condition if it is to guarantee that when the input of T is a quantum state the output will be also a genuine quantum state. The required condition is called completely positivity: a linear map $T \in \mathfrak{B}(\mathbb{C}^{d\times d})$ is *completely positive* if the map $T \otimes \mathrm{id}_n$ is positive for all $n \in \mathbb{N}$ where id_n is the identity on $\mathbb{C}^{n\times n}$. We use sometimes as well $\mathbb{1}_n$ for the identity on $\mathbb{C}^{n\times n}$. The following result [18, 19] is used to decide if a linear map is completely positive. A linear map T is completely positive if and only if the operator $\tau := T \otimes \mathrm{id}(|\Omega\rangle\langle\Omega|)$, known as Jamiołkowski state, is positive. If $\mathrm{Tr}\, T(X) = \mathrm{Tr}\, X$ for all $X \in \mathbb{C}^{d\times d}$ then T is called *trace preserving*. A trace preserving completely positive linear map T is called a *quantum channel*.

Given a Jamiołkowski state $\tau \in \mathcal{B}(\mathbb{C}^d \otimes \mathbb{C}^d)$ we can recover the corresponding linear map $T \in \mathcal{B}(\mathbb{C}^{d\times d})$ via

$$\mathrm{Tr}\, |i\rangle\langle j| T(B) = \langle j|T(B)|i\rangle = d\, \mathrm{Tr}\, \tau |i\rangle\langle j| \otimes B^T.$$

The following are important examples of quantum channels: the *completely depolarizing channel*

$$T(A) = \mathrm{Tr}[A]\frac{\mathbb{1}}{d}, \qquad \text{with } \tau = \frac{\mathbb{1}}{d} \otimes \frac{\mathbb{1}}{d}.$$

The *identity channel*

$$T(A) = A, \qquad \text{with } \tau = |\Omega\rangle\langle\Omega|.$$

The *diagonal channel*

$$T(A) = \mathrm{diag}(A) := \sum_k \langle k|A|k\rangle |k\rangle\langle k|,$$

$$\text{with} \quad \tau = \sum_{k=1}^{d} \frac{1}{d} |k\rangle\langle k| \otimes |k\rangle\langle k|.$$

With these channels we can construct other quantum channels such as the *depolarizing channel*

$$T(A) = \lambda A + (1-\lambda)\,\mathrm{Tr}[A]\frac{\mathbb{1}}{d}, \qquad -\frac{1}{d^2-1} \leq \lambda \leq 1$$

$$\text{with} \quad \tau = \lambda|\Omega\rangle\langle\Omega| + (1-\lambda)\frac{\mathbb{1}}{d} \otimes \frac{\mathbb{1}}{d}.$$

We can also construct new channels out of symmetries: the *Werner-Holevo Channel* [20] $T_{WH} \in \mathcal{B}(\mathbb{C}^{d\times d})$ is the quantum channel whose Jamiołkowski state τ commutes with all the unitary operators of the form $U \otimes U$ where U is a unitary matrix in $\mathbb{C}^d$. For $\lambda \in [0,1]$

$$
\begin{aligned}
T_{WH} &= \lambda T_{WH}^{Sym} + (1-\lambda) T_{WH}^{Asym}, \\
T_{WH}^{Sym} &:= \frac{1}{d+1}\left(\mathrm{Tr}[A]\,\mathbb{1}_d + A^T\right), \\
T_{WH}^{Asym} &:= \frac{1}{d+1}\left(\mathrm{Tr}[A]\,\mathbb{1}_d - A^T\right).
\end{aligned}
$$

The Jamiołkowski state τ corresponding to T_{WH} is also called the *Werner-state* [21]

$$
\begin{aligned}
\tau &= \lambda \frac{2}{d(d+1)} \Pi^{Sym} + (1-\lambda) \frac{2}{d(d+1)} \Pi^{Asym}, \\
\Pi^{Sym} &:= \frac{1}{2}\left(\mathbb{1} + \mathbb{F}\right), \qquad \Pi^{Asym} := \frac{1}{2}\left(\mathbb{1} - \mathbb{F}\right),
\end{aligned}
$$

where $\mathbb{F} \in \mathcal{B}(\mathbb{C}^d \otimes \mathbb{C}^d)$ is the flip operator, which satisfies $\mathbb{F}|kl\rangle = |lk\rangle$. The following symmetric channel includes a parametrization, Eq. (2.3), which will play an important role in Section A.1 (cf. Lemma 1 in [22]).

Lemma 2.3.1. *Let G be the group generated by all diagonal unitaries and permutation matrices in $\mathbb{C}^{d\times d}$ and $T : \mathbb{C}^{d\times d} \to \mathbb{C}^{d\times d}$ a quantum channel, i.e. a completely positive trace-preserving map. Then, the following are equivalent*

(i) $T(A) = UT(U^*AU)U^*$ for all unitaries $U \in G$ and $A \in \mathcal{B}(\mathbb{C}^d)$.

(ii) There exists a pair of real numbers $(\alpha, \beta) \in \mathbb{R}^2$ contained in the triangle with vertices $(\alpha, \beta) \in \{(0, \frac{-1}{d-1}), (\frac{d}{d-1}, 0), (0, 1)\}$ so that

$$
T_G(A) = \alpha\, \mathrm{Tr}[A]\, \mathbb{1}_d + \beta A + (1-\alpha-\beta)\, \mathrm{diag}(A).
$$

Proof. Let us denote

$$
T_G(A) := \int_G UT(U^*AU)U^* dU,
$$

where dU is the Haar measure of G. Then condition (i) says that $T_G = T$. Consider the the Jamiołkowski state τ of $T_G = T$. We show first that (i) implies (ii). Using the identity $\mathbb{1} \otimes A\,|\Omega\rangle = A^T \otimes \mathbb{1}\,|\Omega\rangle$ we can write the Jamiolkowski state of T_G as

$$
\tau = \int_G (U \otimes \bar{U})(T \otimes \mathrm{id})(|\Omega\rangle\langle\Omega|)(U \otimes \bar{U})^* dU.
$$

Thus $[\tau, U \otimes \bar{U}] = 0$ for all $U \in G$. The latter is equivalent to $[(\mathrm{id} \otimes t)(\tau), U \otimes U] = 0$ for all $U \in G(f)$, where t is here the transpose map, $t(A) = A^T$. Note that τ does not have the full $U \otimes U$ symmetry which the Werner state possess. However, we can proceed in the same fashion and find the general structure of the matrix τ: the action of diagonal unitaries and the permutation of the basis elements imply that the only possible non-zero matrix elements are $\langle kk|(\mathrm{id} \otimes t)(\tau)|kk\rangle$, $\langle kl|(\mathrm{id} \otimes t)(\tau)|kl\rangle$ and $\langle kl|(\mathrm{id} \otimes t)(\tau)|lk\rangle$ for $k \neq l \in \{1, \ldots, d\}$. Moreover, since these elements are permutationally invariant they are constant and thus

$$
\tau = \alpha \sum_{k \neq l}^{d} |kl\rangle\langle kl| + \beta \sum_{k \neq l}^{d} |kk\rangle\langle ll| + \gamma \sum_{k=1}^{d} |kk\rangle\langle kk|,
$$

where $\alpha, \beta, \gamma \in \mathbb{C}$. Redefining the variables $d\alpha \to \alpha, d\beta \to \beta$ and $d\gamma \to \gamma$, followed by the invertible transformation $(\alpha, \beta, \gamma - \alpha - \beta) \to (\alpha, \beta, \alpha)$, the corresponding linear map $T_G = T$ of τ can be parametrized as

$$T(A) = \alpha \operatorname{Tr}[A] \mathbb{1}_d + \beta A + \gamma \operatorname{diag}(A). \tag{2.3}$$

If we demand that T is a quantum channel, then $\tau \geq 0$ and in particular T is Hermicity preserving (see Prop. 2.1 in [8]). This implies together with the trace-preserving property that $\alpha, \beta \in \mathbb{R}$ and $\gamma = 1 - \alpha - \beta$. In order that τ is positive, we need that the eigenvalues of τ are positive. For that matters, we write $T(A) = \alpha \operatorname{Tr}[A] \mathbb{1}_d + Y \circ A$, where $\circ$ is the Hadamard product and

$$Y := \begin{pmatrix} 1-\alpha & \beta & \dots & \beta \\ \beta & 1-\alpha & \dots & \beta \\ \vdots & \vdots & \ddots & \vdots \\ \beta & \beta & \dots & 1-\alpha \end{pmatrix} = \beta J_d + (1 - \alpha - \beta) \mathbb{1}_d .$$

The matrix Y can be diagonalized to obtain that $Spec(Y) = \{1 - \alpha + \beta(d-1), 1 - \alpha - \beta\}$. Thus, adding the completely positive requirement imposes that

$$\begin{aligned} \alpha &\geq 0, \\ d\beta + \frac{d}{d-1} &\geq \alpha, \\ \frac{d}{d-1}(1-\beta) &\geq \alpha. \end{aligned}$$

meaning that the values of $(\alpha, \beta) \in \mathbb{R}^2$ which make τ positive are contained in the triangle with vertices $(\alpha, \beta) \in \{(0, \frac{-1}{d-1}), (\frac{d}{d-1}, 0), (0, 1)\}$. Finally, *(ii)* implies *(i)* as the trace, the identity and $\operatorname{diag}(\cdot)$ are invariant under such conjugation of unitaries. □

2.4 Stability

In this work we study the stability of symmetries from a mathematical analysis point of view. At the risk of oversimplification one could say that the general situation is the following: one slightly perturbs the hypothesis of a theorem in terms of an additive error ε and studies whether the conclusion of the theorem is still close to the ideal situation or not. The deviations from the ideal situation, $\varepsilon = 0$, are quantified by a notion of distance such as a norm and the inequalities obtained depend on ε. This is where the realm of analysis enters here. The attention to this sort of stability problems was sparked by the work of S. Ulam and D. Hyers in 1940 [23, 24] and later by Th. M. Rassias in 1978 [25]. They studied the stability of the most important functional equation, namely the functional equation defining an additive map (also known as Cauchy functional equation).

Theorem 2.4.1 (Hyers–Ulam–Rassias). *Let $f : X \to Y$ be a function between Banach spaces. If f satisfies the functional inequality*

$$\|f(x_1 + x_2) - f(x_1) - f(x_2)\|_Y \leq \varepsilon(\|x_1\|_X^p + \|x_2\|_X^p),$$

for some $\varepsilon \geq 0$, $0 \leq p < 1$ and for all $x_1, x_2 \in X$. Then there exists a unique additive function $A : X \to Y$ such that

$$\|f(x) - A(x)\|_Y \leq \frac{2\varepsilon}{2 - 2^p} \|x\|_X^p \qquad \textit{for all } x \in X.$$

If in addition $f(tx)$ is continuous in t for each fixed $x \in X$, then A is a linear function.

Their work has motivated a lot of research in non-linear analysis [26, 27] that nowadays we referred to this notion of stability on functional equations as the *Hyers–Ulam–Rassias (HUR) stability.* Stability problems do not always give a positive answer. The additive functional equation is not stable in the *HUR* sense for $p = 1$ for general Banach spaces. The case $p = 1$ is called the *singular case.* For $d-$dimensional Banach spaces X, Y, the additive functional equation is stable for $p = 1$, but the bound necessarily depends on the dimension d. This "(in-)stability" phenomena appears in other problems as well. For instance, almost commuting matrices need not be nearly commuting if the considered set is $\mathcal{B}(\mathcal{H})$ [28] or if they are unitaries [29]. The stability bounds cannot be independent of the dimension d. On the other hand, almost commuting matrices are always nearly commuting if they are Hermitian [30] (see [31] for an explicit bound). We will encounter a similar situation in the main article of this thesis (see Section A.1).

Our research question deals with the stability of symmetries in quantum mechanics in two important scenarios. On one hand, we study the *linear* stability of Wigner's theorem (see Section 2.2) which is the fundamental result on how we represent any symmetry transformation in quantum theory [4]. More concretely, for an arbitrary Hilbert space $\mathcal{H}$, we study if an *almost-symmetry*, i.e. a mapping $f : \mathbb{P}(\mathcal{H}) \to \mathbb{P}(\mathcal{H})$ which satisfies for some $\varepsilon \geq 0$

$$|\mathrm{Tr}\, f(X) f(Y) - \mathrm{Tr}\, XY| \leq \varepsilon \qquad \text{for all } X, Y \in \mathbb{P}(\mathcal{H}),$$

can be always approximated by a linear map. Our results improve the recent results on the general stability of Wigner's theorem [32]; we clarify the role of the dimension of $\mathcal{H}$ for the stability and obtain a new result for the infinite-dimensional case (see Section A.1). Particularly noteworthy in the proof of [32] is the use of the stability of Herstein's theorem [33]. The latter non-trivial result implies the isometry-stability of Wigner's theorem.

On the other hand, we study the stability of an ubiquitous symmetric state in continous variable quantum information, namely the *Gaussian bosonic state* (GBS). These states are central to continuous variable quantum information just as the normal distribution is to classical probability theory. We investigate an operational characterization of a general GBS. This characterization describes an inherent symmetry that is only shared by Gaussian Bosonic states (see Section 3). The main contribution here is that we show that this characterization of GBS is stable if the symmetry assumption is slightly weakened (see Section B.1).

An important breakthrough in the stability of linear maps is present in the work of N. J. Kalton. In Ref. [34] (Theorem 2.2) Kalton studies the stability of the additive map between Euclidean spaces for $p = 1$ and provides a sharp bound on the dimension dependence. Moreover, he makes a link between the stability of linear maps and the geometry of the Banach spaces on which the almost-linear map is defined. We study in more detail this last novel idea in Section A.2 and use it to obtain an improvement in the dimension dependence of the linear stability of Wigner's theorem.

An arbitrary isometry $f : X \to Y$ between two Banach spaces is a map such that

$$\|f(x) - f(y)\| = \|x - y\| \qquad \text{for all } x, y \in X.$$

A similar result to Wigner's theorem is the classical theorem of Ulam and Mazur [35] which states that any surjective isometry between two real Banach spaces is affine. However, there are some major differences between these two theorems. A Wigner symmetry transformation is not necessarily assumed to be surjective and is initially defined on the set of pure states, which is a non-linear space. Note that from Eq. (2.1) and Wigner's symmetry condition, Eq. (2.2), the distance between *pure states* is preserved. So in principle, the linearity of Wigner's theorem is not directly implied by Ulam-Mazur's theorem as the latter requires that the isometry is global. Furthermore, Wigner's theorem states that the linear isometry has a particular form on $\mathcal{B}(\mathcal{H})$ (in the quantum channel language, it has Kraus rank equal to one). Instead of isometries, one could consider $\varepsilon-$isometries, that is mappings which almost preserve distance. A description of such $\varepsilon-$isometries goes beyond the scope of this introduction. We refer the reader to chapters 14 and 15 of [36] and references therein.

2.5 (Linear) Preservers problems

Although this thesis is mainly concerned with the non-trivial question of the linear stability of Wigner's theorem, let us briefly mention where this result might find applications. A *preserver problem* considers the characterization of maps that preserve certain functional, subset, or invariant in a matrix space or operator algebra. Examples of such problems include the characterization of maps that preserve the spectrum of a matrix or the commutativity properties of a subset of matrices. For a survey on this subfield of mathematics we refer to [37, 38]. It turns out that most of the preserver problems in matrix theory give as a result that the only map that preserves the required property is of the form $f(X) = UXU^*$ or $f(X) = UX^TU^*$ where U is a unitary matrix. This occurs because in most of the cases the given characterization can be reduced to the characterization of a map –not necessarily linear– which preserves the Hilbert-Schmidt scalar product of all rank one projections. That is, they can be reduced to the Wigner theorem. This does not only happen for linear preserver problems where the studied map is initially assumed to be linear. In quantum information theory, the mappings which preserve the von Neumann entropy or the relative entropy are again only the Wigner symmetry transformations [39]. The stability of Wigner's theorem can be then applied to questions on almost preserving problems; for instance, on how near is a map that almost preserves entropy to a Wigner symmetry transformation and what are the optimal bounds for such approximation?

3 Bosonic quantum systems

In this section we introduce the basic elements of bosonic quantum systems which will be needed in Section B.1. There are many good books on this topic. In particular, the work of Holevo [6] establishes the common mathematical and physics background which is still used widely nowadays. We also urged the reader to take a look at chapter 16 and 17 of Ref. [40] for a better understanding of the mathematical subtlety that arises when working with bosonic quantum systems. Here we focus in the description of the ubiquitous Gaussian bosonic states from a symmetry point of view.

A *bosonic quantum system* is a quantum system whose canonical observables Q and P satisfy the commutation relation

$$[Q, P] = i, \tag{3.1}$$

where $i = \sqrt{-1}$. It turns out that any pair of operators (Q, P) that satisfy Eq. (3.1) cannot be both bounded neither representable in a finite dimensional Hilbert space $\mathcal{H}$. This result is attributed to Wintner and Wielandt (see Lemma 17.1-7 and Proposition 17.1-4 in [40]):

Theorem 3.0.1 (Wintner-Wielandt)**.** *Let Q and P be two bounded operators on a real or complex Hilbert space such that the commutation relation $[Q, P] = z\,\mathbb{1}$ is valid for some $z \in \mathbb{R}$ or $z \in \mathbb{C}$, respectively. Then it follows that $z = 0$.*

Therefore when working with bosonic quantum systems we necessarily work in infinite-dimensional Hilbert spaces and in particular with unbounded operators. For a system with multiple degrees of freedom (modes) we denote by

$$R := (Q_1, P_1, \dots, Q_n, P_n),$$

the vector of canonical operators. We represent the Hilbert space of the whole bosonic quantum system as the *Fock space* $\mathcal{H} = \bigotimes_{k=1}^{n} \mathcal{L}^2(\mathbb{R})$. The entries $R_k, k = 1, \dots, 2n$ act on the $k-$tensor factor of $\bigotimes_{k=1}^{n} \mathcal{L}^2(\mathbb{R})$. Then Eq. (3.1) is extended to

$$[R_k, R_l] = i\sigma_{kl}, \tag{3.2}$$

where σ_{ij} are the entries of the symplectic matrix

$$\sigma = \bigoplus_{i=1}^{n} \omega \quad \text{with} \quad \omega := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Equation (3.2) is known as the *canonical commutation relation (CCR)*. There exists however a way to encode the CCR algebra in a family of bounded operators, namely introducing the so-called *Weyl operators* $W_\xi := e^{i\xi\cdot\sigma R}$ where $\xi \in \mathbb{R}^{2n}$. The Weyl operators satisfy the relation

$$W_\xi W_\eta = e^{-\frac{i}{2}\xi\cdot\sigma\eta}\, W_{\xi+\eta}, \qquad \xi, \eta \in \mathbb{R}^{2n}. \tag{3.3}$$

3.1 Phase space description

It will be useful to translate the operator description of an infinite-dimensional quantum system to a description in terms of complex-valued functions in phase space. For $f \in \mathcal{L}^1(\mathbb{R}^{2n})$, the *Weyl Transform* $\hat{f}$ of f is defined as

$$\hat{f} := \frac{1}{(2\pi)^n} \int f(\xi) W_{-\xi} d\xi, \tag{3.4}$$

where the integral is well-defined in the weak sense. The Weyl transform establishes an isometry (up to a constant factor) between square-integrable functions in phase space and Hilbert-Schmidt operators. This is captured by the *quantum Parseval theorem* [6]:

Theorem 3.1.1 (Quantum Parseval relation). *Let $\{W_\xi\}$ be a strongly continuous and irreducible Weyl systems acting on the Hilbert space $\mathcal{H}$ with respective phase space $X \simeq \mathbb{R}^{2n} \ni \xi$. Then $A \mapsto \mathrm{Tr}[W_\xi A]$ extends uniquely to an isometric map from the Hilbert space of Hilbert-Schmidt class operators on $\mathcal{H}$ onto $\mathcal{L}^2(X)$, such that*

$$\mathrm{Tr}\, A^* B = \frac{1}{(2\pi)^n} \int \overline{\mathrm{Tr}[W_\xi A]}\, \mathrm{Tr}[W_\xi B] d\xi.$$

It can be shown from Theorem 3.1.1 (corollary 5.3.5 in Ref. [6]) that any Hilbert-Schmidt operator A in $\mathcal{H}$ is the Weyl transform of $\mathrm{Tr}\, W_\xi A$, i.e.

$$A = \frac{1}{(2\pi)^n} \int (\mathrm{Tr}\, W_\xi A) W_{-\xi}\, d\xi.$$

In fact, the map $A \mapsto \mathrm{Tr}\, W_\xi A$ is the *inverse* of the Weyl Transform and so it is one-to-one map. In these terms, the *characteristic function* $\chi : \mathbb{R}^{2n} \to \mathbb{C}$ of a quantum state ρ is defined as the inverse Weyl transform of ρ, i.e.

$$\chi(\xi) := \mathrm{Tr}[W_\xi \rho].$$

We write χ_A to emphasize that this is the characteristic function of the operator A. The regularity properties of ρ manifest now on the integrability and differentiability conditions of χ_ρ.

Remark on notation: Unfortunately, the mathematical and physics community is not united on the name for the map in Eq. (3.4). We follow the terminology of Ref. [6] and use the name of Weyl transform for this map. We caution that this differs from the terminology of Ref. [41].

The Hilbert-Schmidt norm (HS-norm), $\|A\|_2 := \sqrt{\mathrm{Tr}\, A^* A}$, proves sometimes to be a good choice of norm not only because of Theorem 3.1.1, but rather because of its operational meaning [42]. In a model of equality testing in a two-party scenario in which the preparer and tester do not share a reference frame, the HS-norm appears as the right figure of merit in which the optimal probability of success is expressed. Moreover, the HS-norm is the right measure to distinguish between two equally prepared states if the measurements to be performed are random.

3.1.1 Characteristic functions

A classical characteristic function $\phi : \mathbb{R}^{2n} \to \mathbb{C}$ is the Fourier transform of a classical probability distribution. The necessary and sufficient conditions for a function ϕ to be a valid classical characteristic function are given by the Bochner-Khinchin theorem:

Theorem 3.1.2 (Bochner-Khinchin)**.** *For $\phi : \mathbb{R}^{2n} \to \mathbb{C}$ to be a classical characteristic function, i.e. the Fourier transform of a classical probability distribution, the following conditions are necessary and sufficient:*

1. *$\phi(0) = 1$ and ϕ is continuous at $\xi = 0$,*

2. *ϕ is positive definite, i.e. for any $m \in \mathbb{N}$, any set $\{\xi_1, \xi_2, \dots, \xi_m\}$ of vectors in $\mathbb{R}^{2n}$, and any set $\{c_1, c_2, \dots, c_m\}$ of complex numbers*

$$\sum_{k,l=1}^{m} c_k \overline{c_l}\ \phi(\xi_k - \xi_l) \geq 0. \tag{3.5}$$

Now in order that $\chi : \mathbb{R}^{2n} \to \mathbb{C}$ is a valid characteristic function of a quantum state the second condition of Theorem 3.1.2 has to be changed, namely Eq.(3.5) has to be "twisted" so that it corresponds to a positive definite operator ρ. This new condition is called σ-positive definite: for any $m \in \mathbb{N}$, any set $\{\xi_1, \xi_2, \dots, \xi_m\}$ of vectors in $\mathbb{R}^{2n}$, and any set $\{c_1, c_2, \dots, c_m\}$ of complex numbers

$$\sum_{k,l=1}^{m} c_k \overline{c_l}\ \chi(\xi_k - \xi_l)\ e^{\frac{i}{2}\xi_k \cdot \sigma \xi_l} \geq 0.$$

As in the classical case $\chi(0) = 1$ and χ is continuous at $\xi = 0$, which accounts for $\operatorname{Tr}\rho = 1$ (see Section 5.4 in [6]).

We can now list some useful properties of quantum characteristic functions which are easily verified from the definitions.

Lemma 3.1.3. *Let χ_{ρ_1} and χ_{ρ_2} be the quantum characteristic functions of ρ_1 and ρ_2, respectively. Then*

(i) *for $\lambda_1, \lambda_2 \geq 0$, $\lambda_1 + \lambda_2 = 1$ the function $\chi_\rho := \lambda_1 \chi_1 + \lambda_2 \chi_2$ is the quantum characteristic function of the state $\rho = \lambda_1 \rho_1 + \rho_2 \rho_2$.*

(ii) *The characteristic function of $\rho_1 \otimes \rho_2$ is $\chi_{\rho_1 \otimes \rho_2}(\xi_1, \xi_2) = \chi_{\rho_1}(\xi_1)\chi_{\rho_2}(\xi_2)$ where $\xi_1, \xi_2 \in \mathbb{R}^{2n}$.*

(iii) *$\overline{\chi(\xi)} = \chi(-\xi)$ and $|\chi(\xi)|^2$ are valid quantum characteristic functions.*

(iv) *If $\rho_{AB} \in \mathcal{B}(\mathcal{H}_{AB})$ is a density operator of a bipartite system AB with quantum characteristic function $\chi_{AB}(\xi, \eta)$ where $\xi, \eta \in \mathbb{R}^{2n}$, then the partial trace $\rho_A = \operatorname{Tr}_B \rho_{AB}$ has quantum characteristic function $\chi_A(\xi) := \chi_{AB}(\xi, 0)$. Likewise $\rho_B = \operatorname{Tr}_A \rho_{AB}$ has quantum characteristic function $\chi_B(\eta) := \chi_{AB}(0, \eta)$.*

(v) *The action of displacement is at the level of characteristic functions $\chi_{W_\xi \rho W_\xi^*}(\eta) = e^{i\eta \cdot \sigma \xi} \chi_\rho(\eta)$.*

Characteristic functions are useful because if their derivatives exist, we can obtain the statistical moments of the respective distribution. We will talk more about quantum states whose all moments exist in subsection 3.3. Let us for now assume that we are working with such states. Then the gradient of the characteristic function at zero gives the first moments of ρ, that is $\nabla\chi(0) = i\sigma d$ where d is the vector with entries $d_k := \mathrm{Tr}[R_k\rho]$. The vector d is known as the *displacement vector*. The *covariance matrix* is defined as the matrix Γ with entries

$$\Gamma_{kl} := \mathrm{Tr}[\rho\{R_k - d_k\,\mathbb{1}, R_l - d_l\,\mathbb{1}\}_+],$$

where $\{A, B\}_+ := AB+BA$ is the anticommutator. The second derivatives of the characteristic function are related to Γ by

$$(\sigma\Gamma\sigma^T)_{kl} = -2\left(\frac{\partial^2\chi(0)}{\partial\xi_k\partial\xi_l} - d_k d_l\right).$$

Every bosonic quantum state satisfies $\Gamma \geq i\sigma$, which is just a coordinate-independent way to express Heisenberg's uncertainty principle [43]. Finally, we define *Wigner's phase space distribution function* of ρ as

$$\mathcal{W}_\rho(\xi) := \frac{1}{(2\pi)^{2n}}\int e^{i\xi\cdot\sigma\eta}\chi_\rho(\eta)d\eta,$$

that is, $\mathcal{W}$ is the inverse symplectic Fourier transform of χ_ρ. It is generally not a probability distribution since it can take negative values. However, the marginals of $\mathcal{W}$ are genuine probability distributions of the individual canonical operators R_k. The Wigner distributions of a quantum state of light can be readily reconstructed by means of Tomography (see Chapter 5 in Ref. [44]).

3.2 Gaussian states

A *Gaussian state* is a quantum state ρ whose characteristic function is Gaussian:

$$\chi_\rho(\xi) = \exp[-\frac{\xi\cdot\Gamma\xi}{4} + i\xi\cdot d].$$

This means that the statistic of Gaussian states is fully determined by its first and second moments. Examples of Gaussian states include thermal states and the so-called squeezed states. The importance of Gaussian bosonic states lies in the fact that there is a non-commutative central limit theorem [45, 46] for bosonic systems where these states play the central role; just like the normal distribution does in classical probability. This quantum central limit theorem implies a number of important extremal properties [46, 47]. A *symplectic transformation* is a real matrix S such that $S\sigma S^T = \sigma$. We denote the group of $2n\times 2n$ symplectic transformations by $Sp(2n, \mathbb{R})$. These transformations preserve the CCR relations, Eq.(3.2). A Unitary evolution $\rho \mapsto U_S\rho U_S^*$ where $U_S = \exp[i\sum_{k,l} A_{k,l}\{R_k, R_l\}_+]$, $A^T = A \in \mathbb{R}^{2n\times 2n}$, can be represented by a symplectic transformation S. This is known as the *Metaplectic representation* and the map $U_S \mapsto S$ is a two-to-one homomorphism (see [43] and references therein). In quantum optics the action of beam-splitters, phase-shifters and "squeezers" can be modeled very well by such unitary evolutions coming from quadractic Hamiltonians. We will be particularly interested in the unitary evolution corresponding to a non-trivial two n-mode *Beam splitter* operation

$$S_\theta = \begin{pmatrix} \cos\theta\,\mathbb{1}_{2n} & -\sin\theta\,\mathbb{1}_{2n} \\ \sin\theta\,\mathbb{1}_{2n} & \cos\theta\,\mathbb{1}_{2n} \end{pmatrix}, \qquad \theta \neq m\pi/2, \quad m = 0, 1, 2, \ldots. \tag{3.6}$$

At the level of characteristic functions, the dynamical evolution is represented by

$$\chi_{(U_S \rho U_S^*)}(\xi) = \chi_\rho(S^T \xi).$$

When ρ is a Gaussian state, we see that U_S preserves the Gaussian character of ρ. This is an example of a *Gaussian Channel*. For a general description of general Gaussian channels see [43].

There exist different characterizations of Gaussian states: (i) in terms of its extremal properties (Gaussian states maximize the von Neumann entropy among all the states with the same second moments), (ii) they are completely characterized by the first and second moments (Wick's theorem) and (iii) in terms of a *symmetry*:

Theorem 3.2.1 (Characterization of Gaussian states)**.** *Let $\rho \in \mathcal{B}(\mathcal{H})$ be a bosonic quantum state and let $T_\theta \in \mathcal{B}(\mathcal{H}\otimes\mathcal{H})$ be the unitary operation corresponding to a non-trivial beam splitter operation, Eq.*(3.6)*. Then, ρ is a Gaussian state if and only if $\rho \otimes \rho$ is a fixed point of T_θ, i.e. $T_\theta(\rho \otimes \rho) = \rho \otimes \rho$.*

Of course one implication of Theorem 3.2.1 is trivial since $S_\theta(\Gamma \oplus \Gamma)S_\theta^T = (\Gamma \oplus \Gamma)$. However, *a priori* there is no reason to believe that there does not exist a non-Gaussian state which is also a fixed point of T_θ. This symmetry characterization of Gaussian states and its stability properties will be the main topic of the second core article of this thesis. Theorem 3.2.1 is a consequence of a quantum version of the *Darmois-Skitovich theorem* which will be proven in Section B.1.

3.3 Schwartz operators

The set of quantum states whose all moments of all orders and combinations in Q and P exist is the set of *Schwartz density operators* [41]. One way to characterize this set is in terms of the characteristic function: a quantum state ρ is *Schwartz* if and only if the characteristic function χ_ρ is a Schwartz function. We denote the set of Schwartz functions by $\mathscr{S}(\mathbb{R}^n)$ and the set of Schwartz density operators by $\mathscr{S}(\mathcal{H})$. Since the Fourier transform is a linear bicontinuous bijection from $\mathscr{S}(\mathbb{R}^n)$ onto $\mathscr{S}(\mathbb{R}^n)$, we can say that $\rho \in \mathscr{S}(\mathcal{H})$ if and only if $\mathcal{W}_\rho \in \mathscr{S}(\mathbb{R}^n)$.

Examples of Schwartz quantum states include the set of Gaussian states and the (generally non-Gaussian) set of Fock states. The latter set is the set of eigenstates of the number operator $a^*a|n\rangle = n|n\rangle$, $n \in \mathbb{N} \cup \{0\}$, where $a := (Q + iP)/\sqrt{2}$, $a^* := (Q - iP)/\sqrt{2}$ are the so-called "creation" and "annihilation" operators, respectively. For a single degree of freedom the characteristic function a $n-$Fock state is

$$\chi_{|n\rangle\langle n|}(\xi) = e^{-\frac{\|\xi\|_2^2}{4}} L_n(\|\xi\|_2^2/2),$$

where $\xi \in \mathbb{R}^2$ and L_n is the n Laguerre polynomial. After the set of Gaussian states, the Schwartz density operators correspond to the most regular set of states. With this set, we can manipulate unbounded operators with greater freedom: the cyclicity under the trace is allowed for a pair of Schwartz operators; the differentiation and integration of the quantum characteristic function for any order exist and is finite; we can write the derivatives of the characteristic function in terms of a trace so that they indeed are directly related with the moments of ρ.

For A, a bounded operator in $\mathcal{L}^2(\mathbb{R}^n)$, the *Kernel* of A is defined as the function $K \in \mathcal{L}^2(\mathbb{R}^n \times \mathbb{R}^n)$ such that

$$\langle \psi | A\varphi \rangle = \int_{\mathbb{R}^{2n}} \overline{\psi}(q) K(q,q') \varphi(q') dq dq' \qquad \text{for all } \psi, \varphi \in \mathcal{L}^2(\mathbb{R}^n).$$

Any Hilbert-Schmidt operator A has a unique square-integrable Kernel. *Schwartz operators* are operators whose inverse Weyl transform is a Schwartz function. We denote by $\mathfrak{S}(\mathcal{H})$ the set of general Schwartz operators. The following theorem summarizes the main properties of Schwartz operators [41].

Theorem 3.3.1. *Let $\mathcal{H} = \mathcal{L}^2(\mathbb{R}^{2n})$ and $A \in \mathcal{B}(\mathcal{H})$. Then*

(i) (Range Theorem) $A \in \mathfrak{S}(\mathcal{H})$ *if and only if* $\mathrm{Ran}(A)$ *and* $\mathrm{Ran}(A^*)$ *are Schwartz functions on $\mathbb{R}^{2n}$.*

(ii) $A \in \mathfrak{S}(\mathcal{H})$ if and only if the Kernel of A is a Schwartz function.

(iii) Let f be a polynomial function on the entries of the vector $R = (Q_1, P_1, \ldots, Q_n, P_n)$ and W_ξ the Weyl operator. If $A \in \mathfrak{S}(\mathcal{H})$, then $\mathrm{Tr}[f(R)A] = \mathrm{Tr}[Af(R)]$. Moreover, $f(R)A \in \mathfrak{S}(\mathcal{H})$ and $\mathrm{Tr}[W_\xi f(R)A] = \mathrm{Tr}[f(R)AW_\xi] = \mathrm{Tr}[TW_\xi f(R)]$.

(iv) If $A \in \mathfrak{S}(\mathcal{H})$, then A is trace-class.

(v) If $A \in \mathfrak{S}(\mathcal{H})$, then $|A| \in \mathfrak{S}(\mathcal{H})$.

(vi) If $0 < A \in \mathfrak{S}(\mathcal{H})$, then $\sqrt{A} \in \mathfrak{S}(\mathcal{H})$.

(vii) If $A \in \mathfrak{S}(\mathcal{H})$, then $A^* \in \mathfrak{S}(\mathcal{H})$.

Not every quantum state in $\mathcal{H}$ is so regular like a Schwartz operator as the next example shows.

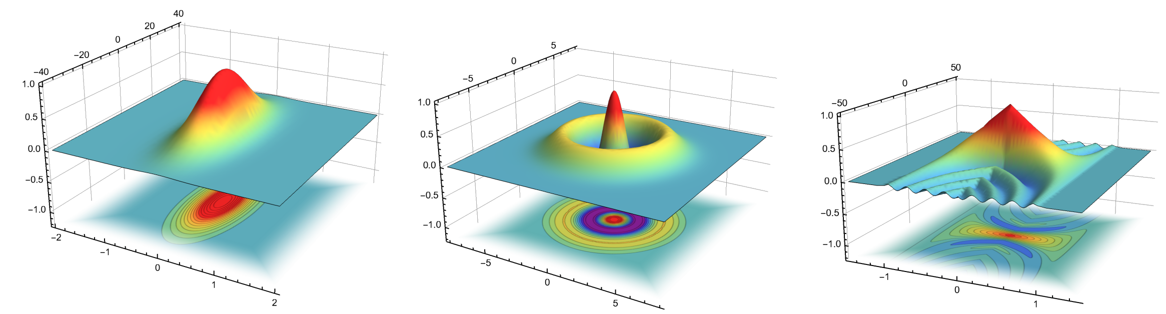


Figure 3.1: The left and center figures are the characteristic functions of a (squeezed) Gaussian state and 2-Fock state, respectively. The right figure is the characteristic function of the unit rectangle of example 4.3.2 with $l = 1$. It is identically zero outside the strip $-l \leq q \leq l$ and not partially differentiable with respect to q at zero.

Example 3.3.2 (Non-Schwartz quantum state)**.** *Let $l > 0$. The state $\psi \in \mathcal{H} = \mathcal{L}^2(\mathbb{R})$ whose position representation is*

$$\psi(x) = \begin{cases} \frac{1}{\sqrt{l}} & |x| \leq \frac{l}{2}, \\ 0 & |x| > \frac{l}{2}, \end{cases}$$

does not correspond to a Schwartz density operator. Its Fourier transform, i.e. its momentum representation, is $\hat{\psi}(p) = \sqrt{l}\,\mathrm{sinc}(pl/2)$ *where* $\mathrm{sinc}(p) := (\sin p)/p$*. While the moments with respect to* Q *are finite, e.g.* $\mathrm{Tr}\, Q|\psi\rangle\langle\psi| = \int_{-\infty}^{\infty} |\psi(x)|^2 x dx = 0$, $\mathrm{Tr}\, Q^2|\psi\rangle\langle\psi| = \int_{-\infty}^{\infty} |\psi(x)|^2 x^2 dx = l^3/12$*, the same does not happen with* P *since* $\int_{-\infty}^{\infty} |\hat{\psi}(p)|^2 p^2 dx = \infty$*. In fact,* $\psi(x)$ *is not classically differentiable at* $\pm l/2$ *and therefore* $\mathrm{Tr}\, P|\psi\rangle\langle\psi|$ *has to be understood in a distributional sense. Moreover, the operator* P^2 *maps* ψ *out of the Hilbert space of square integrable functions as* $\|P\psi\|_2 = \infty$*. These sort of pathologies can be seen as well in the characteristic function of* $|\psi\rangle\langle\psi|$*. Let* $\Lambda(y) = 1 - |y|$ *for* $|y| \le 1$ *and zero otherwise, then*

$$\chi_{|\psi\rangle\langle\psi|}(q,p) = \frac{\sin\left[\frac{pl}{2}\Lambda(q/l)\right]}{\frac{pl}{2}}.$$

The characteristic function of $|\psi\rangle\langle\psi|$ *is not partially differentiable with respect to* q *at zero since* Λ *is not differentiable at this point. See Figure 3.1.*

3.4 Quantum Meyers-Serrin theorem (optional)

The following section is part of a joint unpublished work with Michael M. Wolf and Michael Keyl. The results presented here find an application in the contributed article B.1, but the reader can skip it as it is not needed to understand B.1. Here we study how Schwartz density operators can approximate arbitrarily well quantum states with symmetric moments. This makes the set of Schwartz density operators more useful.

We define a *symmetric moment* of $\rho \in \mathcal{B}(\mathcal{H}^{\otimes n})$ to be the quantity

$$\mathrm{Tr}\, Q^\alpha P^\beta \rho P^\beta Q^\alpha,$$

where

$$Q^\alpha = Q_1^{\alpha_1} \cdots Q_n^{\alpha_n}, \qquad P^\beta = P_1^{\beta_1} \cdots P_n^{\beta_n}.$$

for some

$$\alpha, \beta \in I_n := \{\alpha = (\alpha_1, \dots, \alpha_n) | \alpha_i \in \mathbb{N} \cup \{0\} \text{ for all } i = 1, \dots, n\}.$$

Recall that Q_k and P_k act on the $k-$tensor factor of $\mathcal{H}^{\otimes n}$. We define $|\alpha| := \sum_{i=1}^n \alpha_i$ and $|\beta| := \sum_{i=1}^n \beta_i$.

Definition 3.4.1 (Finite-moment state)**.** *Let* $r \in \mathbb{N}$*. We say that a density operator* $\rho \in \mathcal{B}(\mathcal{H}^{\otimes n})$ *has up to* $2r-$*th symmetric moments if*

$$\|\rho\|_{r,2} := \sum_{\substack{\alpha,\beta \in I_n \\ |\alpha|+|\beta| \le r}} \left\| Q^\alpha P^\beta \sqrt{\rho} \right\|_2 < \infty.$$

The space of density operators on $\mathcal{H}$ *which has up to finite 2r-th symmetric moments is denoted by* $\mathscr{S}_r^2(\mathcal{H})$*.*

For $\rho \in \mathscr{S}_r^2(\mathcal{H})$ the expectation values of the canonical operators are finite up to order $2r$, namely for $|\alpha| + |\beta| \le r$

$$\mathrm{Tr}[Q^\alpha P^\beta \rho] \le \left\| Q^\alpha P^\beta \rho \right\|_1 \le \left\| Q^\alpha P^\beta \sqrt{\rho} \right\|_2 = (\mathrm{Tr}\, Q^\alpha P^\beta \rho P^\beta Q^\alpha)^{1/2} < \infty.$$

The states ρ with up to $2r-$finite moments can be characterized in terms of the twisted derivative of the inverse Weyl transform of $\sqrt{\rho}$. The *twisted (or symplectic) derivatives* L_z and $\bar{L}_z$ of χ_A are defined [48] as

$$\begin{aligned}
L_z\chi_A(\xi) &:= -i\frac{d}{dt}\left(e^{itz\cdot\sigma\xi/2}\chi_A(\xi+tz)\right)\Big|_{t=0}, \\
&= \left(\frac{1}{2}z\cdot\sigma\xi - i\frac{\partial}{\partial z}\right)\chi_A(\xi). \\
\bar{L}_z\chi_A(\xi) &:= -i\frac{d}{dt}\left(e^{-itz\cdot\sigma\xi/2}\chi_A(\xi+tz)\right)\Big|_{t=0}, \\
&= \left(-\frac{1}{2}z\cdot\sigma\xi - i\frac{\partial}{\partial z}\right)\chi_A(\xi).
\end{aligned}$$

In order to avoid to many subindices we write sometimes L instead of L_z for a symplectic derivative in an arbitrary direction. We use L_z when we want to emphasize the direction. Of course one has to specify in which norm the derivative is taken; this will depend on the operator (equivalently, the inverse Weyl transform) as we will see.

Let us write

$$L^{\alpha\beta} := L_{z_1}^{\alpha_1}\cdots L_{z_n}^{\alpha_n}L_{z_{n+1}}^{\beta_1}\cdots L_{z_{2n}}^{\beta_1}, \tag{3.7}$$

where $\{z_k\}_{k=1}^{2n}$ is a basis such that for $i=1,\dots n$, $z_i\cdot\sigma R$ is equal to Q_i and for $i=n+1,\dots 2n$, $z_i\cdot\sigma R$ is equal to P_i.

Definition 3.4.2. *Let ρ be a density operator and denote by $\chi_{\sqrt{\rho}}$ the inverse Weyl transform of the square root of ρ. Let $r\in\mathbb{N}$ and $L^{\alpha\beta}$ be the composed twisted derivative defined as in Eq. (3.7). We say that a density operator ρ on a Hilbert space $\mathcal{H}$ has up to $2r-$th symmetric moments if and only if*

$$\sum_{\substack{\alpha,\beta\in I_n\\|\alpha|+|\beta|\le r}}\left\|L^{\alpha\beta}\chi_{\sqrt{\rho}}\right\|_2<\infty.$$

This definition is equivalent to definition 3.4.1 due to Lemma 3.4.3 (below) and the quantum Parseval theorem. Note also that $L\chi\in\mathcal{L}^2(\mathbb{R}^{2n})$ if and only if the twisted derivative of the (inverse) symplectic-Fourier transform of χ is in $\mathcal{L}^2(\mathbb{R}^{2n})$.

Lemma 3.4.3. *Let T be a Hilbert-Schmidt operator in $\mathcal{H}$ and $\chi_T\in\mathcal{L}^2(\mathbb{R}^{2n})$ its corresponding inverse Weyl transform. Let $R_z=z\cdot\sigma R$ where R is the vector of canonical operators and $z\in\mathbb{R}^{2n}$. If R_zT is Hilbert-Schmidt then*

$$L\chi_T(\xi) := -i\frac{d}{dt}\left(e^{itz\cdot\sigma\xi/2}\chi_T(\xi+tz)\right)\Big|_{t=0} = \chi_{R_zT}(\xi).$$

where the derivative is taken in the $\mathcal{L}^2-$ norm.

Note that here the limit of the derivative is taken with respect to the $\mathcal{L}^2$ norm. We omit the proof as it is analogous to the proof of Lemma 14 in Section B.1.

The definition 3.4.2 is analogous to the definition of a classical Sobolev space; in the quantum case the twisted derivative takes the role of the weak derivative. For example, a one-mode quantum state has finite symmetric second moments if and only if

$$\left\|\chi_{\sqrt{\rho}}\right\|_{\mathcal{L}^2}+\left\|L_{z_1}\chi_{\sqrt{\rho}}\right\|_{\mathcal{L}^2}+\left\|L_{z_2}\chi_{\sqrt{\rho}}\right\|_{\mathcal{L}^2}<\infty,$$

where $z_1, z_2 \in \mathbb{R}^2$ are orthogonal vectors. We will show that the space of Schwartz-density operators $\mathscr{S}(\mathcal{H})$ is dense in $\mathscr{S}^2_r(\mathcal{H})$. In other words, the set of states with finite symmetric moments is by definition the completion of $\mathscr{S}(\mathcal{H})$ with respect to $\|\cdot\|_{r,2}$. This implies that when working in any set of states in which the expectation of the canonical operators exist up to certain order, we can from the start work with the well-behaved set of Schwartz density operators.

The following theorem is the main result of this subsection and could be understood as a quantum version of the classical theorem of Meyers and Serrin in Sobolev spaces [49] (see subsection 3.4.1).

Theorem 3.4.4 ($\mathscr{S}(\mathcal{H})$ is dense in $\mathscr{S}^2_r(\mathcal{H})$)**.** *Let $\rho \in \mathscr{S}^2_r(\mathcal{H})$ be a density operator with finite moments up to $2r$ and $\varepsilon > 0$. Then there exists a Schwartz-density operator $\rho_\varepsilon \in \mathscr{S}(\mathcal{H})$ such that for all $\alpha, \beta \in I_n$ with $|\alpha| + |\beta| \leq r$,*

(i) $\left\|Q^\alpha P^\beta \sqrt{\rho} - Q^\alpha P^\beta \sqrt{\rho_\varepsilon}\right\|_2 \longrightarrow 0$ *as* $\varepsilon \longrightarrow 0$

(ii) $\left|\left\|Q^\alpha P^\beta \rho\right\|_1 - \left\|Q^\alpha P^\beta \rho_\varepsilon\right\|_1\right| \leq \left\|Q^\alpha P^\beta \rho - Q^\alpha P^\beta \rho_\varepsilon\right\|_1 \longrightarrow 0$ *as* $\varepsilon \longrightarrow 0$

This theorem tells us that given a density operator ρ with finite moments up to an even number, say $2r$, there exists a density operator with all finite moments, that is a Schwartz-density operator ρ_ε, which approximates ρ in the strongest possible sense $\|\rho - \rho_\varepsilon\|_1 \to 0$. Moreover, the symmetric moments of this Schwartz-density operator approximate arbitrarily well the symmetric moments of the original state ρ. This can be seen by applying the reverse triangle inequality in (i) and for $x, y \geq 0$, $|x - y| = |\sqrt{x} - \sqrt{y}||\sqrt{x} + \sqrt{y}|$. That is, for all $\alpha, \beta \in I_n$ with $|\alpha| + |\beta| \leq r$

$$\left|\operatorname{Tr} Q^\alpha P^\beta \rho_\varepsilon P^\beta Q^\alpha - \operatorname{Tr} Q^\alpha P^\beta \rho P^\beta Q^\alpha\right| \longrightarrow 0 \qquad \text{as } \varepsilon \to 0.$$

Before proving Theorem 3.4.4 we introduce some needed definitions and tools.

3.4.1 Sobolev operators

Let us recall the notion of weak differentiability: $\varphi \in \mathcal{L}^2(\mathbb{R}^n)$ is *weakly differentiable* in the x_j direction if there exists $f \in \mathcal{L}^2(\mathbb{R}^n)$, called the weak derivative of φ, such that

$$\int \overline{\varphi(x)} \frac{\partial \psi}{\partial x_j}(x) dx = \int \overline{f(x)} \psi(x) dx \qquad \text{for all } \psi \in \mathscr{S}(\mathbb{R}^n).$$

When the weak derivative exist, these linear functionals are Hilbert space bounded, that is

$$\sup \left\{ \left| \left\langle \varphi \middle| \frac{\partial \psi}{\partial x_j} \right\rangle \right| : \psi \in \mathscr{S}(\mathbb{R}^n), \|\psi\|_2 = 1 \right\} = \|f\|_2 < \infty.$$

In the case that the classical derivative of φ exists and is continuous, the function $f(x) = -\frac{\partial \varphi}{\partial x_j}(x)$ is in $\mathcal{L}^2(\mathbb{R}^n)$ (see Theorem 6.10 in [50]). Conversely, if the linear functionals

$$\psi \mapsto \left\langle \varphi \middle| \frac{\partial \psi}{\partial x_j} \right\rangle$$

are continuous and bounded, then by Riesz Lemma there is a function $f \in \mathcal{L}^2(\mathbb{R}^n)$ such that $\left\langle \varphi \middle| \frac{\partial \psi}{\partial x_j} \right\rangle = \langle f | \psi \rangle$. From now on, we denote the weak derivative of φ as $D^\alpha \varphi$.

The first Sobolev space of functions that we consider is the set of functions $\varphi \in \mathcal{L}^2(\mathbb{R}^n)$ that are

(i) weakly differentiable with respect to a basis $x_j \in \mathbb{R}^n$, and

(ii) the Fourier transform of φ is weakly differentiable. Equivalently, for a basis $x_j \in \mathbb{R}^n$, $x_j\varphi(x) \in \mathcal{L}^2(\mathbb{R}^n)$.

This space is denoted by $\mathscr{S}_1^2(\mathbb{R}^n)$. This definition extends to the case of higher derivatives. The following is a useful characterization of Sobolev functions in $\mathscr{S}_r^2(\mathbb{R}^n)$ without assuming differentiability.

Lemma 3.4.5 (Characterization of $\mathscr{S}_r^2(\mathbb{R}^n)$)**.** *Let $\varphi \in \mathcal{L}^2(\mathbb{R}^n)$. Then $\varphi \in \mathscr{S}_r^2(\mathbb{R}^n)$ if and only if for $\alpha, \beta \in I_n$ with $|\alpha| + |\beta| \le r$, the $\mathscr{S}(\mathbb{R}^n)-$continuous linear functional*

$$\mathscr{S}(\mathbb{R}^n) \ni \psi \mapsto \langle\varphi|P^\beta Q^\alpha \psi\rangle \in \mathbb{C}$$

is Hilbert space bounded, that is for $\alpha, \beta \in I_n$ with $|\alpha| + |\beta| \le r$,

$$\|\varphi\|_{\alpha,\beta} := \sup\{|\langle\varphi|P^\beta Q^\alpha\psi\rangle| : \psi \in \mathscr{S}(\mathbb{R}^n), \|\psi\|_2 = 1\} < \infty.$$

When the derivatives of φ exist and are continuous we have in fact

$$\|\varphi\|_{\alpha,\beta} = \left\|Q^\alpha P^\beta \varphi\right\|_2 = \left(\int \left|x^\alpha \frac{\partial^\beta \varphi}{\partial x^\beta}(x)\right|^2 dx\right)^{1/2}.$$

Example 3.4.6 ($\mathscr{S}_1^2(\mathbb{R})$)**.** *The box-car state Ψ of example 3.3.2 is not in $\mathscr{S}_1^2(\mathbb{R})$ since $\|\psi\|_{0,1}$ is not bounded. However, the state*

$$\Lambda(x) := \psi * \psi(x) = \begin{cases} 1 - \frac{|x|}{l} & |x| \le l, \\ 0 & |x| > l, \end{cases}$$

is in $\mathscr{S}_1^2(\mathbb{R})$. The state $|\Lambda\rangle \in \mathcal{L}^2(\mathbb{R})$ has finite moments in Q for all orders, but only finite moments in P up to third order, i.e. $\mathrm{Tr}[|\Lambda\rangle\langle\Lambda|P^m] < \infty$ *for $m \le 3$. Moreover, we have as well that the mixed moment* $\mathrm{Tr}[|\Lambda\rangle\langle\Lambda|PQ] \le \|\Lambda\|_{1,0} \|\Lambda\|_{0,1}$ *is finite. See Figure 3.2.*

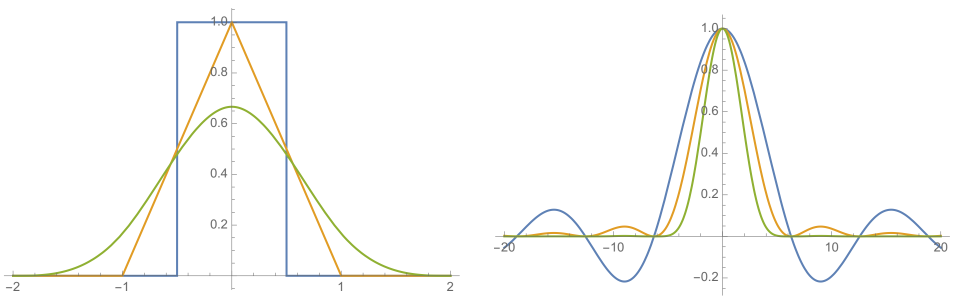


Figure 3.2: (Left) The blue, orange and green line are the space representation of the states ψ (Box-car), Λ and $\Lambda * \Lambda$ with $l = 1$, respectively. (Right) The blue, orange and green line are $\mathrm{Sinc}(k/2)$, $\mathrm{Sinc}^2(k/2)$ and $\mathrm{Sinc}^4(k/2)$, respectively; the Fourier transforms of ψ, Λ and $\Lambda * \Lambda$ are proportional to these functions.

The classical theorem of Meyers and Serrin says that the following spaces are the same

$$H_r(\mathbb{R}^n) := \left\{ \text{The completion of } \varphi \in C^\infty(\mathbb{R}^{2n}) \text{ with respect to the norm } \|\varphi\|_r := \sum_{0\leq\beta\leq r} \|\varphi\|_{0,\beta} \right\}$$
$$= \left\{ \varphi \in \mathcal{L}^2(\mathbb{R}^n) | D^\alpha\varphi \in \mathcal{L}^2(\mathbb{R}^n) \text{ for } 0 \leq |\alpha| \leq r \right\}.$$

In fact, it is shown that the compactly supported functions $C_0^\infty(\mathbb{R}^n)$ are dense in $H_r(\mathbb{R}^n)$. Since $C_0^\infty(\mathbb{R}^n)$ is dense in $\mathscr{S}(\mathbb{R}^n)$ and $\mathscr{S}(\mathbb{R}^n) \subset H_r(\mathbb{R}^n) \subset \mathcal{L}^2(\mathbb{R}^n)$ it follows that the Schwartz space is also dense in the classical Sobolev space $H_r(\mathbb{R}^n)$. It can be shown [50] that $H_r(\mathbb{R}^n)$ is a Hilbert space with inner product

$$(\varphi, \psi)_r := \sum_{0\leq|\alpha|\leq r} \langle D^\alpha\varphi | D^\alpha\psi\rangle,$$

where $\langle\varphi|\psi\rangle = \int_{\mathbb{R}^n} \overline{\varphi(x)}\psi(x)dx$ is the inner product in $\mathcal{L}^2(\mathbb{R}^n)$.

We are now ready to introduce the set of Sobolev operators. Let $A \in \mathfrak{B}(\mathcal{H})$ and consider the sesquilinear form

$$\mathscr{S}(\mathbb{R}^n) \times \in \mathscr{S}(\mathbb{R}^n) \ni (\psi, \varphi) \mapsto \langle P^\beta Q^\alpha \psi | A P^{\beta'} Q^{\alpha'} \varphi\rangle \in \mathbb{C}$$

which is well-defined and jointly continuous.

Definition 3.4.7 (Sobolev operators)**.** *Let $A \in \mathfrak{B}(\mathcal{H})$. If for $\alpha, \beta, \alpha^{'}, \beta^{'} \in I_N$ with $|\alpha| + |\alpha'| + |\beta| + |\beta'| \leq r$*

$$\|A\|_{\alpha,\alpha',\beta,\beta'} := \sup\{|\langle P^\beta Q^\alpha \psi | A P^{\beta'} Q^{\alpha'} \varphi\rangle| : \psi, \varphi \in \mathscr{S}(\mathbb{R}^n), \|\psi\|_2 \leq 1, \|\varphi\|_2 \leq 1\} < \infty,$$

then we say A is a Sobolev operator. The set of Sobolev operators is denoted by $\mathscr{S}_r^2(\mathcal{H})$.

From Riesz Lemma we know there exists a unique $A_{\alpha,\alpha',\beta,\beta'} \in \mathfrak{B}(\mathcal{H})$ such that

$$\langle P^\beta Q^\alpha \psi | A P^{\beta'} Q^{\alpha'} \varphi\rangle = \langle \psi | A_{\alpha,\alpha',\beta,\beta'} \varphi\rangle \qquad \varphi, \psi \in \mathscr{S}(\mathbb{R}^n).$$

We show that when $A \in \mathscr{S}_r^2(\mathcal{H})$, the operator $Q^\alpha P^\beta A P^{\beta'} Q^{\alpha'}$ is well-defined on $\mathscr{S}(\mathbb{R}^n)$ so that from the previous equality $A_{\alpha,\alpha',\beta,\beta'}$ is its bounded extension. It will suffice to show that A maps into the domain of $Q^\alpha P^\beta$:

Lemma 3.4.8. *If $\|A\|_{\alpha,0,\beta,0} < \infty$ for all $|\alpha| + |\beta| \leq r$, then $Ran(A) \subset \mathscr{S}_r^2(\mathbb{R}^n)$ and*

$$A : \mathcal{H} \to \mathscr{S}_r^2(\mathbb{R}^n) \qquad \textit{is continuous.}$$

Proof. Since $\|A\|_{\alpha,0,\beta,0} < \infty$ and $\mathscr{S}(\mathbb{R}^n)$ is dense in $\mathcal{L}^2(\mathbb{R}^n)$ we have

$$\sup\{|\langle P^\beta Q^\alpha \psi | A\varphi\rangle| : \psi \in \mathscr{S}(\mathbb{R}^n), \|\psi\|_2 \leq 1\} < \infty$$

for all $\varphi \in \mathcal{H}$. Hence from Lemma 3.4.5, $A\varphi \in \mathscr{S}_r^2(\mathbb{R}^n)$ and $\|A\varphi\|_{\alpha,\beta} \leq \|A\|_{\alpha,0,\beta,0} \|\varphi\|_2$. The last inequality follows from the definition of $\|\cdot\|_{\alpha,0,\beta,0}$. ☐

The previous Lemma implies that for a Sobolev operator $A \in \mathscr{S}^2_r(\mathcal{H})$, we have the equality $\|A\|_{\alpha,\beta,\alpha',\beta'} = \left\|Q^\alpha P^\beta A P^{\beta'} Q^{\alpha'}\right\|_\infty$. We remark that the order of the operators in the definition of Sobolev operators (and Sobolev vectors) is not relevant since Q and P map $\mathscr{S}(\mathbb{R}^n)$ into itself and the CCR relations Eq. (3.2) hold on $\mathscr{S}(\mathbb{R}^n)$. Moreover, the space of quantum states $\rho \in \mathcal{B}(\mathcal{H})$ such that the symmetric moments $\operatorname{Tr} Q^\alpha P^\beta \rho P^\beta Q^\alpha$ for $|\alpha| + |\beta| \leq r$ are finite is the revelant subset $\mathscr{S}^2_r(\mathcal{H})$ which we are interested here.

Before proving Theorem 3.4.4 we introduce some needed lemmas and tools. The *symplectic (or twisted) convolution* [51, 52] of two square-integrable functions χ_1 and χ_2 is defined as

$$\begin{aligned}\chi_1 *_\sigma \chi_2(\xi) &:= \frac{1}{(2\pi)^n} \int \chi_1(\xi - \eta)\chi_2(\eta)\ e^{\frac{i}{2}\eta\cdot\sigma\xi}\ d\eta, \\ &= \frac{1}{(2\pi)^n} \int \chi_2(\xi - \eta)\chi_1(\eta)\ e^{-\frac{i}{2}\eta\cdot\sigma\xi}\ d\eta =: \chi_2 *_{-\sigma} \chi_1(\xi)\end{aligned}$$

When $\chi_1(\xi)$ and $\chi_2(\xi)$ are the inverse Weyl transform of the Hilbert-Schmidt operators A_1, A_2 respectively, then $\chi_1 *_\sigma \chi_2(\xi)$ is the inverse Weyl transform of the trace-class operator A_1A_2. This fact can be seen by computing the integral in the definition of twisted convolution with the help of the quantum parseval theorem and Eq. (3.3). Although it is not a commutative operation, it is associative, distributive and satisfies Young's inequality

$$\|f *_\sigma g\|_{\mathcal{L}^r} \leq \frac{\|f\|_{\mathcal{L}^p}\,\|g\|_{\mathcal{L}^q}}{(2\pi)^n} \qquad \text{when} \quad \frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1. \tag{3.8}$$

Lemma 3.4.9 (Schwartz Convolution)**.** *Let G be a Schwartz function in $\mathbb{R}^{2n}$ and $\chi \in \mathcal{L}^2(\mathbb{R}^{2n})$. Then the symplectic convolution of these two functions is a Schwartz function.*

Proof. We consider the Hilbert Schmidt operator $T : \mathcal{L}^2(\mathbb{R}^{2n}) \to \mathcal{L}^2(\mathbb{R}^{2n})$ defined by $T(\chi) = G *_\sigma \chi$. Now, recall that an operator is a Schwartz operator if and only if the kernel is a Schwartz function (Theorem 3.3.1 (*ii*)). Then clearly T is a Schwartz operator as the kernel function $G(\xi - \eta)e^{-\frac{i}{2}\eta\cdot\sigma\xi}$ is a Schwartz function. Finally, by Theorem 3.3.1 (*i*) we know that the range of a Schwartz operator is a Schwartz function. This finishes the proof. ☐

Let us consider the following integrable functions and their respective Weyl transforms (see theorem 3.1.1)

$$\begin{aligned}\mathcal{L}^1(\mathbb{R}^{2n}) \ni G_\varepsilon(\eta) := \frac{1}{\varepsilon^{2n}} \exp[-\eta^2/2\varepsilon^2] \quad &\longleftrightarrow \quad \hat{g}_\varepsilon = \frac{1}{(2\pi)^n} \int G_\varepsilon(\eta) W_{-\eta} d\eta, \\ \mathcal{L}^2(\mathbb{R}^{2n}) \ni \chi(\eta) = \operatorname{Tr} W_\eta A \longleftrightarrow \quad &A = \frac{1}{(2\pi)^n} \int (\operatorname{Tr} W_\eta A)\, W_{-\eta} d\eta,\end{aligned}$$

Using these functions we can write the inverse Weyl transform of the operator $A\hat{g}_\varepsilon$,

$$\begin{aligned}\chi_\varepsilon := \chi *_\sigma G_\varepsilon \quad &\longleftrightarrow \quad A\hat{g}_\varepsilon, \\ \chi_\varepsilon(\xi) = \frac{1}{(2\pi)^n} &\int \chi(\xi - \eta)\ G_\varepsilon(\eta)\ e^{\frac{i}{2}\eta\cdot\sigma\xi}\ d\eta.\end{aligned}$$

Lemma 3.4.10 (Approximation by Schwartz operators)**.** *Let $G_\varepsilon(\xi) = \frac{1}{\varepsilon^{2n}} \exp[-\xi^2/2\varepsilon^2]$ with $\varepsilon > 0$ and A a Hilbert-Schmidt operator with corresponding inverse Weyl transform $\chi(\xi) \in \mathcal{L}^2(\mathbb{R}^{2n})$. Define*

$$\chi_\varepsilon := \chi *_\sigma G_\varepsilon.$$

Then

(i) χ_ε *is an Schwartz function with* $\|\chi_\varepsilon\|_{\mathcal{L}^2} \leq \|\chi\|_{\mathcal{L}^2}$*. Equivalently* Ag_ε *is a Schwartz operator with* $\|Ag_\varepsilon\|_2 \leq \|A\|_2$*.*

(ii) $\|\chi_\varepsilon - \chi\|_{\mathcal{L}^2(\mathbb{R}^{2n})} \longrightarrow 0$ *and* $\|\overline{\chi}_\varepsilon - \chi\|_{\mathcal{L}^2(\mathbb{R}^{2n})} \longrightarrow 0$ *as* $\varepsilon \longrightarrow 0$*. Equivalently* $\|Ag_\varepsilon - A\|_2 \longrightarrow 0$ *and* $\|g_\varepsilon A - A\|_2 \longrightarrow 0$ *as* $\varepsilon \longrightarrow 0$*.*

Note that from this lemma we obtain $\|g_\varepsilon A g_\varepsilon - Ag_\varepsilon\|_2 = \|(g_\varepsilon A - A)g_\varepsilon\|_2 \leq \|g_\varepsilon A - A\|_2$ which goes to zero as ε approaches zero. Therefore, one can show via the triangle inequality and induction that for $m \in \mathbb{N}$, $\|g_\varepsilon^m A g_\varepsilon^m - A\|_2$ vanishes in the limit.

Proof of lemma 3.4.10. For part (i) we have that χ_ε is Schwartz from Lemma 3.4.9 and the fact that a Gaussian function is a Schwartz function. Using Young's inequality Eq. (3.8) and $\|G_\varepsilon\|_{\mathcal{L}^1} = (2\pi)^n$ we get an upper bound for the norm $(2\pi)^n \|\chi *_\sigma G_\varepsilon\|_{\mathcal{L}^2} \leq \|\chi\|_{\mathcal{L}^2} \|G_\varepsilon\|_{\mathcal{L}^1} = (2\pi)^n \|\chi\|_{\mathcal{L}^2}$. For part (ii) we use the triangle inequality to bound

$$\|\chi_\varepsilon - \chi\|_{\mathcal{L}^2(\mathbb{R}^{2n})} \leq \|\chi *_\sigma G_\varepsilon - \chi * G_\varepsilon\|_{\mathcal{L}^2(\mathbb{R}^{2n})} + \|\chi * G_\varepsilon - \chi\|_{\mathcal{L}^2(\mathbb{R}^{2n})}.$$

The second term goes to zero as ε goes to zero due to the classical result on approximating $\mathcal{L}^p$ functions by smooth functions (see for instance Theorem 2.16 in Ref. [50]). For the remaining term, we use a standard dominated convergence argument. Indeed,

$$\begin{aligned}
\|\chi *_\sigma G_\varepsilon - \chi * G_\varepsilon\|^2_{\mathcal{L}^2(\mathbb{R}^{2n})} &= \int |\chi *_\sigma G_\varepsilon(\xi) - \chi * G_\varepsilon(\xi)|^2 \, d\xi, \\
&\leq \frac{1}{(2\pi)^{2n}} \int \left(\int |\chi(\xi - \eta)||e^{i\eta\cdot\sigma\xi/2} - 1| G_\varepsilon(\eta) d\eta \right)^2 d\xi, \\
&\leq \frac{1}{(2\pi)^{2n}} \int \left(\int |\chi(\xi - \eta)|^2 |e^{i\eta\cdot\sigma\xi/2} - 1|^2 G_\varepsilon(\eta) d\eta \right) \left(\int G_\varepsilon(\eta') d\eta' \right) d\xi, \\
&= \frac{1}{(2\pi)^n} \int G_\varepsilon(\eta) \left(\int |\chi(\xi - \eta)|^2 |e^{i\eta\cdot\sigma\xi/2} - 1|^2 d\xi \right) d\eta, \\
&= \frac{1}{(2\pi)^n} \int e^{-\eta^2/2} \left(\int |\chi(\xi - \varepsilon\eta)|^2 |e^{i\varepsilon\eta\cdot\sigma\xi/2} - 1|^2 d\xi \right) d\eta, \\
&= \frac{1}{(2\pi)^n} \int e^{-\eta^2/2} \left(\int |\chi(\xi)|^2 |e^{i\varepsilon\eta\cdot\sigma\xi/2} - 1|^2 d\xi \right) d\eta.
\end{aligned}$$

Here we have just used the Cauchy-Schwarz inequality in the second inequality, Fubini-Tonelli's theorem in the second inequality and finally the change of variables $\eta \mapsto \varepsilon\eta$ and $\chi \mapsto \chi - \varepsilon\eta$ in the last equalities. Since for $x \in \mathbb{R}$, $|e^{ix} - 1| \leq 2$ and $|e^{ix} - 1| \to 0$ as $x \to 0$, we have by the dominated convergence theorem that $\|\chi *_\sigma G_\varepsilon - \chi * G_\varepsilon\|^2_{\mathcal{L}^2(\mathbb{R}^{2n})}$ vanishes in the limit $\varepsilon \to 0$. From the quantum parseval theorem we obtain the equivalent statement in terms of Hilbert-Schmidt operators. Finally, the other limit is the same as $\overline{\chi}_\varepsilon(\xi) = G_\varepsilon *_{-\sigma} \chi(-\xi)$.

□

In comparison with the classical convolution, the twisted derivative of a twisted convolution does not commute in general. Instead, we have the following identities which follow directly from the definitions.

Lemma 3.4.11. *Let χ_1, χ_2 be such that $L\chi_1, L\chi_2 \in \mathcal{L}^2(\mathbb{R}^{2n})$ and denote the corresponding Weyl transforms of these functions by A_1 and A_2. Then*

$$\begin{aligned}
L(\chi_1 *_\sigma \chi_2)(\xi) &= (L\chi_1) *_\sigma \chi_2(\xi) = \operatorname{Tr} W_\xi R A_1 A_2,\\
\chi_1 *_\sigma L\chi_2(\xi) &= \bar{L}\chi *_\sigma \chi_2(\xi) = \operatorname{Tr} W_\xi A_1 R A_2,\\
\chi_1 *_\sigma \bar{L}\chi_2(\xi) &= L\chi_1 *_\sigma \chi_2(\xi) - (z\cdot\sigma\xi)\chi_1 *_\sigma \chi_2(\xi),\\
\operatorname{Tr} W_\xi A_1 A_2 R &= \operatorname{Tr} W_\xi R A_1 A_2 - (z\cdot\sigma\xi) \operatorname{Tr} W_\xi A_1 A_2.
\end{aligned}$$

However, we still have the following important result.

Lemma 3.4.12 (Approximation of twisted derivatives). *Let $\chi \in \mathcal{L}^2(\mathbb{R}^{2n})$ be such that $L\chi \in \mathcal{L}^2(\mathbb{R}^{2n})$. Then*

$$\|LG_\varepsilon *_\sigma \chi - L\chi\|_{\mathcal{L}^2} \longrightarrow 0 \qquad \text{as } \varepsilon \to 0.$$

Proof. The proof consists in a repetitive use of Fubini-Tonelli and dominated convergence theorem as in the proof of part (ii) of Lemma 3.4.10. We use that for $\varepsilon > 0$, G_ε is a Schwartz function to separate the integral in the definition of the symplectic convolution of LG_ε and χ

$$\begin{aligned}
LG_\varepsilon *_\sigma \chi(\xi) = \chi *_{-\sigma} LG_\varepsilon(\xi) &= \frac{1}{(2\pi)^n}\int \left(\frac{1}{2}z\cdot\sigma\eta\right) G_\varepsilon(\eta)\chi(\xi-\eta)e^{-i\eta\cdot\sigma\xi/2}d\eta\\
&\quad - \frac{i}{(2\pi)^n}\int \frac{\partial G_\varepsilon}{\partial z}(\eta)\chi(\xi-\eta)e^{-i\eta\cdot\sigma\xi/2}d\eta.
\end{aligned}$$

Since $\chi(\xi) \in \mathcal{L}^2(\mathbb{R}^{2n})$, it vanishes when $\xi \to \infty$. Thus we can use partial integration and $\chi_1 *_\sigma \chi_2(\xi) = \chi_2 *_{-\sigma} \chi_1(\xi)$ to rewrite

$$\begin{aligned}
LG_\varepsilon *_\sigma \chi(\xi) &= \frac{1}{(2\pi)^n}\int \left(\frac{1}{2}z\cdot\sigma\eta\right) G_\varepsilon(\eta)\chi(\xi-\eta)e^{-i\eta\cdot\sigma\xi/2}d\eta + \frac{i}{(2\pi)^n}\int G_\varepsilon(\eta)\frac{\partial\chi}{\partial z}(\xi-\eta)e^{-i\eta\cdot\sigma\xi/2}d\eta\\
&\quad - \frac{1}{(2\pi)^n}\left(\frac{1}{2}z\cdot\sigma\xi\right)\int G_\varepsilon(\eta)\chi(\xi-\eta)e^{-i\eta\cdot\sigma\xi/2}d\eta,\\
&= \left(\Delta G_\varepsilon *_\sigma \chi\right)(\xi) + i\left(G_\varepsilon *_\sigma \frac{\partial\chi}{\partial z}\right)(\xi) - \left(\frac{1}{2}z\cdot\sigma\xi\right)(G_\varepsilon *_\sigma \chi)(\xi),
\end{aligned} \tag{3.9}$$

where $\Delta G_\varepsilon(\eta) := \left(\frac{1}{2}z\cdot\sigma\eta\right)G_\varepsilon(\eta)$. From now on, for any function f we write in short $\Delta f(\eta) := \left(\frac{1}{2}z\cdot\sigma\eta\right)f(\eta)$. We consider

$$\|LG_\varepsilon *_\sigma \chi - L\chi\|^2_{\mathcal{L}^2} \le \left|\langle LG_\varepsilon *_\sigma \chi, LG_\varepsilon *_\sigma \chi\rangle - \langle L\chi, L\chi\rangle\right| + 2\left|\langle L\chi, L\chi\rangle - \langle L\chi, LG_\varepsilon *_\sigma \chi\rangle\right|, \tag{3.10}$$

and show first that

$$\left|\langle LG_\varepsilon *_\sigma \chi, LG_\varepsilon *_\sigma \chi\rangle - \langle L\chi, L\chi\rangle\right| \longrightarrow 0 \qquad \text{as } \varepsilon \to 0. \tag{3.11}$$

We insert Eq. (3.9) in the previuos equation and bound each term separately. The first term we bound is

$$\begin{aligned}
(2\pi)^n \|\Delta G_\varepsilon *_\sigma \chi\|_{\mathcal{L}^2} &\le \|\Delta G_\varepsilon\|_{\mathcal{L}^1}\|\chi\|_{\mathcal{L}^2},\\
&= \varepsilon\|\Delta G\|_{\mathcal{L}^1}\|\chi\|_{\mathcal{L}^2},
\end{aligned}$$

where we have used Young's inequality. Since G is Schwartz, $\|\Delta G\|_{\mathcal{L}^1} \|\chi\|_{\mathcal{L}^2} < \infty$ and this term vanishes in the limit. As a consequence, all the inner products with a square integrable function and $\Delta G_\varepsilon *_\sigma \chi$ vanish by the Cauchy-Schwarz inequality. Indeed,

$$\begin{aligned}\langle \Delta G_\varepsilon *_\sigma \chi, G_\varepsilon *_\sigma \frac{\partial \chi}{\partial z}\rangle &\leq \|\Delta G_\varepsilon *_\sigma \chi\|_{\mathcal{L}^2} \left\| G_\varepsilon *_\sigma \frac{\partial \chi}{\partial z}\right\|_{\mathcal{L}^2}, \\ &\leq \|\Delta G_\varepsilon *_\sigma \chi\|_{\mathcal{L}^2} \left\| \frac{\partial \chi}{\partial z}\right\|_{\mathcal{L}^2},\end{aligned}$$

where we have used again Young's inequality in the last step and Lemma 3.4.10(i). Likewise

$$\int (\Delta G_\varepsilon *_\sigma \chi)(\xi) \left(\frac{1}{2} z\cdot \sigma\xi\right) (\bar{\chi} *_\sigma G_\varepsilon)(\xi) d\xi \leq \|\Delta G_\varepsilon *_\sigma \chi\|_2 \left(\int \left|\frac{1}{2}z\cdot\sigma\xi\right|^2 |\bar{\chi} *_\sigma G_\varepsilon(\xi)|^2 d\xi \right)^{1/2}.$$

The last integral on the previous equation can be shown to be finite by the same arguments as in the proof of Lemma 3.4.10:

$$\begin{aligned}\int \left|\frac{z\cdot\sigma\xi}{2}\right|^2 |\bar{\chi} *_\sigma G_\varepsilon(\xi)|^2 d\xi &\leq \frac{1}{(2\pi)^{2n}} \int \left|\frac{z\cdot\sigma\xi}{2}\right|^2 \left(\int |\chi(\xi-\eta)| G_\varepsilon(\eta) d\eta\right)^2 d\xi, \\ &\leq \frac{1}{(2\pi)^n} \int \left|\frac{z\cdot\sigma\xi}{2}\right|^2 |\chi(\xi-\eta)|^2 G_\varepsilon(\eta) d\xi\, d\eta, \\ &= \frac{1}{(2\pi)^n} \left(\int \left|\frac{z\cdot\sigma\xi}{2} + \frac{z\cdot\sigma\eta}{2}\right|^2 |\chi(\xi)|^2 d\xi \right) G_\varepsilon(\eta)\, d\eta, \\ &\leq 2\|\Delta\chi\|^2_{\mathcal{L}^2} + \frac{\varepsilon^2 \|\chi\|_2^2}{(2\pi)^n} \int \left|\frac{z\cdot\sigma\eta}{2}\right|^2 G(\eta) d\eta.\end{aligned}$$

The next term we bound is

$$\left| \left\| G_\varepsilon *_\sigma \frac{\partial\chi}{\partial z}\right\|_{\mathcal{L}^2} - \left\|\frac{\partial \chi}{\partial z}\right\|_{\mathcal{L}^2} \right| \leq \left\| G_\varepsilon *_\sigma \frac{\partial\chi}{\partial z} - \frac{\partial\chi}{\partial z}\right\|_{\mathcal{L}^2},$$

which vanishes in the limit because $\frac{\partial\chi}{\partial z} \in \mathcal{L}^2(\mathbb{R}^{2n})$ and Lemma 3.4.10 applies. We consider now

$$\begin{aligned}&\left| \int |z\cdot\sigma\xi|^2 |\bar{\chi} *_\sigma G_\varepsilon(\xi)|^2 d\xi - \int |z\cdot\sigma\xi|^2 |\chi(\xi)|^2 d\xi \right| \leq \\ &\left| \int |z\cdot\sigma\xi|^2 |\bar{\chi} *_\sigma G_\varepsilon(\xi)|^2 d\xi - \int |z\cdot\sigma\xi|^2 |\bar{\chi} * G_\varepsilon(\xi)|^2 d\xi \right| \\ &\quad + \left| \int |z\cdot\sigma\xi|^2 |\bar{\chi} * G_\varepsilon(\xi)|^2 d\xi - \int |z\cdot\sigma\xi|^2 |\chi(\xi)|^2 d\xi \right|. \end{aligned} \tag{3.12}$$

The first term of the RHS can be bounded again as before

$$\left|\int |z\cdot\sigma\xi|^2\,|\bar{\chi}*_\sigma G_\varepsilon(\xi)|^2\,d\xi - \int |z\cdot\sigma\xi|^2\,|\bar{\chi}*G_\varepsilon(\xi)|^2\,d\xi\right| \le \int |z\cdot\sigma\xi|^2\,|\bar{\chi}*_\sigma G_\varepsilon(\xi)-\bar{\chi}*G_\varepsilon(\xi)|^2\,d\xi,$$
$$\le \frac{1}{(2\pi)^{2n}}\int |z\cdot\sigma\xi|^2\left(\int |\chi(\xi-\eta)|G_\varepsilon(\eta)|e^{i\eta\cdot\sigma\xi/2}-1|d\eta\right)^2 d\xi,$$
$$\le \frac{1}{(2\pi)^{n}}\int |z\cdot\sigma\xi|^2|\chi(\xi-\eta)|^2G_\varepsilon(\eta)|e^{i\eta\cdot\sigma\xi/2}-1|^2\,d\eta\,d\xi,$$
$$= \frac{1}{(2\pi)^{n}}\int |z\cdot\sigma\xi|^2|\chi(\xi-\varepsilon\eta)|^2G(\eta)|e^{i\varepsilon\eta\cdot\sigma\xi/2}-1|^2\,d\eta\,d\xi,$$
$$= \frac{1}{(2\pi)^{n}}\int |z\cdot\sigma\xi+\varepsilon\cdot\sigma\eta|^2|\chi(\xi)|^2G(\eta)|e^{i\varepsilon\eta\cdot\sigma\xi/2}-1|^2\,d\xi\,d\eta,$$
$$= 2\int |z\cdot\sigma\xi|^2|e^{i\varepsilon\eta\cdot\sigma\xi/2}-1|^2\,d\xi + \frac{2\varepsilon\,\|\chi\|_2^2}{(2\pi)^n}\int |z\cdot\sigma\eta|^2\,G(\eta)d\eta,$$

which vanishes in the limit by the dominated convergence theorem. For the second term of Eq. (3.12) we use the classical parseval theorem $\|\chi\|_{\mathcal{L}^2}=(2\pi)^n\,\|\mathfrak{F}_\sigma[\chi]\|_{\mathcal{L}^2}$ where

$$\mathfrak{F}_\sigma[\chi](\eta)=\frac{1}{(2\pi)^{2n}}\int e^{i\eta\cdot\sigma\xi}\chi(\xi)\,d\xi,$$

is the inverse symplectic Fourier transform $\mathfrak{F}_\sigma[\chi]$ of χ (whenever χ_ρ is a quantum characteristic function, $\mathfrak{F}_\sigma[\chi_\rho]$ is just the Wigner function of the density operator ρ). The inverse symplectic Fourier transform of G_ε is

$$\mathfrak{F}_\sigma[G_\varepsilon](\eta)=\frac{1}{(2\pi)^n}e^{-\frac{\varepsilon^2\eta^2}{2}},$$

so with the classical convolution theorem $\mathfrak{F}_\sigma[G_\varepsilon*\chi](\eta)=(2\pi)^n\mathfrak{F}_\sigma[G_\varepsilon](\eta)\mathfrak{F}_\sigma[\chi](\eta)$ and derivative formulas we obtain

$$\left|\int |z\cdot\sigma\xi|^2\,|\bar{\chi}*G_\varepsilon(\xi)|^2\,d\xi - \int |z\cdot\sigma\xi|^2\,|\chi(\xi)|^2\,d\xi\right| \le \int |z\cdot\sigma\xi|^2\,|G_\varepsilon*\chi(\xi)-\chi(\xi)|^2\,d\xi,$$
$$=\int\left|\frac{\partial\mathfrak{F}_\sigma[\chi]}{\partial z}(\eta)\left((2\pi)^n\mathfrak{F}_\sigma[G_\varepsilon](\eta)-1\right)+(2\pi)^n\mathfrak{F}_\sigma[\chi](\eta)\frac{\partial\mathfrak{F}_\sigma[G_\varepsilon]}{\partial z}(\eta)\right|^2 d\eta,$$
$$\le 2\int\left|\frac{\partial\mathfrak{F}_\sigma[\chi]}{\partial z}(\eta)\right|^2\left|e^{-\varepsilon^2\eta^2/2}-1\right|^2\,d\eta+\varepsilon(2\pi)^n\int|\mathfrak{F}_\sigma[\chi](\eta)(z\cdot\sigma\eta)|^2\,d\eta,$$

Using the fact that $L\mathfrak{F}_\sigma[\chi]\in\mathcal{L}^2(\mathbb{R}^{2n})$ and applying dominated convergence theorem gives that this term vanishes in the limit. Finally, we bound the last term using the triangle inequality and the previous bounds. Here we abuse a bit in notation in order to keep it short and write the argument of the function inside the $\mathcal{L}_2$ inner product

$$\left|\langle(z\cdot\sigma\xi)(G_\varepsilon*_\sigma\chi)(\xi),G_\varepsilon*_\sigma\frac{\partial\chi}{\partial z}\rangle-\langle(z\cdot\sigma\xi)\chi(\xi),\frac{\partial\chi}{\partial z}\rangle\right|\le\left|\langle(z\cdot\sigma\xi)(G_\varepsilon*_\sigma\chi-G_\varepsilon*\chi)(\xi),G_\varepsilon*_\sigma\frac{\partial\chi}{\partial z}\rangle\right|$$
$$+\left|\langle(z\cdot\sigma\xi)(G_\varepsilon*\chi-\chi)(\xi),G_\varepsilon*_\sigma\frac{\partial\chi}{\partial z}\rangle\right|-\left|\langle(z\cdot\sigma\xi)\chi(\xi),G_\varepsilon*_\sigma\frac{\partial\chi}{\partial z}-\frac{\partial\chi}{\partial z}\rangle\right|.$$

Using Cauchy-Schwarz inequality, the previous estimates and Lemma 3.4.10 we find that this term vanishes in the limit.

In the same fashion we bound the second term of the RHS of Eq. (3.10)

$$
\begin{aligned}
|\langle L\chi, L\chi\rangle - \langle L\chi, LG_\varepsilon *_\sigma \chi\rangle| &\leq |\langle L\chi, \Delta G_\varepsilon *_\sigma \chi\rangle| + \left|\langle L\chi, \left(\frac{1}{2}z\cdot\sigma\xi\right)(\chi - G_\varepsilon *_\varepsilon \chi)(\xi)\rangle\right| \\
&\quad + \left|\langle L\chi, G_\varepsilon *_\sigma \frac{\partial\chi}{\partial z} - \frac{\partial\chi}{\partial z}\rangle\right|, \\
&\leq \|L\chi\|_{\mathcal{L}^2}\left(\|\Delta G_\varepsilon *_\sigma \chi\|_{\mathcal{L}^2} + \left\|\left(\frac{1}{2}z\cdot\sigma\xi\right)(\chi - G_\varepsilon *_\varepsilon \chi)(\xi)\right\|_{\mathcal{L}^2} + \left\|G_\varepsilon *_\sigma \frac{\partial\chi}{\partial z} - \frac{\partial\chi}{\partial z}\right\|_{\mathcal{L}^2}\right),
\end{aligned}
$$

which vanishes in the limit from what we just have done for Eq. (3.11). □

We are finally ready to prove the main result of this section, Theorem 3.4.4.

Proof of Theorem 3.4.4. For part (i) we set $\sqrt{\rho_\varepsilon} = g_\varepsilon\sqrt{\rho}g_\varepsilon/c_\varepsilon$ with $c_\varepsilon := \left(\operatorname{Tr} g_\varepsilon^2\sqrt{\rho}g_\varepsilon^2\sqrt{\rho}\right)^{1/2}$ as the square root of the Schwartz operator ρ_ε. Note that $\sqrt{\rho_\varepsilon}$ is indeed Schwartz from lemma 3.4.10(i) and that for sufficiently small ε, $c_\varepsilon > 0$. Moreover, $c_\varepsilon \to 1$ as $\varepsilon \to 0$. This can be seen from lemma 3.4.10 and Cauchy-Schwarz inequality

$$
\begin{aligned}
|c_\varepsilon^2 - 1| &= |\operatorname{Tr}(g_\varepsilon^2\sqrt{\rho}g_\varepsilon^2 - \sqrt{\rho})\sqrt{\rho}|, \\
&\leq \left\|g_\varepsilon^2\sqrt{\rho}g_\varepsilon^2 - \sqrt{\rho}\right\|_2.
\end{aligned}
$$

We consider

$$
\left\|Q^\alpha P^\beta\sqrt{\rho} - \frac{Q^\alpha P^\beta g_\varepsilon\sqrt{\rho}g_\varepsilon}{c_\varepsilon}\right\|_2 \leq \left\|Q^\alpha P^\beta\sqrt{\rho} - \frac{Q^\alpha P^\beta\sqrt{\rho}g_\varepsilon}{c_\varepsilon}\right\|_2 + \left\|\frac{Q^\alpha P^\beta\sqrt{\rho}g_\varepsilon}{c_\varepsilon} - \frac{Q^\alpha P^\beta g_\varepsilon\sqrt{\rho}g_\varepsilon}{c_\varepsilon}\right\|_2,
$$

and bound the first term by

$$
\begin{aligned}
\left\|Q^\alpha P^\beta\sqrt{\rho} - \frac{Q^\alpha P^\beta\sqrt{\rho}g_\varepsilon}{c_\varepsilon}\right\|_2 &\leq \left\|Q^\alpha P^\beta\sqrt{\rho} + Q^\alpha P^\beta\sqrt{\rho}g_\varepsilon\right\|_2 + \left|\frac{c_\varepsilon - 1}{c_\varepsilon}\right|^2 \left\|Q^\alpha P^\beta\sqrt{\rho}g_\varepsilon\right\|_2, \\
&\leq \left\|Q^\alpha P^\beta\sqrt{\rho} + Q^\alpha P^\beta\sqrt{\rho}g_\varepsilon\right\|_2 + \left|\frac{c_\varepsilon - 1}{c_\varepsilon}\right|^2 \left\|Q^\alpha P^\beta\sqrt{\rho}\right\|_2
\end{aligned}
$$

which vanishes from Lemma 3.4.10 because $Q^\alpha P^\beta\sqrt{\rho}$ is Hilbert-Schmidt.

The second term is an operator version of an extension of Lemma 3.4.12 for higher derivatives. We do not show here the extension of Lemma 3.4.12 for higher derivatives as it follows from an inductive argument; using 3.4.10(i) and quantum parseval theorem we find

$$
\begin{aligned}
\frac{1}{|c_\varepsilon|}\left\|Q^\alpha P^\beta\sqrt{\rho}g_\varepsilon - Q^\alpha P^\beta g_\varepsilon\sqrt{\rho}g_\varepsilon\right\|_2 &= \frac{1}{|c_\varepsilon|}\left\|Q^\alpha P^\beta\sqrt{\rho} - Q^\alpha P^\beta g_\varepsilon\sqrt{\rho}\right\|_2, \\
&= \frac{1}{|c_\varepsilon|(2\pi)^n}\left\|L^{\alpha\beta}\left(\chi_{\sqrt{\rho}}\right) - L^{\alpha\beta}G_\varepsilon *_\sigma \chi_{\sqrt{\rho}}\right\|_{\mathcal{L}^2},
\end{aligned}
$$

which vanishes in the limit.

Part (ii) follows from part (i) since

$$\begin{aligned}
\left|\left\|Q^\alpha P^\beta \rho\right\|_1 - \left\|Q^\alpha P^\beta \rho_\varepsilon\right\|_1\right| &\leq \left\|Q^\alpha P^\beta \rho - Q^\alpha P^\beta \rho_\varepsilon\right\|_1 , \\
&= \left\|Q^\alpha P^\beta \sqrt{\rho}\sqrt{\rho} - Q^\alpha P^\beta \sqrt{\rho_\varepsilon}\sqrt{\rho} + Q^\alpha P^\beta \sqrt{\rho_\varepsilon}\sqrt{\rho} - Q^\alpha P^\beta \sqrt{\rho_\varepsilon}\sqrt{\rho_\varepsilon}\right\|_1 , \\
&\leq \left\|(Q^\alpha P^\beta \sqrt{\rho} - Q^\alpha P^\beta \sqrt{\rho_\varepsilon})\sqrt{\rho}\right\|_1 + \left\|(Q^\alpha P^\beta \sqrt{\rho_\varepsilon})(\sqrt{\rho} - \sqrt{\rho_\varepsilon})\right\|_1 , \\
&\leq \left\|Q^\alpha P^\beta \sqrt{\rho} - Q^\alpha P^\beta \sqrt{\rho_\varepsilon}\right\|_2 \|\sqrt{\rho}\|_2 + \left\|Q^\alpha P^\beta \sqrt{\rho_\varepsilon}\right\|_2 \|\sqrt{\rho} - \sqrt{\rho_\varepsilon}\|_2 .
\end{aligned}$$

We have used the reversed and the standard triangle inequality in the first two inequalities and the trace Cauchy-Schwarz inequality in the last step. □

4 Banach space theory

This chapter introduces the main tools from Banach space theory used in the core articles A.1 and A.2. The linear approximation of an almost-symmetry will rely on the Hahn-Banach theorem. Since this theorem takes many forms, we present them here. In a sense, the linear stability problem of Wigner's theorem is a problem about extensions of non-linear maps. The best tool that we have at hand for extending a linear map is precisely the Hahn-Banach theorem. Thus, at the risk of oversimplification, one could say that much of the effort in the linear stability of an almost-symmetry consists in trying to use the Hahn-Banach theorem in an almost-linear setting.

In section 4.2 we present a sophisticated use of the Hahn-Banach theorem, namely *Maurey's extension principle* [53]. Section 4.2 and section 4.3 provide the basic notions and tools that we employ in order to obtain a better upper bound on the linear stability of Wigner's theorem (see Section A.2).

4.1 Hahn-Banach theorems

A *Banach space* is a complete normed vector space X. We will focus here on real Banach spaces, but most of the results carry over to complex Banach spaces. It is a basic fact in functional analysis that finite dimensional normed spaces are always complete because they all have equivalent norms. The *dual* of a Banach space X is the Banach space X^* of all continuous linear mappings $x^* : X \to \mathbb{R}$, endowed with the norm $\|x^*\| = \sup\{|x^*(x)| : x \in \mathcal{B}_X\}$. The first analytic version of the Hahn-Banach theorem is the following (see Chapter 3 in [54]).

Theorem 4.1.1 (Hahn-Banach extension theorem)**.** *Let X be a real vector space and $E \subset X$ a subspace.*

1. *Suppose $p : X \to \mathbb{R}$ is sub-linear, i.e. for all $x, y \in X$ and $t \geq 0$ we have*

$$p(x+y) \leq p(x) + p(y), \qquad p(tx) = tp(x).$$

If $h : X \to \mathbb{R}$ is a linear functional satisfying

$$h(x) \leq p(x) \qquad \text{for all } x \in E,$$

then there exists a linear functional $\tilde{h} : X \to \mathbb{R}$ that extends h, i.e. $h(x) = \tilde{h}(x)$ for all $x \in E$, and

$$-p(-x) \leq \tilde{h}(x) \leq p(x) \qquad \text{for all } x \in X.$$

2. *If in addition X has a norm, then for every $y^* \in E^*$ there exists a linear functional $x \in X^*$ that extends y^* and $\|x^*\| = \|y^*\|$.*

For $x \in X$ let us denote $\hat{x}$ the linear functional on X^* which acts as $\hat{x}(y^*) := y^*(x)$ for all $y^* \in X^*$. Theorem 4.1.1 implies that the natural map $x \mapsto \hat{x}$ is a norm-preserving isomorphism of a normed space X into its second dual X^{**}(see p. 52 in [55]). Thus, it is natural to consider X as a subspace of X^{**}.

We call a point x in a real vector space X an *algebraic interior point* of a set $E \subseteq X$ if and only if for all $y \in X$ there exists a $\varepsilon > 0$ such that for all $t \in \mathbb{R}, |t| \leq \varepsilon$ we have $x + ty \in E$.

let C be a convex subset of a real vector space X with 0 as algebraic interior point. The *Minkowski functional* of C is defined for all $x \in C$

$$p_C(x) := \inf \left\{ t \geq 0 : \frac{x}{t} \in C \right\}.$$

It allows to translate statements in functional analysis to convex geometry (and vice versa).

It is useful to think about linear functionals $f : X \to \mathbb{R}$ as hyperplanes: for some fixed $c \in \mathbb{R}$, $H_f := \{x \in X : f(x) = c\}$ defines a hyperplane. Using the previously introduced notions and Theorem 4.1.1, one can obtain the following geometric version of the Hahn-Banach theorem (see p. 54 in [55]).

Theorem 4.1.2 (Geometric Hahn-Banach theorem)**.** *Let A and B be disjoint non-empty convex subsets of a real vector space X. If A contains an algebraic interior point, then A and B can be separated by a hyperplane, i.e. there is a non-zero linear functional $h \in X'$ and a $c \in \mathbb{R}$ such that*

$$h(x) \leq c \leq h(y) \qquad \text{for all } x \in A \text{ and } y \in B.$$

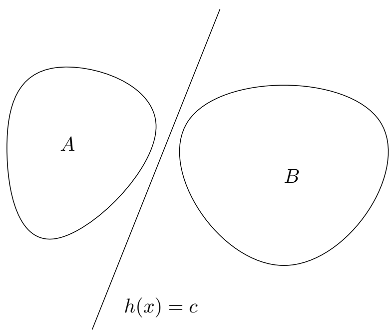


Figure 4.1: Geometric Hahn-Banach theorem

4.2 Type and Cotype

In this section we introduce the notions of type and cotype that are intimately linked with the geometry of Banach spaces. The mathematical material in this section can be found in many good textbooks such as [56, 57, 58]. We refer the reader to them for a more detailed exposition. We denote by $\mathbb{E}$ the expectation value of a random variable.

The simplest and most familiar Banach space is the Banach space in which the norm is induced by an inner product, namely the *Hilbert space*. It was soon realized [59] by two of the (mathematical) fathers of quantum theory that a Hilbert space is the only Banach space where the norm satisfies the Parallelogram identity

$$\frac{\|x + y\|^2}{2} + \frac{\|x - y\|^2}{2} = \|x\|^2 + \|y\|^2 . \tag{4.1}$$

A real *Rademacher variable* r is a uniformly distributed random variable taking values in the set $\{-1, 1\}$. Eq. (4.1) can be generalized for arbitrary Hilbert spaces. Indeed, consider a finite sequence $(x_j)_{j=1}^n \in \mathcal{H}$ and a sequence of independent Rademacher variables $(r_j)_{j=1}^n$. Then we have

$$\begin{aligned}
\mathbb{E}\left\|\sum_j^n r_j x_j\right\|^2 &= \mathbb{E}\langle \sum_j^n r_j x_j, \sum_k^n r_k x_k \rangle, \\
&= \sum_{j,k}^n \langle x_j, x_k \rangle \mathbb{E}(r_j r_k), \\
&= \sum_{j,k}^n \langle x_j, x_k \rangle \delta_{jk} = \sum_{j=1}^n \|x_j\|^2 .
\end{aligned}$$

The notions of type and cotype study how the average $\left(\mathbb{E}\left\|\sum_j^n r_j x_j\right\|^p\right)^{1/p}$ behaves for general Banach spaces. They were introduced by Hoffmann-Jørgensen [60] and developed by Maurey and Pisier in the 1970s [61, 62]. Let X be a Banach space and let $p \in [1,2]$ and $q \in [2,\infty)$. For every positive interger n we define $T_{p,n}(X)$ and $C_{q,n}$ to be the smallest constants such that for arbitrary finite sequences $(x_j)_{j=1}^n \subset X$, we have

$$\left(\mathbb{E}\left\|\sum_{j=1}^n r_j x_j\right\|^2\right)^{1/2} \leq T_{p,n}(X)\left(\sum_{j=1}^n \|x_j\|^p\right)^{1/p},$$

$$C_{p,n}(X)^{-1}\left(\sum_{j=1}^n \|x_j\|^q\right)^{1/q} \leq \left(\mathbb{E}\left\|\sum_{j=1}^n r_j x_j\right\|^2\right)^{1/2}.$$

The space X is said to be of *(Rademacher) type p* if $T_p(X) = \sup_n T_{p,n}(X) < \infty$. Similarly, X is said to have *(Rademacher) cotype q* if $C_q(X) = \sup_n C_{q,n}(X) < \infty$. The quantities $T_p(X)$ and $C_q(X)$ are called the (Rademacher) type p and cotype q constants of X, respectively. By replacing the Rademacher sums by Gaussian sums of elements of X we obtain the related notion of Gaussian type and cotype.

If X has type p, then X has type k for $k < p$. Thus, the type 2 and cotype 2 constant play an important role. Let $\alpha > 1$ be a real constant. We say that X contains a *$\alpha-$isomorphic copy of Y* and write $Y \subseteq_\alpha X$ if there exists a linear map $T : X \to Y$ and constanst α_1, α_2 such that $\alpha_1\alpha_2 \leq \alpha$ and

$$\frac{1}{\alpha_1}\|y\| \leq \|Ty\| \leq \alpha_2\|y\|, \qquad y \in Y.$$

The type and cotype constants are isomorphic invariants and are inherited by subspaces. More generally, if $Y \subseteq_\alpha X$, then $T_p(Y) \leq \alpha T_p(X)$ and $C_q(Y) \leq \alpha C_q(X)$.

The following theorem of Maurey comes from a clever use of Theorem 4.1.1. It is a powerful tool in Banach space theory and will be use in section A.2.

Theorem 4.2.1 (Maurey extension theorem)**.** *Let X and Y be real Banach spaces. Suppose E is a closed linear subspace of X and $T : E \to Y$ is a linear operator. If X has type 2 and*

E has cotype 2, then there exists a Hilbert space $\mathcal{H}$ and linear operators $R_1 : X \to \mathcal{H}$ and $R_2 : \mathcal{H} \to Y$ with $\|R_1\| \, \|R_2\| \leq C_2(Y)T_2(X) \, \|T\|$ such that $T = R_2 \circ R_1|_E$. In other words, T has a continuous linear extension $\tilde{T} = R_2 \circ R_1$ such that

$$\left\|\tilde{T}\right\| \leq C_2(Y)T_2(X) \, \|T\| \, .$$

In particular, if $X = Y$ and T is the identity map, then there exists a projection $P : X \to E$ with $\|P\| \leq C_2(Y)T_2(X)$.

We remark that Maurey's extension theorem is valid for both notions of type: Rademacher and Gaussian. Moreover, the norm in the statement is the operator norm. The operator norm of projections between Hilbert spaces is always one. This is not longer true for general Banach spaces.

$$\begin{tikzcd}
X \arrow[r, "R_1"] & \mathcal{H} = R_1(E) \oplus R_1(E)^{\perp} \arrow[d, "\pi"] & \\
E \arrow[u, "j"] \arrow[r, "R_1"] & R_1(E) \arrow[r, "R_2"] & Y
\end{tikzcd}$$

Figure 4.2: Maurey extension Theorem. The map $T : E \to Y$ is extended to X by factorizing through a Hilbert space $\mathcal{H}$. Here $\pi : \mathcal{H} \to R_1(E)$ is the orthogonal projection of a Hilbert space into a closed subspace, i.e. into a Hilbert subspace.

$$\begin{tikzcd}
X \arrow[r, "R_1"] \arrow[d, bend right, "P = R_1^{-1} \circ \pi \circ R_1"'] & \mathcal{H} = R_1(E) \oplus R_1(E)^{\perp} \arrow[d, "\pi"] \\
E \arrow[u, "j"'] & R_1(E) \arrow[l, "R_1^{-1}"']
\end{tikzcd}$$

Figure 4.3: A projection can be considered as an extension of the identity map. Factorization through a Hilbert space in the case $T = \mathrm{id}$.

The next important characterization of Hilbert spaces follows from Maurey's extension theorem:

Theorem 4.2.2 (Kwapien [63])**.** *For a Banach space X the following assertions are equivalent*

(i) X has type 2 and cotype 2;

(ii) X is isomorphic to a Hilbert space $\mathcal{H}$.

If these equivalent conditions are satisfied, an isomorphism $\Phi : X \to \mathcal{H}$ can be constructed such that

$$\|\Phi\| \left\|\Phi^{-1}\right\| \leq T_2(X)C_2(X).$$

Explicit type and cotype constants of finite dimensional l_p-spaces can be found in [57], Proposition 7.1.7. In particular, l_∞ does not have finite type. The Rademacher type and cotype constants of the Schatten classes were first computed by Tomczak-Jaegermann [58]; they behave similarly as the commutative l_p-spaces.

4.2.1 Proof of Maurey's extension theorem

Albiac and Kalton [56] (Theorem 7.3.4) have considerably simplified the proof of Maurey's extension theorem. Their modern proof can be also found in [57] (Theorem 7.3.2) where a small typo is corrected. We provide here their proofs for the ease of the reader. This theorem will be used in the contributed article of section A.2 in order to obtain a better upper bound of the linear stability of Wigner's theorem. We remark that the conclusion of Theorem 4.2.1 does not depend on which definition of type and cotype we use, i.e. Rademacher or Gaussian. In short, we can say that the latter fact is a consequence of the central limit theorem (see Theorem 7.4.4 in [56]). Furthermore, Maurey extension theorem works also for complex Banach spaces, but we do not present this here as we will be working with the Hermitian part of the Schatten classes.

The following Lemma has a crucial role in the proof of Maurey's extension theorem. It is a beautiful use of the functional version of the Hahn-Banach theorem 4.1.1. We postpone its proof until the end of this subsection.

Lemma 4.2.3. *Let V be a real vector space and $\mathcal{A}, \mathcal{B}$ two subsets of V such that*

$$V = cone(\mathcal{B}) - cone(\mathcal{A}),$$

and two functions $f : \mathcal{A} \to \mathbb{R}$ and $g : \mathcal{B} \to \mathbb{R}$. Then the following are equivalent:

(i) There is a linear functional Φ on V such that

$$\begin{aligned} f(a) \leq &\Phi(a) && \text{for all} && a \in \mathcal{A}, \\ &\Phi(b) \leq g(b) && \text{for all} && b \in \mathcal{B}. \end{aligned}$$

(ii) If $(\alpha_i)_{i=1}^m, (\beta_j)_{j=1}^m$ are two sequences of non-negative scalars such that

$$\sum_{i=1}^m \alpha_i a_i = \sum_{j=1}^n \beta_j b_j,$$

for some $(a_i)_{i=1}^m \subset \mathcal{A}, (b_j)_{j=1}^n \subset \mathcal{B}$. Then

$$\sum_{i=1}^m \alpha_i f(a_i) \leq \sum_{j=1}^n \beta_j g(b_j).$$

We also need the following lemma whose proof can be found in [56] (Lemma 7.4.3.) or in section 6.1.d of [57].

Lemma 4.2.4 (Covariance domination)**.** *Let $(\gamma_j)_{j=1}^\infty$ be a sequence of i.i.d. standard Gaussian random variables and $x_1, \dots, x_n$ and $y_1, \dots, y_m$ be elements of a Banach space X satisfying*

$$\sum_{j=1}^m |x^*(y_j)|^2 \leq \sum_{i=1}^n |x^*(x_i)|^2 \qquad \text{for all } x^* \in X^*,$$

Then, for all $1 \leq p < \infty$,

$$\mathbb{E} \left\| \sum_{j=1}^m \gamma_j y_j \right\|^p \leq \mathbb{E} \left\| \sum_{i=1}^n \gamma_i x_i \right\|^p.$$

The name of Lemma 4.2.4 stems from the fact that one can define a Gaussian random variable for a Banach space X. An $X-$valued random variable z is Gaussian if the real-valued random variable $x^*(z)$ is a real Gaussian random variable for all $x^* \in X^*$. This can be also understood in terms of characteristic functions (see Appendix E.1.c in [57]).

Proof of Theorem 4.2.1. Let $\mathcal{F}(X^*)$ denote the set of all functions from X^* to $\mathbb{R}$ and consider the map $X \to \mathcal{F}(X^*)$ given by $x \mapsto \hat{x}$, where $\hat{x}$ is the evaluation functional, i.e. $\hat{x}(x^*) = x^*(x)$ for all $x^* \in X^*$. Let V be the linear subspace of $\mathcal{F}(X^*)$ of all finite linear combinations of functions of the form $\hat{x}\hat{z}$ with $x, z \in X$. That is

$$V := \left\{ \sum_{k=1}^{N} \lambda_k \hat{x}_k \hat{z}_k : (\lambda_k)_{k=1}^N \in \mathbb{R}, (x_k)_{k=1}^N \text{ and } (z_k)_{k=1}^N \text{ in } X \text{ and } N \in \mathbb{N} \right\}.$$

The set V is actually generated by the subsets $\mathcal{A} = \mathcal{B} = \{\hat{x}^2 \in V : \ x \in X\}$ in the sense that

$$V = \operatorname{cone}(\mathcal{A}) - \operatorname{cone}(\mathcal{B}).$$

Indeed, by linearity it suffices to consider the element $\lambda \hat{x}\hat{z}$ with $x, z \in X$. Now, since these are linear functionals we have by polarization

$$\hat{x}\hat{z} = \frac{1}{4}\left((\hat{x} + \hat{z})^2 - (\hat{x} - \hat{z})^2 \right),$$

so any element in V can be written as claimed. We intend to apply Lemma 4.2.3 to construct a linear functional Φ on V such that

$$0 \le \Phi(\hat{x}^2) \le T_2^2(X) C_2^2(E) \|T\|^2 \|x\|^2, \qquad x \in X$$

and $\|Tx\|^2 \le \Phi(\hat{x}^2)$ for all $x \in E$. For that matters, let $\mathcal{A} = \mathcal{B} = \left\{\hat{x}^2 : \ x \in X\right\}$ with

$$\begin{aligned} f(\hat{x}^2) &:= \begin{cases} 0 & x \in X \setminus E, \\ \|Tx\|^2 & x \in E, \end{cases} \\ g(\hat{x}^2) &:= (\|T\| \, C_2(X) T_2(X))^2 \|x\|^2 . \end{aligned}$$

Assume $\sum_{j=1}^{n} \beta_j^2 \hat{z}_j^2 = \sum_{i=1}^{m} \alpha_i^2 \hat{x}_i^2$ for some $(x_i)_{i=1}^m, (z_j)_{j=1}^n \in X$, and some real scalars $(\alpha_i)_{i=1}^m, (\beta_j)_{j=1}^m$. Without lost of generality, suppose $z_1, \dots, z_l \in E$ and $z_{l+1}, \dots, z_n \in X \setminus E$. Then, since $\beta^2 \hat{z}^2, \alpha^2 \hat{x}^2$ are positive functionals

$$\sum_{j=1}^{l} \beta_j^2 \hat{z}_j^2 \le \sum_{i=1}^{m} \alpha_i^2 \hat{x}_i^2 .$$

From the covariance domination principle for $p = 2$, Lemma 4.2.4, and the definitions of type and cotype 2 we obtain

$$\begin{aligned}\sum_{j=1}^{l} \|T(\beta_j z_j)\|^2 &\leq \|T\|^2 \sum_{j=1}^{l} \|\beta_j z_j\|^2 ,\\ &\leq \|T\|^2 C_2^2(E)\, \mathbb{E} \left\| \sum_{j=1}^{l} \gamma_j \beta_j z_j \right\|^2 ,\\ &\leq \|T\|^2 C_2^2(E)\, \mathbb{E} \left\| \sum_{i=1}^{m} \gamma_j \alpha_i x_i \right\|^2 ,\\ &\leq (\|T\|\, T_2(X) C_2(E))^2 \sum_{i=1}^{m} \|\alpha_i x_i\|^2 ,\end{aligned}$$

which by definition of f and g is the inequality

$$\sum_{j=1}^{l} \beta_j^2 f(\hat{z}_j^2) \leq (\|T\|\, T_2(X) C_2(E))^2 \sum_{i=1}^{m} \alpha_i^2 g(\hat{x}_i^2).$$

Therefore, from Lemma 4.2.3 there exists a linear functional $\Psi : V \to \mathbb{R}$ with

$$f(\hat{x}^2) \leq \Psi(\hat{x}^2) \leq g(\hat{x}^2), \qquad \text{for all } x \in X.$$

Define a bilinear form on X by $(x, y) := \Psi(\hat{x}\hat{y})$ and consider the seminorm $p(x) = \sqrt{\Psi(\hat{x}^2)}$. Let $\mathcal{N} := \{x \in X : p(x) = 0\}$ and consider the real vector space $X_0 := X/\mathcal{N}$. Let $\mathcal{H}$ be the completion of X_0 with respect to the norm p. Then the induced bilinear form on X_0 has a unique continuous extension to $\mathcal{H}$. This makes $\mathcal{H}$ a real Hilbert space. We denote the norm of $\mathcal{H}$ by $\|\cdot\|_{\mathcal{H}}$.

Let $R_1 : X \to \mathcal{H}$ be defined by $x \mapsto [x]$ where $[x]$ is the equivalence class module $\mathcal{N}$. That is $[x] \sim [y]$ if and only if $x - y = z$ with $p(z) = 0$. Thus we have that

$$\|R_1 x\|_{\mathcal{H}} = \|[x]\|_{\mathcal{H}} = p(x) = \sqrt{\Psi(\hat{x}^2)} \leq \sqrt{g(\hat{x}^2)} = \|T\|\, T_2(X) C_2(E)\, \|x\| .$$

Now let $R_2 : R_1(E) \to Y$ be defined by $R_2(R_1 x) = T(x)$. Then

$$\|R_2(R_1 x)\| = \|T(x)\| = \sqrt{f(\hat{x}^2)} \leq p(x) = \|[x]\|_{\mathcal{H}} , \qquad \text{for all } x \in E.$$

Therefore, R_2 is well defined on R_1 and $\|R_2\| \leq 1$. Finally, by the projection Theorem of Hilbert spaces (Theorem 11.3 in [9]) every $x \in \mathcal{H}$ can be written uniquely as $x = y + w$ where $y \in R_1(E)$ and $w \in R_1(E)^{\perp}$. Let us denote by π the orthogonal projection of $\mathcal{H}$ onto closure of $R_1(E)$. Then the extension $\tilde{T}$ of T is $\tilde{T} := R_2 \circ \pi \circ R_1$ (see Fig. 4.2) with $\left\|\tilde{T}\right\| \leq C_2(E) T_2(E) \|T\|$.

□

Proof of Lemma 4.2.3. (i) implies (ii) follows from the chain of inequalities

$$\begin{aligned}\sum_{i=1}^{m} \alpha_i f(a_i) \leq \sum_{i=1}^{m} \alpha_i \Phi(a_i) &= \Phi\left(\sum_{i=1}^{m} \alpha_i a_i\right), \\ &= \Phi\left(\sum_{j=1}^{n} \beta_j b_j\right), \\ &= \sum_{j=1}^{n} \beta_j \Phi\left(b_j\right) \leq \sum_{j=1}^{n} \beta_j g\left(b_j\right).\end{aligned}$$

We proceed to show that (ii) implies (i). In order to achieve this, define the functional $p : V \to [-\infty, \infty)$ by

$$p(v) := \inf \left\{ \sum_{j=1}^{n} \beta_j g\left(b_j\right) - \sum_{i=1}^{m} \alpha_i f(a_i) \right\}$$

where the infimum is taken over all the decompositions of $v = \sum_{j=1}^{n} \beta_j b_j - \sum_{i=1}^{m} \alpha_i a_i$ where $a_i \in \mathcal{A}, b_j \in \mathcal{B}$ and $\alpha_i, \beta_j \geq 0$. Note that since $V = \mathrm{cone}(\mathcal{A}) - \mathrm{cone}(\mathcal{B})$, the functional p is well-defined. We will show that p is sub-linear and then appeal to the functional version of the Hahn-Banach theorem. It is not difficult to check that p is sub-linear, i.e. that $p(\lambda v_1 + v_2) \leq \lambda p(v_1) + p(v_2)$ for all $\lambda > 0$ and $v_1, v_2 \in V$. In order to appeal to the functional version of the Hahn-Banach theorem we need to check that $p(v) > -\infty$ for every $v \in V$. We first show that $p(0) = 0$. Indeed, $p(0) \leq 0$ since $0 = 0b - 0a$ is valid decomposition of $0 \in V$. Moreover, if we represent 0 as

$$0 = \sum_{j=1}^{n} \beta_j b_j - \sum_{i=1}^{m} \alpha_i a_i,$$

then by (ii)

$$\sum_{i=1}^{m} \alpha_i f(a_i) \leq \sum_{j=1}^{n} \beta_j g\left(b_j\right),$$

so $p(0) \geq 0$. Now that we know that $p(0) = 0$, we find that $0 = p(0) \leq p(v) + p(-v)$. This implies that $p(v) > -\infty$ for every $v \in V$. Thus, from the first part of Lemma 4.1.1 with $E = \{0\}$ there exists a linear functional $\Phi : V \to \mathbb{R}$ such that

$$-p(-v) \leq \Phi(v) \leq p(v) \qquad \text{for all } v \in V.$$

Finally, we show that for $v \in \mathcal{A}$, $p(-v) \leq f(v)$, and for $v \in \mathcal{B}$, $p(v) \leq g(v)$. For the first inequality, consider for $-v$ a fixed $b \in \mathcal{B}$ and the decomposition $-v = 0b - 1v$. Then from the definition of p, $p(-v) \leq 0g(b) - 1f(v) = f(v)$. Likewise, consider $v = 1v - 0a$ to obtain the last claimed inequality. □

4.3 Twisted sums

Twisted sums are in correspondence with almost-linear maps and are therefore a useful tool for the study of the stability of almost-symmetries. The results presented in this section are mainly due to Kalton and Peck [64, 65]. We follow the presentation of [36, 66].

A *quasi-norm* on a real vector space X is a function $|||\cdot||| : X \to \mathbb{R}$ satisfying

(i) $|||x||| > 0$ for all $x \neq 0$,

(ii) $|||\lambda x||| = |\lambda| |||x|||$ for $x \in X$ and $\lambda \in \mathbb{R}$,

(iii) $|||x + y||| \leq K (|||x||| + |||y|||)$ where $K > 1$.

The constant K is called *modulus of convavity* and $(X, |||\cdot|||)$ a quasi-Banach space. A *twisted sum* of X and Y is a quasi-Banach space Z containing a subspace Y_0 isomorphic to Y and such that Z/Y_0 is isomorphic to X. A map $F : X \to Y$ between real normed spaces is called a *quasi-linear* map if it satisfies

(i) $F(\lambda x) = \lambda F(x)$ for $\lambda \in \mathbb{R}, x \in X$ and

(ii) there exists a constant $\delta \geq 0$ such that for all $x_1, x_2 \in X$

$$\|F(x_1 + x_2) - F(x_1) - F(x_2)\| \leq \delta (\|x_1\| + \|x_2\|) .$$

Kalton showed that there is a one-to-one correspondence between twisted sums and quasi-linear maps, i.e. twisted sums arise from and give rise to quasi-linear maps [65]. Given a quasi-linear map $F : X \to Y$ we can construct a quasi-norm on the vector space $Y \times X$ via

$$|||(y, x)|||_F := \|y - F(x)\|_Y + \|x\|_X .$$

Here the modulus of concavity is $1 + \delta$. The subspace $Y_0 := \{(y, 0) : y \in Y\}$ of Z is isometric to Y and the quotient Z/Y_0 is isometric to X. Then Z is a twisted sum. This correspondence suggest an alternative notation for the twisted sum Z generated by the quasi-linear $F : X \to Y$, namely $Z = Y \oplus_F X$ (the order of the spaces is important). The name "twisted" is basically due to the fact that the unit balls of twisted sums are twisted by the quasi-linear map F (see Figure 4.4).

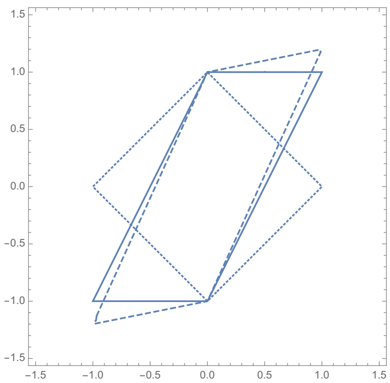


Figure 4.4: The figure shows the twisting induce by the quasi-linear map F on the unit ball of $\mathbb{R} \oplus_F \mathbb{R}$ with $|||(y, x)||| = |y - F(x)| + |x|$. The dotted lines correspond to $F = 0$; the straight lines to $F(x) = x$ and the dashed line to $F(x) = x + 0.02x$.

It was shown by Ribe and Kalton that there exist non-trivial twisted sums of Banach spaces which are not locally convex -thus the Hahn-Banach theorems do not work in these spaces- and which are not isomorphic to any direct sum [67, 64, 65]. An exposition of such twisted sums is beyond the scope of this introduction; we refer the reader to chapter 16 in [36] and the book [66]. Such singular twisted sums do not appear directly in our work since we restrict our attention to finite-dimensional spaces. However, the Ribe-Kalton-Peck twisted sums are important examples of possible obstructions that can appear in the study of quasi-linear maps between infinite-dimensional Banach spaces. In Section A.2, finite-dimensional twisted sums are used together with Theorem 4.2.1 in the case that $T = \mathrm{id}$. This will induce a linear map that approximates well an almost-symmetry.

Bibliography

[1] J. Cuesta and M. M. Wolf. Are almost-symmetries almost linear? *J. Math. Phys.*, 60(8), 2019.

[2] J. Cuesta. Type and cotype constants and the linear stability of Wigner's symmetry theorem. *Symmetry*, 11(9):1107, 2019.

[3] J. Cuesta. A stable quantum Darmois-Skitovich theorem. *To appear in J. Math. Phys.*, 2019.

[4] S. Weinberg. *The quantum theory of fields*, volume I, Foundations. Cambridge university press, 1995.

[5] T. Heinosaari and M. Ziman. *The Mathematical Language of Quantum Theory.* Cambridge University Press, 2012.

[6] A. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory.* Publications of the Scuola Normale Superiore, 2011.

[7] L. Ballentine. *Quantum Mechanics, a modern development.* World Scientific Publishing, 1998.

[8] M. M. Wolf. Quantum channels and operations. Lecture notes available at: http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf.

[9] M. Reed and B. Simon. *Methods of Modern Mathematical Physics.* Volume I: Functional Analysis. Academic Press, 1980.

[10] J. Watrous. *The Theory of Quantum Information.* Cambridge University Press, 2018.

[11] E. P. Wigner. *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren.* Fredrik Vieweg und Sohn, 1931.

[12] V. Bargmann. Note on Wigner's theorem on symmetry operation. *J. Math. Phys.*, (5), 1964.

[13] G. Gehér. An elementary proof for the non-bijective version of Wigner's theorem. *Phys. Lett. A*, (378):2054–2057, 2014.

[14] J. Barviňek and J. Hamhalter. Linear algebraic proof of Wigner theorem and its consequences. *Math. Slovaca*, 67(2), 2017.

[15] L. Molnár. Transformations on the set of all $n-$dimensional subspaces of a hilbert space preserving principal angles. *Commun. Math. Phys.*, 217, 2001.

[16] I. Herstein. Jordan homomorphisms. *Trans. Am. Math. Soc.*, 81, 1956.

[17] J. Hou. Rank preserving linear maps on $\mathcal{B}(x)$. *Sci. China Ser. A*, (32):29–940, 1989.

[18] M. D. Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975.

[19] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3(4):275–278, 1972.

[20] R. F. Werner and A. S. Holevo. Counterexample to an additivity conjecture for output purity of quantum channels. *J. Math. Phys.*, 43(9), 2002.

[21] R. F. Werner. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40(8):4277–4281, 1989.

[22] M. M. W. A. L. K. Hashagen. Universality and optimality in the information-disturbance tradeoff. *Annales Henri Poincar 'e*, 20(1):1–40, 2018.

[23] S. M. Ulam. *A Collection of Mathematical Problems.* Interscience, 1960.

[24] D. H. Hyers. On the stability of the linear functional equation. *Proc. Natl. Acad. Sci. USA*, (27), 1941.

[25] T. M. Rassias. On the stability of the linear mapping in banach spaces. *Proc. Amer. Math. Soc.*, (72):297–300, 1978.

[26] H. S. P. Pardalos, P. Georgiev, editor. *Nonlinear Analysis.* Springer, 2012.

[27] S.-M. Jung. *Hyers–Ulam–Rassias Stability of Functional Equations in Nonlinear Analysis*, volume 48 of *Springer Optimization and Its Applications.* Springer, 2011.

[28] M.-D. Choi. Almost commuting matrices need not be nearly commuting. *Proc. Amer. Math. Soc.*, (102), 1988.

[29] D. Voiculescu. Asymptotically commuting finite rank unitary operators without commuting approximants. *Acta. Sci. Math.*, (45):429–431, 1983.

[30] H. X. Lin. Almost commuting unitary elements in purely infinite simple C^*-algebras. *Math. Ann.*, 303(4):599–616, 1995.

[31] M. B. Hastings. Making almost commuting matrices commute. *Commun. Math. Phys.*, 291:321–345, 2009.

[32] D. Ilišević and A. Turnšek. Stability of the Wigner equation – a singular case. *J. Math. Anal. Appl.*, (429):273–287, 2015.

[33] J. Alaminos, J. Extremera, and A. R. Villena. Metric versions of hersteins theorems on jordan maps. *Linear Multilinear Algebra*, 60:941, 2012.

[34] N. J. Kalton. A remark on quasi-isometries. *Proc. A.M.S.*, 131(4), 2002.

[35] S. Mazur and S. Ulam. Sur les transformations isométriques d'espaces vectoriels norm 'es. *C.R.A.S. Paris*, 194, 1932.

[36] Y. Benyamini and J. Lindenstrauss. *Geometric Nonlinear Functional Analysis*, volume 1. A.M.S. Colloquium publications, 2000.

[37] C.-K. Li and N.-K. Tsing. Linear preserver problems: A brief introduction and some special techniques. *Linear Algebra Appl.*, 162-164, 1992.

[38] L. Molnár. *Selected Preserver Problems on Algebraic Structures of Linear Operators and on Function Spaces*, volume 1895 of *Lecture Notes in Mathematics*. Springer, 2007.

[39] L. Molnár, G. Nagy, and P. Szokol. Maps on density operators preserving quantum f-divergences. *Quantum Information Processing*, 12(7), 2013.

[40] A. R. R. Honegger. *Photons in Fock Space and Beyond*, volume Volume 1: From Classical to Quantized Radiation Systems. World Scientific, 2015.

[41] M. Keyl, J. Kiukas, and R. Werner. Schwartz operators. *Rev. Math. Phys.*, 28, 2016.

[42] A. W. Harrow, A. Montanaro, and A. J. Short. Limitations on quantum dimensionality reduction. In *Automata, Languages and Programming*, pages 86–97. Springer, 2011.

[43] J. Eisert and M. M. Wolf. *Quantum Information with Continuous Variables of Atoms and Light*, chapter 2, page 23–42. World Scientific, 2007.

[44] U. Leonhardt. *Measuring the Quantum State of Light*. Cambridge university press, 1997.

[45] C. Cushen and R. Hudson. A quantum-mechanical central limit theorem. *Journal of Applied Probability*, 8(3), 1971.

[46] M. M. Wolf, G. Giedke, and J. I. Cirac. Extremality of gaussian quantum states. *Phys. Rev. Lett.*, 96(8), 2006.

[47] M. M. Wolf, J. Eisert, and M. B. Plenio. Entangling power of passive optical elements. *Phys. Rev. Lett.*, (90), 2003.

[48] J. Maillard. On the twisted convolution product and the weyl transformation of tempered distributions. *Jour. Geom. Phys.*, 3(231), 1986.

[49] N. Meyers and J. Serrin. $H = W$. *Proc. Nat. Acad. Sci. U.S.A.*, 51, 1964.

[50] E. Lieb and M. Loss. *Analysis*. American Mathematical Society, ii edition, 2001.

[51] D. Kastler. The C^*-algebras of a free boson field. *Commun. math. Phys.*, 1:14–48, 1965.

[52] A. Holevo. Quantum characteristic functions. *Problems Inform. Transmission*, 6(4):316–321, 1970. (in Russian).

[53] B. Maurey. Un théorème de prolongement. *C. R. Acad. Sc. Paris*, 279:329–332, 1974.

[54] W. Rudin. *Functional Analysis*. McGraw-Hill, second edition edition, 1991.

[55] B. Bollobás. *Linear Analysis*. Cambridge University Press, second edition edition, 1999.

[56] F. Albiac and N. J. Kalton. *Topics in Banach space theory*, volume 233 of *Graduate Texts in Mathematics*. Springer, 2006.

[57] T. Hytönen, J. Neerven, M. Veraar, and L. Weis. *Analysis in Banach Spaces*, volume II: Probabilistic Methods and Operator Theory. Springer, 2017.

[58] N. Tomczak-Jaegermann. *Banach-Mazur distances and finite-dimensional operator ideals.* Longman Scientific & Technical, 1989.

[59] P. Jordan and J. V. Neumann. On inner products in linear, metric spaces. *Ann. Math.*, 36(3):719–723, 1935.

[60] J. Hoffmann-Jørgensen. Sums of independent banach space valued random variables. *Studia Math.*, 52:159–186., 1974.

[61] B. Maurey. Type, cotype and k-convexity. *Handbook of the Geometry of Banach Spaces*, 2:1299–1332., 2003.

[62] B. Maurey and G. Pisier. Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de banach. *Studia Math.*, 58:45–90, 1976. (in French).

[63] S. Kwapien. Isomorphic characterizations of inner product spaces by orthogonal series with vector coefficients. *Studia Math.*, 44:583, 1972.

[64] N. J. Kalton. The three space problem for locally bounded f spaces. *Compositio Math.*, 49:214–224, 1978.

[65] N. J. Kalton and N. T. Peck. Twisted sums of sequence spaces and the three space problem. *Trans. Amer. Math. Soc.*, 255:1–30, 1979.

[66] Castillo and M. González. *Three-space problems in Banach space theory.* Springer, 1997.

[67] M. Ribe. Examples for the nonlocally convex three space problem. *Proc. Amer. Math. Soc.*, 73, 1979.

A Core Articles

A.1 Are almost-symmetries almost linear?

Are almost-symmetries almost linear?

Javier Cuesta and Michael M. Wolf

Wigner's symmetry theorem is a foundational result in quantum theory that imposes how we can mathematically represent physical symmetries in quantum physics. Moreover, it is a strong advocate of the linearity of quantum mechanics. This theorem tells us that a transformation between pure states that preserves the probability amplitudes muss be necessarily *linear* and expressed as an inner automorphism of a unitary or anti-unitary map. It is a natural question to ask whether this theorem is stable in the case that the probability amplitudes are just almost preserved, i.e. if we consider an almost-symmetry. After a recent series of work an affirmative answer to this problem was obtained for finite dimensional Hilbert spaces. The proof basically consists on two non-trivial steps: first, the linear stability of an almost-symmetry and second, the stability of Herstein's theorem on Jordan maps. However, the obtained bound presented there depends in a rather complicated form on the dimension of the Hilbert space. In our work, we focus on the first part of the stability of Wigner's theorem and clarify the role of the dimension dependency. We show in Theorem 2 (*i*) that in infinite dimensional Hilbert spaces we can approximate an almost-symmetry by a linear map in a weak sense, i.e. we can approximate the inner product between the value of an almost-symmetry and a fix observable. Moreover, in Theorem 3 we show that in Hilbert spaces with large dimension there exist a non-linear almost-symmetry which cannot be approximated by any linear map. This in turn implies that the linear stability of Wigner's theorem cannot be independent of the dimension of the Hilbert space. On the other hand, we show in Lemma 2 that an almost-symmetry can be extended to an almost-linear map. This is cleverly used in the proof of Theorem 2 (*ii*) together with the geometric version of the Hahn-Banach theorem to prove an upper bound on the quality of approximation of an almost-symmetry in finite dimensional spaces. The latter bound depends linearly on the dimension of the Hilbert space.

I was significantly involved in finding the ideas and carrying out the scientific work of all parts of this article. Furthermore, I was in charge of writing the article and the Journal submission.

Permission to include:

Javier Cuesta and Michael M. Wolf
Are almost-symmetries almost linear?
J. Math. Phys. 60, 082101 (2019)

Permissions - AIP Publishing LLC

Reusing AIP Publishing Content

Permission from AIP Publishing is required to:

- republish content (e.g., excerpts, figures, tables) if you are not the author
- modify, adapt, or redraw materials for another publication
- systematically reproduce content
- store or distribute content electronically
- copy content for promotional purposes

To request permission to reuse AIP Publishing content, use RightsLink® for the fastest response or contact AIP Publishing directly at rights@aip.org and we will respond within one week:

For RightsLink, use Scitation to access the article you wish to license, and click on the Reprints and Permissions link under the TOOLS tab. (For assistance click the "Help" button in the top right corner of the RightsLink page.)

To send a permission request to rights@aip.org, please include the following:

- Citation information for the article containing the material you wish to reuse
- A description of the material you wish to reuse, including figure and/or table numbers
- The title, authors, name of the publisher, and expected publication date of the new work
- The format(s) the new work will appear in (e.g., print, electronic, CD-ROM)
- How the new work will be distributed and whether it will be offered for sale

Authors do **not** need permission from AIP Publishing to:

- quote from a publication (please include the material in quotation marks and provide the customary acknowledgment of the source)
- reuse any materials that are licensed under a Creative Commons CC BY license (please format your credit line: "Author names, Journal Titles, Vol.#, Article ID#, Year of Publication; licensed under a Creative Commons Attribution (CC BY) license.")
- reuse your own AIP Publishing article in your thesis or dissertation (please format your credit line: "Reproduced from [FULL CITATION], with the permission of AIP Publishing")
- reuse content that appears in an AIP Publishing journal for republication in another AIP Publishing journal (please format your credit line: "Reproduced from [FULL CITATION], with the permission of AIP Publishing")
- make multiple copies of articles–although you must contact the Copyright Clearance Center (CCC) at www.copyright.com to do this

Reusing Content Published by Others

To request another publisher's permission to reuse material in AIP Publishing articles, please use our Reuse of Previously Published Material form. (We require documented permission for all reused content.)

Reuse of Previously Published Material Form (pdf)

Unless the publisher requires a specific credit line, please format yours like this:

Reproduced with permission from J. Org. Chem. 63, 99 (1998). Copyright 1998, American Chemical Society.

You do not need permission to reuse material in the public domain, but you should still include an appropriate credit line which cites the original source.

Journal of Mathematical Physics

Are almost-symmetries almost linear?

Cite as: J. Math. Phys. **60**, 082101 (2019); https://doi.org/10.1063/1.5087539
Submitted: 02 January 2019 . Accepted: 14 July 2019 . Published Online: 01 August 2019

Javier Cuesta, and Michael M. Wolf

Journal of Mathematical Physics | ARTICLE | scitation.org/journal/jmp

Are almost-symmetries almost linear?

Cite as: J. Math. Phys. 60, 082101 (2019); doi: 10.1063/1.5087539
Submitted: 2 January 2019 • Accepted: 14 July 2019 •
Published Online: 1 August 2019

View Online Export Citation CrossMark

Javier Cuesta[a)] and Michael M. Wolf[b)]

AFFILIATIONS

Department of Mathematics, Technische Universität München, 85748 Garching, Germany and Munich Center for Quantum Science and Technology (MCQST), München, Germany

[a)] j.cuesta@tum.de
[b)] m.wolf@tum.de

ABSTRACT

It d-pends. Wigner's symmetry theorem implies that transformations that preserve transition probabilities of pure quantum states are linear maps on the level of density operators. We investigate the stability of this implication. On the one hand, we show that any transformation that preserves transition probabilities up to an additive ε in a separable Hilbert space admits a weak linear approximation, i.e., one relative to any fixed observable. This implies the existence of a linear approximation that is $4\sqrt{\varepsilon}d$–close in Hilbert-Schmidt norm, with d the Hilbert space dimension. On the other hand, we prove that a linear approximation that is close in norm and independent of d does not exist in general. To this end, we provide a lower bound that depends logarithmically on d.

I. INTRODUCTION AND SUMMARY OF RESULTS

Wigner's theorem[1] is a cornerstone for the mathematical representation of symmetries in quantum physics. It tells us that an arbitrary transformation on the set of pure states that preserves the "transition probabilities" must necessarily correspond to a unitary or antiunitary operation. In particular, the transformation is representable by a *linear map* on the space spanned by the density operators. Hence, Wigner's theorem is arguably one of the reasons for the linear structure of quantum theory (besides various forms of locality[2,3] and the probabilistic framework[4]).

Wigner's theorem was proven in the general case, which does not assume bijectivity of the map, by Bargmann[5] 30 years after Wigner's original idea. Recently,[6,7] new and simpler proofs of this theorem have appeared where neither bijectivity of the map nor separability of the underlying complex Hilbert space $\mathcal{H}$ is assumed. If we denote by $\mathbb{P}(\mathcal{H})$ the set of pure states, identified with rank-one self-adjoint projections, Wigner's theorem reads as follows:

Theorem 1 (Wigner). *Let $f : \mathbb{P}(\mathcal{H}) \to \mathbb{P}(\mathcal{H})$ be a map that preserves transition probabilities, i.e., such that* $\mathrm{Tr}\, f(X)f(Y) = \mathrm{Tr}XY$ *for all $X, Y \in \mathbb{P}(\mathcal{H})$. Then, there exists a linear or antilinear isometry $U : \mathcal{H} \to \mathcal{H}$ such that $f(X) = UXU^\dagger$.*

The theorem can be seen to establish two things: linearity and isometry. In the present work, we focus on the stability of the linearity property. That is, we address the question: if a map between pure states almost preserves transition probabilities, how well can it be approximated by a linear map?

Recently, a sequence of works[8–10] culminated in the result that Wigner's theorem is stable for finite-dimensional Hilbert spaces. We want to shed new light at least on the linear part of the problem and investigate, in particular, the role of the dimension of the underlying Hilbert space. To this end, we follow a different route than[8,10] and employ (convex) geometry rather than analysis of operator algebras for the main argument. Our results are twofold. On the one hand, any map that approximately satisfies Wigner's symmetry condition in any separable Hilbert space is shown to admit a weak linear approximation. That is, when evaluated through an arbitrary but fixed observable, there exists a linear approximation even in case of infinite dimensional Hilbert spaces. As a corollary, we obtain a linear approximation in Hilbert-Schmidt-norm whose approximation error is bounded linearly in terms of the Hilbert space dimension. This improves on the corresponding result of Ref. 8.

In the second part, we address the problem from the other end and prove that a linear approximation in norm does not always exist in infinite dimensions—not even with respect to the operator norm. For that purpose, we study a componentwise logarithmic spiral map and prove that its operator norm distance to the set of linear maps scales essentially logarithmic with the dimension of the Hilbert space. This holds despite the fact that the action of the map is arbitrarily close to that of a symmetry in Wigner's sense.

II. PRELIMINARIES

We now introduce some notation and definitions. We denote by $\mathcal{H}$ a complex Hilbert space, which we assume to be separable in the following. The space of bounded linear operators on $\mathcal{H}$ is denoted by $\mathfrak{B}(\mathcal{H})$ and its identity element by $\mathbb{1}$. The adjoint of an operator X is written as X^*. For $p \in [1, \infty)$, we denote by $\mathcal{T}_p(\mathcal{H}) := \{X \in \mathfrak{B}(\mathcal{H}) | X = X^*, \|X\|_p := (\mathrm{Tr}|X|^p)^{1/p} < \infty\}$ the real Banach space known as the hermitian p-Schatten class and its respective unit ball by $\mathcal{B}_p(\mathcal{H}) := \{X \in \mathcal{T}_p(\mathcal{H}) | \|X\|_p \le 1\}$. $\|\cdot\|_\infty$ will be the operator norm on $\mathfrak{B}(\mathcal{H})$.

Occasionally, we will make use of the Dirac "bra-ket" notation where a vector in $\mathcal{H}$ is written as $|x\rangle$ and the scalar product of two vectors $|x\rangle$, $|y\rangle$ as $\langle x|y\rangle$. A rank-one projection in $\mathfrak{B}(\mathcal{H})$ with range spanned by a unit vector $|x\rangle$ is then $|x\rangle\langle x|$. Using this, we define $\mathbb{P}(\mathcal{H}) := \{|\psi\rangle\langle\psi| \mid |\psi\rangle \in \mathcal{H}, \|\psi\|_2 = 1\}$, which is the set of pure quantum states written as density operators.

III. ALMOST-SYMMETRIES ARE CLOSE TO LINEAR

The main result of this section is summarized in the following theorem:

Theorem 2 (Linear approximation of almost-symmetries). *Let $\mathcal{H}$ be a separable complex Hilbert space and $f : \mathbb{P}(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})$ a map satisfying*

$$|\mathrm{Tr} f(X)f(Y) - \mathrm{Tr}XY| \le \varepsilon \quad \textit{for all } X, Y \in \mathbb{P}(\mathcal{H}). \tag{1}$$

(i) *For any $A \in \mathcal{B}_2(\mathcal{H})$, there is a linear map $T_A : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})$ such that for all $X \in \mathbb{P}(\mathcal{H})$,*

$$\left|\mathrm{Tr}\left[A\left(f(X) - T_A(X)\right)\right]\right| \le 4\sqrt{\varepsilon}.$$

(ii) *If $\mathcal{H} = \mathbb{C}^d$, there exists a linear map $T : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})$ such that for all $X \in \mathbb{P}(\mathbb{C}^d)$,*

$$\|f(X) - T(X)\|_2 \le 4\, d\sqrt{\varepsilon}.$$

Note that Eq. (1) is a slight relaxation of Wigner's condition since we allow f to map into $\mathcal{T}_2(\mathcal{H})$. That is, we do not restrict its range to the set of pure states. We will see that this generalization comes at no additional cost in the proof.

The overall strategy of the proof is the following: we first extend f to a map F that is defined on the entire space $\mathcal{T}_1(\mathcal{H})$. Exploiting the condition in Eq. (1), we will show that F almost preserves convex combinations. This enables the use of convex analysis and, in particular, of the geometric Hahn-Banach theorem to prove the existence of a linear approximation as stated in part (i) of the theorem. Part (ii) is then derived as a consequence of (i) by exploiting the existence of a finite basis.

We begin the proof of the theorem by extending the function f to a larger domain. To this end, we choose a spectral decomposition $X = \sum_k \lambda_k X_k$ for every $X \in \mathcal{T}_1(\mathcal{H})$ for which $\|X\|_1 = 1$. Here, $X_k \in \mathbb{P}(\mathcal{H})$ are assumed to be orthogonal with respect to the Hilbert-Schmidt inner product $\langle A, B\rangle := \mathrm{Tr}A^*B$. There might be more than one spectral decomposition; however, we just need to work consistently with a fixed choice. Then, we define $F : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})$ by extending

$$F(X) := \sum_k \lambda_k f(X_k) \tag{2}$$

in a homogeneous way from the unit sphere to the entire space $\mathcal{T}_1(\mathcal{H})$. Equation (1) then ensures that $\|F(X)\|_2^2 \le \|X\|_2^2 + \varepsilon\|X\|_1^2$ so that, indeed, $F(X) \in \mathcal{T}_2(\mathcal{H})$. By construction, f is then the restriction of F to the set $\mathbb{P}(\mathcal{H})$ of pure states and $F(\lambda X) = \lambda F(X)$ for all $\lambda \ge 0$. Moreover, Wigner's condition from Eq. (1) easily extends to F.

Lemma 1 (Wigner-condition for the extended map). *Let $f : \mathbb{P}(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})$ satisfy Eq. (1) and $F : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})$ be its extension as defined above. Then,*

$$|\langle F(X), F(Y)\rangle - \langle X, Y\rangle| \le \|X\|_1 \|Y\|_1 \varepsilon, \qquad \textit{for all } X, Y \in \mathcal{T}_1(\mathcal{H}). \tag{3}$$

Proof. Let $X = \sum_k \lambda_k X_k$ and $Y = \sum_j \mu_j Y_j$ be the spectral decompositions that define F on X, Y and recall that for elements of $\mathcal{T}_1(\mathcal{H})$ the trace-norm $\|\cdot\|_1$ is the sum of the absolute values of eigenvalues. The Lemma then follows from applying Eq. (1) to:

$$\left|\langle \sum_k \lambda_k f(X_k), \sum_j \mu_j f(Y_k)\rangle - \langle X, Y\rangle\right| \le \sum_{j,k} |\lambda_k\|\mu_j||\langle f(X_k), f(Y_j)\rangle - \langle X_k, Y_j\rangle|.$$

□

From here, we can show that F is almost-linear in the following sense:

Lemma 2 (Almost-linearity of the extended map). *Let $F : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})$ be any map satisfying Eq. (3). Then, for all $m \in \mathbb{N}, \lambda \in \mathbb{R}^m$ and $X_1, \ldots, X_m \in \mathcal{T}_1(\mathcal{H})$, we have*

$$\left\|\sum_{i=1}^m \lambda_i F(X_i) - F\left(\sum_{i=1}^m \lambda_i X_i\right)\right\|_2 \le 2\sqrt{\varepsilon}\sum_{i=1}^m |\lambda_i|\|X_i\|_1.$$

Proof. Consider $Z \in \mathcal{T}_1(\mathcal{H})$ and use Eq. (3) to bound

$$\begin{aligned}
&\left|\langle \sum_{i=1}^m \lambda_i F(X_i) - F\left(\sum_{i=1}^m \lambda_i X_i\right), F(Z)\rangle\right| \\
&\le \left|\sum_{i=1}^m \lambda_i(\langle F(X_i), F(Z)\rangle - \langle X_i, Z\rangle)\right| + \left|\langle F\left(\sum_{i=1}^m \lambda_i X_i\right), F(Z)\rangle - \langle \sum_{i=1}^m \lambda_i X_i, Z\rangle\right|, \\
&\le \left(\sum_{i=1}^m \|\lambda_i X_i\|_1 + \left\|\sum_{i=1}^m \lambda_i X_i\right\|_1\right)\|Z\|_1 \varepsilon \ \le\ 2\varepsilon\|Z\|_1 \sum_{i=1}^m \|\lambda_i X_i\|_1.
\end{aligned}$$

Then, by linearity of the Hilbert-Schmidt inner product

$$\begin{aligned}
\left\|\sum_{i=1}^m \lambda_i F(X_i) - F\left(\sum_{i=1}^m \lambda_i X_i\right)\right\|_2^2 &\le \left|\langle \sum_{i=1}^m \lambda_i F(X_i) - F\left(\sum_{i=1}^m \lambda_i X_i\right), \sum_{i=1}^m \lambda_i F(X_i)\rangle\right| \\
&\quad + \left|\langle \sum_{i=1}^m \lambda_i F(X_i) - F\left(\sum_{i=1}^m \lambda_i X_i\right), F\left(\sum_{i=1}^m \lambda_i X_i\right)\rangle\right|, \\
&\le 2\left(\sum_{i=1}^m \|\lambda_i X_i\|_1\right)^2 \varepsilon + 2\sum_{i=1}^m \|\lambda_i X_i\|_1 \left\|\sum_{i=1}^m \lambda_i X_i\right\|_1 \varepsilon, \\
&\le \left(2\sum_{i=1}^m \|\lambda_i X_i\|_1\right)^2 \varepsilon.
\end{aligned}$$

□

Now, we have all prerequisites for the proof of the main theorem.

Proof of Theorem 2. To show part (i), we define $\delta := 2\sqrt{\varepsilon}$ and consider the action of $\hat{F}(X) := \mathrm{Tr}[AF(X)]$ on the unit-ball $\mathcal{B}_1(\mathcal{H})$. If $\|\lambda\|_1 \le 1, \|A\|_2 \le 1$ and $\|X_j\|_1 \le 1$, then Cauchy-Schwarz and Lemma 1 imply

$$\begin{aligned}
\left|\hat{F}\left(\sum_{j=1}^m \lambda_j X_j\right) - \sum_{j=1}^m \lambda_j \hat{F}(X_j)\right| &= \left|\mathrm{Tr} A\left(\sum_{j=1}^m \lambda_j F(X_j) - F\left(\sum_{j=1}^m \lambda_j X_j\right)\right)\right| \\
&\le \|A\|_2 \left\|\sum_{j=1}^m \lambda_j F(X_j) - F\left(\sum_{j=1}^m \lambda_j X_j\right)\right\|_2 \ \le\ \delta.
\end{aligned}$$

Thus,

$$\sum_{j=1}^m \lambda_j \hat{F}(X_j) - \delta \le \hat{F}\left(\sum_{j=1}^m \lambda_j X_j\right) \le \sum_{j=1}^m \lambda_j \hat{F}(X_j) + \delta. \tag{4}$$

Let g_- and g_+ be the convex and concave envelopes of $\hat{F}$ over $\mathcal{B}_1(\mathcal{H})$. These are defined as

$$g_-(X) := \inf\left\{\sum_{j=1}^n \lambda_j \hat{F}(X_j) \;\middle|\; X = \sum_{j=1}^n \lambda_j X_j\right\},$$
$$g_+(X) := \sup\left\{\sum_{j=1}^n \lambda_j \hat{F}(X_j) \;\middle|\; X = \sum_{j=1}^n \lambda_j X_j\right\},$$

taken over all finite convex decompositions of X within the unit-ball $\mathcal{B}_1(\mathcal{H})$. Using Eq. (4), one verifies

$$g_+(X) - \delta \le \hat{F}(X) \le g_-(X) + \delta \quad \text{for all } X \in \mathcal{B}_1(\mathcal{H}). \tag{5}$$

Let Λ_+ and Λ_- denote the subgraph of $X \mapsto g_+(X) - \delta$ and the supergraph of $X \mapsto g_-(X) + \delta$, respectively. Since g_- and $-g_+$ are convex, $\Lambda_\pm$ are convex subsets of the direct-sum Banach space $\mathcal{T}_1(\mathcal{H}) \oplus \mathbb{R}$. By construction, they have nonempty interiors and due to Eq. (5) the interiors are nonintersecting, i.e., $\mathrm{Int}(\Lambda_+) \cap \mathrm{Int}(\Lambda_-) = \emptyset$. By the geometric Hahn-Banach theorem, Λ_+ and Λ_- can be separated by a closed hyperplane [cf. Fig. 1(b)]. Since, due to convexity, $\overline{\mathrm{Int}(\Lambda_\pm)} = \overline{\Lambda_\pm}$, this implies that there exists a continuous affine map $h : \mathcal{T}_1(\mathcal{H}) \to \mathbb{R}$ such that for all $X \in \mathcal{B}_1(\mathcal{H})$, $g_+(X) - \delta \le h(X) \le g_-(X) + \delta$. Using that $\hat{F} \le g_+$ and $g_- \le \hat{F}$, the previous inequality implies

$$-\delta \le g_+(X) - \hat{F}(X) - \delta \le h(X) - \hat{F}(X) \le g_-(X) - \hat{F}(X) + \delta \le \delta,$$

and so $|\hat{F}(X) - h(X)| \le \delta$. As $F(0) = 0$, we can choose h linear at the cost of $|\hat{F}(X) - h(X)| \le 2\delta$. Defining $T_A : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})$ as $T_A(X) := h(X)A/\|A\|_2^2$ then completes the proof of part (i).

Proof of part (ii). Let $\{A_j\}_{j=1}^{d^2}$ be a Hilbert-Schmidt orthonormal basis of self-adjoint operators on $\mathcal{H} = \mathbb{C}^d$ and $h_j : \mathcal{T}_1(\mathcal{H}) \to \mathbb{R}$ the corresponding linear maps from part (i). Define a linear map $T := \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})$, $T(X) := \sum_{j=1}^{d^2} h_j(X)A_j$. Then, for any $A = \sum_j a_j A_j$,

$$\begin{aligned}\mathrm{Tr}A(F(X) - T(X)) &= \sum_i a_i \mathrm{Tr}A_i F(X) - \sum_{i,j} a_i h_j(X)\mathrm{Tr}A_i A_j\\ &= \sum_i a_i(\mathrm{Tr}A_i F(X) - h_i(X))\\ &\le \|a\|_1 2\delta \le 2\delta d\|A\|_2.\end{aligned}$$

Therefore,

$$\|F(X) - T(X)\|_2 = \sup_{\|A\|_2 \le 1} \mathrm{Tr}A(F(x) - T(x)) \le 2\delta d.$$

□

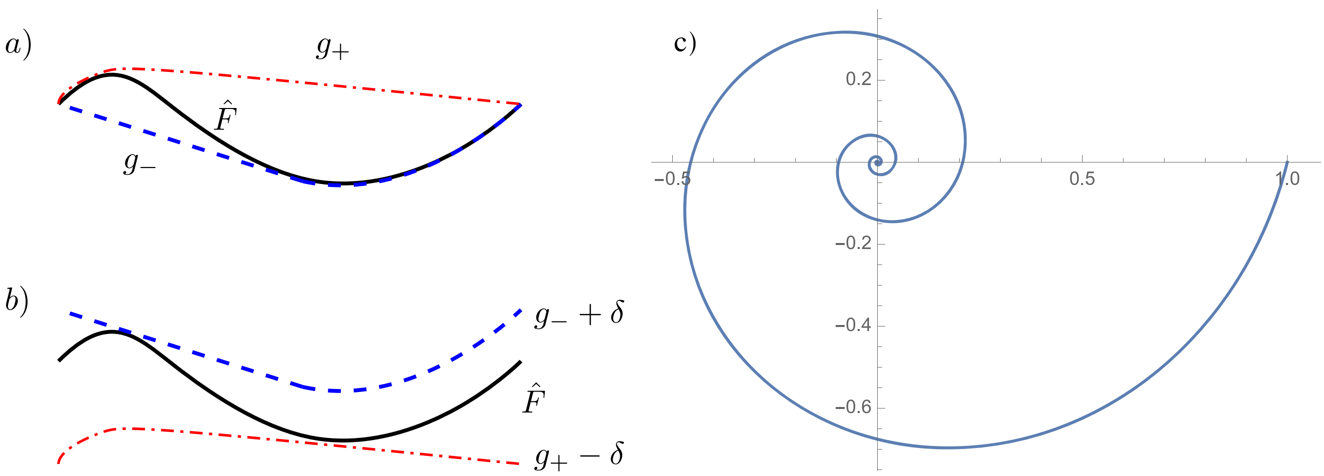


FIG. 1. (a) The convex (g_-) and concave (g_+) envelopes of $\hat{F}$. (b) Shifting them by an appropriate δ, the corresponding supergraph and subgraph can be separated by a hyperplane, which then serves as a linear approximation of $\hat{F}$. (c) The plot shows the image of the interval [0, 1] under the spiral map. Each point $z \in \mathbb{C}$ undergoes a rotation around the origin by an angle that is proportional to ln|z| (for better visibility a large value $\varepsilon = 8$ is chosen for the plot, while $\varepsilon \sim 1/\ln d$ is considered in the proof).

IV. ALMOST-SYMMETRIES FAR FROM LINEAR

In this section, we will address the question from the other end and show that no linear approximation [as in Theorem 2 (ii)] exists if the level of approximation is not allowed to depend on the dimension. This is even true with respect to the operator norm, for which the intrinsic dimension dependence is minimal. The result is summarized in the following theorem where $S_d := \{\psi \in \mathbb{C}^d : \|\psi\|_2 = 1\}$ denotes the unit sphere of $\mathbb{C}^d$.

Theorem 3 (Inapproximability). *Let $\varepsilon > 0$ and $d \in \mathbb{N}$ such that $d \geq e^{\frac{4\pi}{\varepsilon}} + 1$. There is a map $g : S_d \to S_d$ with the following properties:*

(i) $\forall \psi, \phi \in S_d$: $\left| |\langle g(\phi)|g(\psi)\rangle|^2 - |\langle \phi|\psi\rangle|^2 \right| \leq \varepsilon$.

(ii) *For every linear map $T : \mathbb{C}^{d\times d} \to \mathbb{C}^{d\times d}$, we have*

$$\sup_{\psi\in S_d} \| T(|\psi\rangle\langle\psi|) - |g(\psi)\rangle\langle g(\psi)| \|_\infty \geq \frac{1}{3}. \tag{6}$$

Particular instances for T would be $T(\cdot) = V \cdot V^*$, where V is a unitary or antiunitary on $\mathbb{C}^n$. The heart of the Proof of Theorem 3 is the *spiral map* [see Fig. 1(c)],

$$\mathbb{C} \ni z \mapsto z|z|^{i\frac{\varepsilon}{2}} = z e^{i\frac{\varepsilon}{2}\ln|z|}. \tag{7}$$

Its use goes back at least to the work of John[11] and it has since then been used in various similar proofs, e.g., in Refs. 12 and 13. It enters our discussion through the following:

Lemma 3. For any $\varepsilon > 0$, let $g : \mathbb{C}^d \to \mathbb{C}^d$ be the map that acts as in Eq. (7) componentwise.

(i) *For all $\psi, \phi \in S_d$, we have $\left| |\langle g(\psi)|g(\phi)\rangle|^2 - |\langle \psi|\phi\rangle|^2 \right| \leq \varepsilon$.*

(ii) *If $\varepsilon \ln(d-1) = 4\pi$, then $|\langle \varphi|g(\varphi)\rangle| = 0$ holds for $|\varphi\rangle := \frac{1}{\sqrt{2}}(1, \frac{1}{\sqrt{d-1}}, \ldots, \frac{1}{\sqrt{d-1}})$.*

Proof. Let $\psi = (\psi_1, \ldots, \psi_d)$ and $\phi = (\phi_1, \ldots, \phi_d)$. We bound each term separately to obtain

$$\begin{aligned}
\left| |\langle g(\psi)|g(\phi)\rangle|^2 - |\langle \psi|\phi\rangle|^2 \right| &\leq \sum_{k,l} |\phi_k\phi_l\psi_k\psi_l| \left| e^{i\frac{\varepsilon}{2}\ln\left|\frac{\phi_k\psi_l}{\phi_l\psi_k}\right|} - 1 \right|, \\
&\leq \frac{\varepsilon}{2}\sum_{k,l} \left| \phi_k\phi_l\psi_k\psi_l \ln\left|\frac{\phi_k\psi_l}{\phi_l\psi_k}\right| \right|, \\
&\leq \frac{\varepsilon}{2}\, 2 \sum_{k,l} |\phi_k\psi_l|^2 \\
&= \varepsilon \|\phi\|_2^2 \|\psi\|_2^2 = \varepsilon.
\end{aligned}$$

For the second inequality, we used $|e^{i\alpha} - 1| \leq |\alpha|$ for $\alpha \in \mathbb{R}$. The third inequality follows from considering the cases where $c := \left|\frac{\phi_k\psi_l}{\phi_l\psi_k}\right|$ is bigger or less than one and applying $|\ln c| \leq c$ for $c > 1$ and $|\ln c| \leq c^{-1}$ for $c < 1$.

Part (ii) of the Lemma follows from inserting $\varepsilon \ln(d-1) = 4\pi$ into

$$|\langle \varphi|g(\varphi)\rangle| = \left| \frac{1}{2} + \frac{1}{2}\exp\left[-\frac{i}{4}\varepsilon \ln(d-1) \right] \right|.$$

□

Proof of Theorem 3. We use the spiral map $g : S_d \to S_d$ that acts componentwise as in Eq. (7). If $d > \exp[4\pi/\varepsilon] + 1$, then we decrease ε until equality is achieved. In this way, Lemma 3 proves part (i) of the theorem and at the same time guarantees that g maps φ onto an orthogonal vector. In order to prove a bound on the best linear approximation, we exploit the symmetry of g. Let G be the subgroup of $U(d) \subseteq \mathbb{C}^{d\times d}$ that consists of all unitaries of the form $D\Pi$ where D is a diagonal unitary and Π a permutation matrix. Then, $\forall \psi \in \mathbb{C}^d$: $U^{-1}g(U\psi) = g(\psi)$ holds for all $U \in G$. The idea is now to argue that without loss of generality, the best linear approximation has the same symmetry.

For every unitarily invariant norm on $\mathbb{C}^{d\times d}$, in particular, for the operator norm, and for any linear map $T : \mathbb{C}^{d\times d} \to \mathbb{C}^{d\times d}$ consider the following chain of inequalities:

$$\sup_{\psi\in S_d}\|T(|\psi\rangle\langle\psi|)-|g(\psi)\rangle\langle g(\psi)|\| = \sup_{\psi\in S_d}\|UT(U^*|\psi\rangle\langle\psi|U)U^*-|g(\psi)\rangle\langle g(\psi)|\|$$
$$\geq \sup_{\psi\in S_d}\int\|UT(U^*|\psi\rangle\langle\psi|U)U^*-|g(\psi)\rangle\langle g(\psi)|\|dU$$
$$\geq \sup_{\psi\in S_d}\left\|\int UT(U^*|\psi\rangle\langle\psi|U)U^*dU-|g(\psi)\rangle\langle g(\psi)|\right\|,$$

where $U\in G$, dU is the Haar measure of G, and the first inequality uses $\sup\sum_k g_k \leq \sum_k \sup g_k$. Following these inequalities, we can lower bound the quality of approximation of any linear map T by the one of its symmetrized counterpart

$$T_G(A) := \int_G UT(U^*AU)U^*dU.$$

As proven in Lemma 1 of Ref. 14, any linear map with this symmetry is specified by three parameters $\alpha,\beta,\gamma\in\mathbb{C}$ and has the form

$$T_G(A) = \alpha \mathrm{Tr}[A]\mathbb{1} + \beta A + \gamma\, \mathrm{diag}(A), \tag{8}$$

where diag(A) is the diagonal part of the matrix A [strictly speaking, Ref. 14 considers quantum channels, but since the relevant commutant is a vector space that is closed under taking adjoints, the parametrization in Eq. (8) holds for all linear maps].

Using the state $|\varphi\rangle = \frac{1}{\sqrt{2}}(1,\frac{1}{\sqrt{d-1}},\dots,\frac{1}{\sqrt{d-1}})$ from Lemma 3 for which $\langle\varphi|g(\varphi)\rangle=0$, we can bound

$$\sup_{\psi\in S_d}\|T(|\psi\rangle\langle\psi|)-|g(\psi)\rangle\langle g(\psi)|\|_\infty \geq \|T_G(|\varphi\rangle\langle\varphi|)-|g(\varphi)\rangle\langle g(\varphi)|\|_\infty$$
$$\geq \max\left\{|\langle g(\varphi)|(T_G(|\varphi\rangle\langle\varphi|)-|g(\varphi)\rangle\langle g(\varphi)|)|g(\varphi)\rangle|,\ |\langle\varphi|(T_G(|\varphi\rangle\langle\varphi|)-|g(\varphi)\rangle\langle g(\varphi)|)|\varphi\rangle|\right\},$$
$$= \max\left\{\left|\alpha+\gamma\frac{d}{4(d-1)}-1\right|,\left|\alpha+\beta+\gamma\frac{d}{4(d-1)}\right|\right\} \geq \frac{|\beta+1|}{2},$$

where the last step used that for $x,y\in\mathbb{C}$, $\max\{|x|,|y|\}\geq(|x|+|y|)/2\geq|x-y|/2$. In order to eventually arrive at a parameter-independent lower bound, we need a second inequality in which β appears in a different way. For that purpose, let us denote the matrix of ones by $J_d\in\mathbb{R}^{d\times d}$, $J_{ij}=1$, and the projection $P:=\mathbb{1}-|1\rangle\langle 1|$. Since the operator-norm is submultiplicative, we can obtain another lower bound via

$$\|T_G(|\varphi\rangle\langle\varphi|)-|g(\varphi)\rangle\langle g(\varphi)|\|_\infty \geq \|P(T_G(|\varphi\rangle\langle\varphi|)-|g(\varphi)\rangle\langle g(\varphi)|)P\|_\infty,$$
$$= \left\|\left(\alpha+\frac{\gamma}{2(d-1)}\right)\mathbb{1}_{d-1}+\frac{\beta-1}{2(d-1)}J_{d-1}\right\|_\infty,$$
$$= \max\left\{\left|\alpha+\frac{\gamma}{2(d-1)}\right|,\left|\alpha+\frac{\gamma}{2(d-1)}+\frac{\beta-1}{2}\right|\right\},$$
$$\geq \frac{|\beta-1|}{4}.$$

Finally, combining the two β-dependent bounds, we obtain

$$3\sup_{\psi\in S_d}\|T(|\psi\rangle\langle\psi|)-|g(\psi)\rangle\langle g(\psi)|\|_\infty \geq \frac{|\beta+1|}{2}+2\frac{|\beta-1|}{4}\geq\frac{\beta+1}{2}-\frac{\beta-1}{2} = 1.$$

□

V. DISCUSSION

The inapproximability result of Theorem 3 shows that a dimension-independent linear approximation result is not possible. This raises the question about the optimal dimension-dependence of a positive result of the form in Theorem 2 (ii). Theorem 3 imposes a logarithmic lower bound in the following way:

For any map $f:\mathbb{P}(\mathbb{C}^d)\to\mathcal{T}_2(\mathbb{C}^d)$ that fulfills Wigner's condition up to ε according to Eq. (1) define $\Delta(f) := \inf_T \sup_{\psi\in S^d}\|T(|\psi\rangle\langle\psi|)-f(|\psi\rangle\langle\psi|)\|_\infty$ where the infimum is taken over all linear maps $T:\mathbb{C}^{d\times d}\to\mathbb{C}^{d\times d}$. Assume that $\sup_f\Delta(f)\leq\kappa(d)\varepsilon^p$ for some function κ and some $p>0$. Choosing $\varepsilon=4\pi/\ln(d-1)$ for sufficiently large d, Theorem 3 provides a map g that fulfills Eq. (1) together with $\Delta(g)\geq 1/3$. Therefore,

$$\kappa(d) \geq \frac{\sup_f \Delta(f)}{\varepsilon^p}, \\ \geq \frac{1}{3}\left(\frac{\ln(d-1)}{4\pi}\right)^p.$$

On the other hand, Theorem 2 guarantees that $\sup_f \Delta(f) \leq 4d\sqrt{\varepsilon}$ so that a significant gap between upper and lower bound remains. In order to close this gap, more sophisticated tools from Banach space theory might be useful (see end remark in Ref. 13).

ACKNOWLEDGMENTS

M.M.W. acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—Grant No. EXC-2111 - 390814868.

REFERENCES

[1] E. P. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren* (Fredrik Vieweg und Sohn, 1931).
[2] T. F. Jordan, "Assumptions that imply quantum dynamics is linear," Phys. Rev. A **73**, 022101 (2006).
[3] C. Simon, V. Buzek, and N. Gisin, "No-signaling condition and quantum dynamics," Phys. Rev. Lett. **87**, 170405 (2001).
[4] L. Masanes and M. P. Mueller, "A derivation of quantum theory from physical requirements," New J. Phys. **13**, 063001 (2011).
[5] V. Bargmann, "Note on Wigner's theorem on symmetry operations," J. Math. Phys. **5**, 862 (1964).
[6] G. Gehér, "An elementary proof for the non-bijective version of Wigner's theorem," Phys. Lett. A **30**, 2054 (2014).
[7] R. Simon, N. Mukunda, S. Chaturvedi, and V. Srinivasan "Two elementary proofs of the Wigner theorem on symmetry in quantum mechanics," Phys. Lett. A **372**, 6847 (2008); R. Simon, N. Mukunda, S. Chaturvedi, V. Srinivasan, and J. Hamhalter, "Comment on: 'Two elementary proofs of the Wigner theorem on symmetry in quantum mechanics [Phys. Lett. A 372, 6847 (2008)],'" *ibid.* **378**, 2332 (2014).
[8] D. Ilišević and A. Turnšek, "Stability of the Wigner equation – a singular case," J. Math. Anal. Appl. **429**, 273–287 (2015).
[9] A. Turnšek, "On mappings approximately preserving orthogonality," J. Math. Anal. Appl. **336**, 625 (2007).
[10] J. Alaminos, J. Extremera, and A. R. Villena, "Metric versions of Herstein's theorems on Jordan maps," Linear Multilinear Algebra **60**, 941 (2012).
[11] F. John, "Rotation and strain," Commun. Pure Appl. Math. **XIV**, 391 (1961).
[12] E. Matouskova, "Almost isometries of balls," J. Funct. Anal. **190**, 507 (2002).
[13] N. J. Kalton, "A remark on quasi-isometries," Proc. A.M.S. **131**, 1225 (2003).
[14] A. L. K. Hashagen and M. M. Wolf, "Universality and optimality in the information-disturbance tradeoff," Ann. Henri Poincaré **20**, 219–258 (2018).

A.2 Type and cotype constants and the linear stability of Wigner's symmetry theorem

Type and cotype constants and the linear stability of Wigner's symmetry theorem

Javier Cuesta

In the core article [1] we had shown that on a finite dimensional Hilbert space an almost-symmetry can be approximated by a linear map with an upper bound which depends linearly on the dimension of the Hilbert space. However, there the lower bound is logarithmically on the dimension and so an exponential improvement on the upper bound seems possible. The main idea of the upper bound in [1] relies on an almost-linear extension of an almost-symmetry together with a clever use of the geometric Hahn-Banach theorem. In this work, we study how the quality of the approximation would improve if we consider other almost-linear extensions with different domains and codomains. We do this by developing an idea of N. J. Kalton which uses sofisticated tools of geometric functional analysis. In Theorem 2 of this article we show that an almost-linear map can be approximated by a linear map and the quality of the approximation depends on the type 2 and cotype 2 constants of the domain and codomain of the almost-linear map. We use this in Theorem 4 in order to show that an almost-symmetry can be linearly approximated with an upper bound of the order square-root of the dimension of the Hilbert space. This improves the result of [1], but does not close the gap. However, it does provide a systematic study of the possible almost-linear extensions that we can consider and points out the optimal extension, namely the case that the domain and codomain are Hilbert spaces. In the latter the order of approximation is logarithmic and it would therefore close the gap.

I am the single author of this article and was thus solely involved in all parts of it.

Permission to include:

Javier Cuesta
Type and cotype constants and the linear stability of Wigner's symmetry theorem
Symmetry 11(9), 1107 (2019)

Symmetry

Aims

Symmetry (ISSN 2073-8994), an international and interdisciplinary scientific journal, publishes reviews, regular research papers and short notes. Our aim is to encourage scientists to publish their experimental and theoretical research in as much detail as possible. There is no restriction on the length of the papers. Full experimental and/or methodical details must be provided, so that results can be reproduced. There are in addition three unique features:

- Manuscripts regarding research proposals and research ideas will be particularly welcomed.
- Comments on any related papers published in this journal and other journals can be published as short letters.
- Electronic files regarding the full details of the calculation and experimental procedure, if unable to be published in a normal way, will be deposited as supplementary material.

Subject Areas

- **Physics:** conservation laws, Noether's theorem, spatial parity, charge parity, time parity, G-parity, standard model, internal symmetry, Lorentz symmetry, transformations, invariance, conservation, local and global symmetries, laws and symmetry, symmetry breaking, color symmetry, periodic and quasiperiodic crystals, time-reversal symmetry breaking, symmetry and complexity, Curie-Rosen symmetry principles, constants, biophysics, entropy, and indistinguishability.
- **Chemistry:** crystal and crystallography; chiral molecules, chiral resolution and asymmetric synthesis, asymmetric induction, chiral auxiliaries and chiral catalysts, stereochemistry, diastereomers, stereogenic, stereoisomers (enantiomers, atropisomers, diastereomers), stability, mixing, and phase separation.
- **Biology:** symmetry in biology, radial symmetry (tetramerism, pentamerism, *etc.*), diversity, preservation, sustainability, morphology, origin of life, and molecular evolution (homochirality).
- **Mathematics:** invariance, transformation, group theory, Lie groups, chirality, achiral or amphichiral, helix and Möbius strip, knot theory, graph theory, isometry, plane of symmetry, skewness, vertex algebra, asymmetry, dissymmetry, nonsymmetry and antisymmetry, supergroups and nonlinear algebraic structures, supersymmetry and supergravity, strings and branes, integrability and geometry, information theory, Felix Klein's Erlangen Program, and continuous symmetry.

- **Computer Science, Theory and Methods:** computer-aided design, computational geometry, computer graphics, visualization, image compression, data compression, pattern recognition, diversity, similarity, and conservation and sustainability.
- **Symmetry and other scientific disciplines and engineering.**

MDPI Publication Ethics Statement

C O P E *Symmetry* is a member of the Committee on Publication Ethics (COPE). MDPI takes the responsibility to enforce a rigorous peer-review together with strict ethical policies and standards to ensure to add high quality scientific works to the field of scholarly publication. Unfortunately, cases of plagiarism, data falsification, inappropriate authorship credit, and the like, do arise. MDPI takes such publishing ethics issues very seriously and our editors are trained to proceed in such cases with a zero tolerance policy. To verify the originality of content submitted to our journals, we use iThenticate to check submissions against previous publications. MDPI works with Publons to provide reviewers with credit for their work.

Book Reviews

Authors and publishers are encouraged to send review copies of their recent related books to the following address. Received books will be listed as *Books Received* within the journal's *News & Announcements* section.

MDPI
St. Alban-Anlage 66
CH-4052 Basel
Switzerland

E-mail:

Copyright / Open Access

Articles published in *Symmetry* will be Open-Access articles distributed under the terms and conditions of the Creative Commons Attribution License (CC BY). The copyright is retained by the author(s). MDPI will insert the following note at the end of the published text:

© 2019 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/).

Reprints

Reprints may be ordered. Please contact for more information on how to order reprints.

Journal History

See https://www.mdpi.com/journal/symmetry/history

Announcement and Advertisement

Announcements regarding academic activities such as conferences are published for free. Advertisement can be either published or placed on the pertinent website. For further information, please contact the editorial office.

Editorial Office

Ms. Dalia Su

Managing Editor
E-Mail

For further MDPI contacts, see here.

symmetry

MDPI

Article

Type and Cotype Constants and the Linear Stability of Wigner's Symmetry Theorem

Javier Cuesta [1,2]

[1] Department of Mathematics, Technische Universität München, 85748 Garching, Germany; j.cuesta@tum.de
[2] Munich Center for Quantum Science and Technology (MCQST), D-80799 München, Germany

Received: 5 August 2019; Accepted: 23 August 2019; Published: 3 September 2019

Abstract: We study the relation between almost-symmetries and the geometry of Banach spaces. We show that any almost-linear extension of a transformation that preserves transition probabilities up to an additive error admits an approximation by a linear map, and the quality of the approximation depends on the type and cotype constants of the involved spaces.

Keywords: Wigner's theorem; stability; almost-symmetry; almost-linear; type; cotype; Banach spaces

---

1. Introduction and Preliminaries

In the work of N. J. Kalton [1–3], we can find novel ideas and methods for the stability of functional equations that depart from the classical methods of Hyers, Ulam and Rassias [4]. In Ref. [3] (see Theorem 2.2), Kalton provides a sharp bound on the stability of the additive map in $\mathbb{R}^n$ for the so-called singular case. His proof makes use of probabilistic and geometric methods in Banach space theory. This paper ends with a sketch on how the theory of twisted sums in Banach space theory could be used to obtain the same result. In this note, we develop this last idea and use it to obtain an improvement in the linear stability of Wigner's symmetry theorem (see Theorem 3).

Wigner's celebrated symmetry theorem [5] is not only central for physics, but it also finds an important role in many preservers' problems. A *preserver problem* deals with the characterization of maps, primarily on matrix spaces and operator algebras that preserve certain functional, subset, or an invariant. In particular, in the field of Quantum Information Theory (QIT) it has been shown [6] that the only mapping T that preserves the $f-$divergences (this includes the von Neumann and relative entropy) is a Wigner symmetry transformation, i.e., of the form $T(x) = UxU^*$, where U is either a unitary or antiunitary transformation on $\mathbb{C}^d$. It turns out that most of the proofs of different preservers problems can be reduced to Wigner's theorem. Therefore, it is natural to expect that sharp bounds on the stability of Wigner's theorem could provide good approximations for a wide range of almost-preserving problems. It is worth pointing out that there exists a close relation between geometric functional analysis and many questions in QIT [7]. This is the point of view that we want to motivate here.

It has been recently shown [8] that an arbitrary *almost-symmetry* in quantum theory, i.e., a transformation on the set of pure states $\mathbb{P}(\mathcal{H})$ in a separable complex Hilbert space $\mathcal{H}$ that almost preserves the transition probabilities up to an error ε, can be approximated by a linear map H if and only if $\mathcal{H}$ is a finite-dimensional Hilbert space. For an infinite-dimensional Hilbert space, the approximation is in a weak sense (see Theorem 2-(i) in [8]). The quality of the approximation for a d-dimensional Hilbert space $\mathcal{H}$ was obtained to be

$$\frac{1}{3}\sqrt{\frac{\ln(d-1)\varepsilon}{4\pi}} \leq \|f(x) - H(x)\|_2 \leq 4d\sqrt{\varepsilon}, \tag{1}$$

where $\|\cdot\|_2$ is the Hilbert–Schmidt norm. The main idea for the upper bound in Equation (1) was to consider an almost-linear extension of f with some particular domain and codomain, followed by an application of the geometric Hahn–Banach theorem. In this work, we explore how the quality of the approximation depends on the consideration of various classes of almost-linear extensions. These extensions now have arbitrary finite-dimensional Banach spaces as domain and codomain.

Throughout this note, we will be entirely concerned with finite-dimensional Banach spaces and the twisted sums generated by almost-linear maps. A map $F : X \to Y$ between Banach spaces will be called *almost-linear* if it satifies the following two conditions:

(i) $F(\lambda x) = \lambda F(x)$ for all $\lambda \in \mathbb{R}$ and $x \in X$,
(ii) there exists a $\delta > 0$ such that for any finite sequence $(x_i)_{i=1}^m \subset X$, $m \in \mathbb{N}$ and $\lambda \in \mathbb{R}^m$,

$$\left\| \sum_{i=1}^m \lambda_i F(x_i) - F\left(\sum_{i=1}^m \lambda_i x_i\right) \right\|_Y \le \delta \sum_{i=1}^m |\lambda_i| \, \|x_i\|_X . \tag{2}$$

We will show that, for every almost-linear map F, there exists a linear map H whose distance to F depends additively on δ and on some geometric invariants of the domain and target space of F (see Theorem 1). The Banach space numbers used to express the results are the type and cotype constants which we introduce now. Let $\{\gamma_j\}_{j=1}^n$ be a sequence of independent real Gaussian random variables, i.e., for each Borel subset $B \subset \mathbb{R}$, each random variable has a distribution

$$\mu(\gamma \in B) = \frac{1}{(2\pi)^{1/2}} \int_B e^{-\frac{t^2}{2}} dt.$$

Let X be a Banach space with norm $\|\cdot\|$ and let $p \in [1,2]$, $q \in [2,\infty)$. For every positive interger n, we define $T_{p,n}(X), C_{q,n}(X)$ to be the smallest constants such that for arbitrary sequences $\{x_j\}_{j=1}^n \subset X$, we have

$$\left(\mathbb{E} \left\| \sum_{j=1}^n \gamma_j x_j \right\|^2 \right)^{1/2} \le T_{p,n}(X) \left(\sum_{j=1}^n \|x_j\|^p \right)^{1/p},$$

$$C_{p,n}(X)^{-1} \left(\sum_{j=1}^n \|x_j\|^q \right)^{1/q} \le \left(\mathbb{E} \left\| \sum_{j=1}^n \gamma_j x_j \right\|^2 \right)^{1/2}.$$

The space X is said to be of Gaussian type p (resp. Gaussian cotype q) if $T_p(X) = \sup_n T_{p,n}(X) < \infty$ (resp. $C_q(X) = \sup_n C_{q,n}(X) < \infty$). One can analogously define the Rademacher type and cotype by exchanging the Gaussian sequence by a Rademacher sequence. The results shown in this note are valid for both notions of type and cotype.

For $r \in [1,\infty)$, we denote by S_r^d the Hermitian part of the d-dimensional r-Schatten class and by l_r^d the classical space of r-summable sequences in $\mathbb{R}^d$; the space S_r^d is a real Banach space with norm $\|x\|_r := (\mathrm{Tr}\,|x|^r)^{1/r}$. Table 1 summarizes the behaviour of the type and cotype constants for the r-Schatten classes that we use (see Ref. [9] for details).

We now introduce some notation. The set of rank-one projections in $\mathbb{C}^{d\times d}$ is denoted by $\mathbb{P}(\mathbb{C}^d)$. The unit ball of a space Z is written as B_Z. The convex hull of a set S is the set of convex combinations of elements of S, which we denote by $\mathrm{conv}(S)$. The set of linear maps between X and Y is $L(X,Y)$. A linear projection $P \in L(X,Y)$ is a linear map such that $P^2 = P$. Finally, we denote by $\langle x, y\rangle := \mathrm{Tr}(xy)$ the Hilbert–Schmidt inner product in the real vector space of Hermitian matrices $\mathcal{H}_d$.

In the next section, we introduce a special space which will generate the linear approximation to the almost-linear map $F : X \to Y$. This space is an extension of X and Y and is called a twisted sum (basically because it “twists” the unit ball of X and Y according to F). Twisted sums were extensively studied by Kalton [1] in the context of the three-space problem. In particular, Kalton showed that twisted sums are in correspondence with quasi-linear maps; this is a weaker condition than almost-linearity, but, for our purposes, it suffices to say that any almost-linear map is a quasi-linear map. See Ref. [10] for a detailed exposition of this topic.

Table 1. Upper bounds for the Rademacher type and cotype constants of the spaces l_r^d and S_r^d. The Gaussian type and cotype for these spaces behave in the same way, up to a factor of $\sqrt{2/\pi}$, as the Rademacher type and cotype. For a Hilbert space, the type and cotype constants are always equal to one.

	Type $p \in [1,2]$	**Cotype $q \in [2,\infty]$**
l_1^d	$d^{1-\frac{1}{p}}$	$\sqrt{2}$
Hilbert space	1	1
l_∞^d	$(4\log d)^{1-1/p}$	$d^{1/q}$
S_1^d	$d^{1-1/p}$	$\sqrt{e}$
S_∞^d	$(4\log d)^{1-1/p}$	$d^{1/q}$

2. Finite-Dimensional Twisted Sums

Let X, Y be two Banach spaces with dimension d_1, d_2, respectively. The twisted sum of Y and X is a (d_1+d_2)-dimensional space Z that contains a subspace Y_0 that is isomorphic to Y and such that Z/Y_0 is isomorphic to X. The twisted sums that interest us are constructed with an almost-linear function F. Consider $\delta > 0$ and the Cartesian product $Y \oplus X$ (the order is important) endowed with the quasi-norm:

$$|||(y,x)|||_F := \frac{\|y - F(x)\|_Y}{\delta} + \|x\|_X . \tag{3}$$

Then, $Y_0 = \{(y,0) : y \in Y\}$ is δ^{-1}-isometric to Y and Z/Y_0-isometric to X. Note that, since F is homogeneous, $|||(-y,-x)|||_F = |||(y,x)|||_F$ and $|||(y,x)|||_F = 0$ implies that $(y,x) = 0$. Although $|||(y_1,x_1)+(y_2,x_2)|||_F \le 2(|||(y_1,x_1)|||_F + |||(y_2,x_2)|||_F)$, we can still endow Z with a norm. The twisted sum Z can be made into a Banach space with the norm

$$\|(y,x)\| := \inf\left\{\sum_j |||(y_j,x_j)|||_F : (y,x) = \sum_j (y_j,x_j)\right\}. \tag{4}$$

The fact that the above expression defines a norm will be shown below. The completion of a quasi-Banach space Z whose dual is non-trivial with respect to this norm is known as the Banach envelope of Z [11]. In order to avoid charged notation, we also denote the Banach envelope by Z.

Lemma 1. *Let $|||\cdot|||$ be a quasi-norm on Z, then the following equivalent expressions define a norm on Z. For $z \in Z$,*

$$\|z\| = \inf\left\{\sum_{j=1}^n |||z_j||| : z = \sum_{j=1}^n z_j\right\}, \tag{5}$$

$$= \inf\{\lambda > 0 : z/\lambda \in \mathrm{conv}(B_Z)\}, \tag{6}$$

$$= \inf\{\xi(z) : \xi \in Z^*, \|\xi\| \le 1\}. \tag{7}$$

Moreover, for the quasi-norm defined by Equation (3), we have the following equivalence:

$$\|(y,x)\| \le |||(y,x)|||_F \le 2\,\|(y,x)\| . \tag{8}$$

Proof. We show first that the first expression indeed defines a norm. Since $|||\cdot|||$ is a quasi-norm, the only property that we need to check is the triangle inequality. This can be verified by

$$\begin{aligned}
\|z_1+z_2\| &= \inf\left\{\sum_{j=1}^{n}|||w_j||| : z_1+z_2=\sum_{j=1}^{n}w_j=\sum_{j=1}^{n_1}w_j+\sum_{j=1}^{n_2}w_j\right\}, \\
&\leq \inf\left\{\sum_{j=1}^{n_1}|||w_j||| : z_1=\sum_{j=1}^{n_1}w_j\right\}+\inf\left\{\sum_{j=1}^{n_2}|||w_j||| : z_2=\sum_{j=1}^{n_2}w_j\right\}, \\
&= \|z_1\|+\|z_2\|,
\end{aligned}$$

as those are valid decompositions of z_1+z_2. We show now that Equations (5) and (6) are the same. Let $\alpha=\|z\|$ be the infimum of Equation (6). Then, there exist $m\in\mathbb{N}$, positive real numbers $(\lambda_j)_{j=1}^{m}$, $\sum_{j=1}^{m}\lambda_j=1$ and $(z_j)_{j=1}^{m}$ with quasi-norm one such $z=\alpha\sum_{j=1}^{m}\lambda_j z_j$. This is a valid decomposition of z and $\sum_{j=1}^{m}|||\alpha\lambda_j z_j|||\leq\alpha$. On the other hand, let $z=\sum_{j=1}^{m}z_j$ be the decomposition that achieves the infimum in Equation (5) so that $\|z\|=\sum_{j=1}^{m}|||z_j|||$. Then,

$$\frac{z}{\sum_{k=1}^{m}|||z_k|||}=\sum_{j=1}\left(\frac{|||z_j|||}{\sum_{k=1}^{m}|||z_k|||}\right)\frac{z_j}{|||z_j|||}\in\mathrm{conv}(B_Z).$$

The norm of $\xi\in Z^*$ can be computed as

$$\|\xi\|=\sup_{z\in\mathrm{conv}(B_Z)}|\xi(z)|=\sup_{z\in B_Z}|\xi(z)|=\sup\{|\xi(z)| : |||z|||\leq 1\},$$

as the supremum over a convex function is achieved at the extremal points. Thus, the dual of the quasi-Banach space Z and its Banach envelope coincide. Thus, Equation (7) is just the usual expression in terms of the dual. We now compare the quasi-norm in Equation (3) with the norm of its envelope.

Since $\|z\|$ is defined by the infimum of $\sum_j|||z_j|||$ over all the decompositions of z, Equation (5), we immediately have the first inequality in Equation (8). For the second inequality, let $(y,x)=\sum_j(y_j,x_j)$, then, using Equation (2),

$$\begin{aligned}
|||(y,x)|||_F &= \frac{\|F(x)-y\|_Y}{\delta}+\|x\|_X, \\
&= \frac{\left\|F(\sum_j x_j)-\sum_j F(x_j)+\sum_j F(x_j)-\sum_j y_j\right\|_Y}{\delta}+\left\|\sum_j x_j\right\|_X, \\
&\leq \frac{\left\|F(\sum_j x_j)-\sum_j F(x_j)\right\|_Y}{\delta}+\sum_j\frac{\|F(x_j)-y_j\|_Y}{\delta}+\sum_j\|x_j\|_X, \\
&\leq \frac{\delta\sum_j\|x_j\|_X}{\delta}+\sum_j\frac{\|F(x_j)-y_j\|_Y}{\delta}+\sum_j\|x_j\|_X, \\
&\leq 2\sum_j|||(y_j,x_j)|||_F.
\end{aligned}$$

□

Additionally, we can understand the resulting twisted sum Z with a norm as in Equation (5) as the space with unit ball [12]

$$B_Z:=\mathrm{conv}\left(\{(y,0):\|y\|_Y\leq 1\}\cup\{(F(x),x):\|x\|_X\leq 1\}\right).$$

We write $Z=Y\oplus_F X$ for the (Banach envelope of) twisted sum of Y and X generated by the almost-linear map $F: X\to Y$.

3. Main Result

We are now ready to put all the pieces together and to make the connection explicitly between (co)type constants and the linear stability of almost-linear maps.

Theorem 1. *Let $F : X \to Y$ be an almost-linear map between finite-dimensional real Banach spaces, i.e., F is a real homogeneous map and there exists a $\delta > 0$ such that, for any finite sequence $(x_i)_{i=1}^m \subset X, m \in \mathbb{N}$ and $\lambda \in \mathbb{R}^m$,*

$$\left\| \sum_{i=1}^m \lambda_i F(x_i) - F\left(\sum_{i=1}^m \lambda_i x_i\right) \right\|_Y \leq \delta \sum_{i=1}^m |\lambda_i| \, \|x_i\|_X .$$

Let $Z = Y \oplus_F X$ be the respective twisted sum generated by this map. Then,

$$\inf_{H \in L(X,Y)} \sup_{x \in X} \frac{\|F(x) - H(x)\|_Y}{\|x\|_X} \leq 2\delta \min\{T_2(Z)C_2(X), 1 + T_2(Z^*)C_2(Y^*)\}, \tag{9}$$

where T_2 and C_2 are the type 2 and cotype 2 constants.

Proof of Theorem 1. We need the following important theorem of Maurey [13] (see Theorem 7.4.4 in Ref. [14] for a modern proof).

Theorem 2 (Maurey's Extension)**.** *Let E be a Banach space and S a closed subspace of E. Let $T_2(E)$ be either the Gaussian or Rademacher type 2 constant of E and $C_2(S)$ either the Gaussian or Rademacher cotype 2 constant of S. Then, there exists a projection $P : E \to S$ with*

$$\|P\| \leq T_2(E)C_2(S).$$

We remark that the norm $\|\cdot\|$ in Theorem 2 is the operator norm. This might seem odd at first sight as usually the projections are considered between Hilbert spaces and in that case they always have a norm equal to one. This is no longer true when we leave the special world of Hilbert spaces and consider general Banach spaces. Maurey's theorem is proven by factorizing through a Hilbert space though. In a sense, the notions of type and cotype measure how far we are from the Hilbert space scenario.

Let $Z = Y \oplus_F X$ be the twisted sums of Y and X and consider the Banach envelope of Z. Let us denote by Z as well the Banach envelope of Z. From Maurey's theorem, we know there exists a projection $P : Z \to X$ such that $\|P\| \leq T_2(Z)C_2(X)$. Since P is a projection, it has the general form $P(y, x) = (y - H(x), 0)$, where $H : X \to Y$ is a linear map. Then, using Equation (8),

$$\begin{aligned}
\|P\| = \sup_{(y,x) \in Z} \frac{\|P(y,x)\|}{\|(y,x)\|} &\geq \sup_{x \in X} \frac{\|P(F(x),x)\|}{\|(F(x),x)\|} \\
&= \sup_{x \in X} \frac{\|(F(x) - H(x), 0)\|}{\|(F(x),x)\|} \\
&\geq \sup_{x \in X} \frac{|||(F(x) - H(x), 0)|||}{2|||(F(x),x)|||} \\
&= \sup_{x \in X} \frac{\|F(x) - H(x)\|_Y}{2\delta \, \|x\|_X} \\
&\geq \inf_{H \in L(X,Y)} \sup_{x \in X} \frac{\|F(x) - H(x)\|_Y}{2\delta \, \|x\|_X}.
\end{aligned}$$

We can also consider a dual construction for a different bound. Let Z^* be the dual of the twisted sum $Y \oplus_F X$. It is known [15] that the dual of Z is isomorphic to $X^* \oplus_{F^*} Y^*$, where F^* is in some sense the dual map of F (see [15] for details). Since we are dealing with finite-dimensional spaces, Z^{**} can be identified with Z. Let $Q : Z^* \to Y^*$ be the projection obtained by Maurey's extension theorem when applied to the Banach spaces Z^* and Y^*. Let us consider the projection $\tilde{P} : Z \to X$ defined via $\tilde{P} := \mathrm{id} - Q^*\pi$, where π is the quotient map $\pi : Z \to X$, $\pi(y, x) = x$. Indeed, let $\Omega : X \to Y$ be the linear map induced by Q^*. Then,

$$\begin{aligned}\tilde{P}(y,x) &= (y,x) - Q^*\pi(y,x),\\ &= (y,x) - Q^*x,\\ &= (y,x) - (\Omega(x),x) = (y-\Omega(x),0) \in X.\end{aligned}$$

Analogously with the previous calculation, we find

$$\inf_{\Omega \in L(X,Y)} \sup_{x \in X} \frac{\|F(x) - \Omega(x)\|_Y}{\|x\|_X} \leq 2\delta \left\|\tilde{P}\right\| \leq 2\delta(1 + \|Q\|).$$

The final results then follows from the upper bound that Maurey's theorem provides on the norm of such projections. □

4. Applications

The following result gives an improvement on Theorem 2-(ii) in [8].

Theorem 3 (Linear Stability of Wigner's theorem). *Let $f : \mathbb{P}(\mathbb{C}^d) \to \mathbb{P}(\mathbb{C}^d)$ be a function that satisfies*

$$|\langle f(x), f(y)\rangle - \langle x, y\rangle| \leq \varepsilon \qquad \text{for all } x, y \in \mathbb{P}(\mathbb{C}^d). \tag{10}$$

Then, there exists a universal constant C and a linear map $H : \mathcal{H}_d \to \mathcal{H}_d$ such that, for all $x \in \mathbb{P}(\mathbb{C}^d)$,

$$\|f(x) - H(x)\|_2 \leq (C \log_2 d)^\beta \sqrt{d\varepsilon},$$

where $\beta = 2 + \frac{1}{2} \log_2 \log_2 2d$.

We call a map $f : \mathbb{P}(\mathbb{C}^d) \to \mathbb{P}(\mathbb{C}^d)$ that satisfies Equation (10) an *almost-symmetry*. In order to prove Theorem 3, we make use of the following lemmas (c.f. Theorem 1 in [12]). First, we need the type constant of a twisted sum (cf. Lemma 16.6-7 in [16]).

Lemma 2. *Let Z be the twisted sum of Y and X; then,*

$$T_{2,n^2}(Z) \leq T_{2,n}(Y)T_{2,n}(Z) + T_{2,n}(Y)T_{2,n}(X) + T_{2,n}(Z)T_{2,n}(X). \tag{11}$$

The type 2 constant of a Banach space of dimension d can be obtained from the type constant restricted to families of size $d(d+1)/2$ as stated by the following lemma. This result follows from a cone version of Caratheodory's theorem (see Lemma 6.1 in [17]).

Lemma 3. *Let X be a d-dimensional Banach space. Then, $T_{2,n}(X) = T_{2,d(d+1)/2}(X)$ and $C_{2,n}(X) = C_{2,d(d+1)/2}(X)$ for any $n \geq d(d+1)/2$.*

Proof. The first step of the proof consists of extending the function f to $F : S_1^d \to S_1^d$ such that $F|_{\mathbb{P}(\mathbb{C}^d)} = f$. We take x in the unit sphere of S_1^d and identify it with its antipodal point $-x$. We choose a fixed spectral decomposition for both elements, say $x = \sum_{j=1}^d \lambda_j x_j$, and define $F(x) := \sum_{j=1}^d \lambda_j f(x_j)$. Then, we can extend F homogeneously from the unit sphere to any $y \in S_1^d$ by multiplying x or $-x$ with $\lambda \geq 0$ so that $\lambda x = y$ or $-\lambda x = y$. We call again this extension F. By construction, F is a real homogeneous map. Note that this extension is not unique, but we do not need this here.

As proven in Lemma 2 of Ref. [8], F is an almost-linear map

$$\left\| \sum_{i=1}^{m} \lambda_i F(x_i) - F\left(\sum_{i=1}^{m} \lambda_i x_i \right) \right\|_2 \leq \delta \sum_{i=1}^{m} |\lambda_i| \, \|x_i\|_1 ,$$

with $\delta = 2\sqrt{\varepsilon}$. If we use Theorem 1 with the twisted sum $Z := S_2^d \oplus_F S_1^d$, we will see that we cannot obtain anything better than a linear dependence on d. However, we will be able to obtain a better dimension dependence if we consider a dual construction, namely with $Z^* := S_\infty^d \oplus_{F^*} S_2^d$. For that matter, we use Lemmas 2 and 3 in order to estimate the type 2 constant of Z^*. From Equation (11) and $T_2(S_\infty) \leq \sqrt{4 \log d}$, we obtain $T_{2,n^2}(Z^*) \leq 2\sqrt{8 \log_2 d}\; T_{2,n}(Z^*)$ for all $n \in \mathbb{N}$. It is known that, for a general Banach space E, $T_2(E) \leq \sqrt{\dim(E)}$ (Proposition 12.3 in [9]). Thus, for all two-dimensional subspaces of Z, the type is less than $\sqrt{2}$ and $T_{2,2}(Z) \leq \sqrt{2}$ (this can be alternatively derived from a classical result of John and the relation between the Banach–Mazur distance and type 2 constants). It follows from induction that

$$T_{2,2^{2^k}}(Z^*) \leq (2\sqrt{8 \log_2 d})^k \sqrt{2},$$

which, in turn, implies

$$T_{2,n}(Z^*) \leq \sqrt{2}(\log_2 n)(8 \log_2 d)^{\frac{\log_2 \log_2 n}{2}}.$$

The dimension of the real vector space of Hermitian matrices $\mathcal{H}_d$ is d^2. Therefore, we obtain from Lemma 3 with $n = 2d^4$

$$T_2(Z^*) \leq 2(8 \log_2 d)^{2+\frac{\log_2 \log_2 2d}{2}}.$$

It follows from Theorem 1 and $C_2(S_\infty^d) \leq \sqrt{d}$ that there exists a linear map $H : S_1^d \to S_2^d$ such that

$$\sup_{x \in \mathbb{P}(\mathbb{C}^d)} \|f(x) - H(x)\|_2 \leq 4(8 \log_2 d)^{2+\frac{\log_2 \log_2 2d}{2}} \sqrt{d\varepsilon}.$$

□

The following proposition is essentially due to Kalton. It can be shown using Theorem 2.2 in [3] as S_2^d and $\mathbb{R}^{d^2}$ are isomorphic Hilbert-spaces. We present here a proof using the notions of (co)type and Theorem 1.

Proposition 1 (Stability of Global Symmetries). *Let $f : B_{S_2^d} \to S_2^d$ be a continuous function that satisfies*

$$|\langle f(x), f(y) \rangle - \langle x, y \rangle| \leq \varepsilon \qquad \text{for all } x, y \in B_{S_2^d}. \tag{12}$$

Then, there exists a linear map $H : S_2^d \to S_2^d$ and an absolute constant C such that, for all $X \in B_{S_2^d}$,

$$\|f(x) - H(x)\|_2 \leq C\sqrt{\varepsilon} \log_2 d.$$

Proof of Proposition 1. The first step consists of showing that the function f can be extended to a continuous homogeneous function on the whole space without paying much.

Lemma 4. *Let $f : B_{S_2^d} \to S_2^d$ be a continuous function that satisfies*

$$|\langle f(x), f(y) \rangle - \langle x, y \rangle| \leq \varepsilon \qquad \text{for all } x, y \in B_{S_2^d}.$$

Then, there exists a continuous and homogeneous function $F : S_2^d \to S_2^d$ *such that*

$$\left\| \sum_{j=1}^{n} F(x_j) - F\left(\sum_{j=1}^{n} x_j\right) \right\|_2 \le 4\sqrt{\varepsilon} \sum_{j=1}^{n} \|x_j\|_2 \qquad \text{for all } x_j \in S_2^d, \tag{13}$$

and

$$\sup_{X \in B_{S_2^d}} \|f(x) - F(x)\|_2 \le 3\sqrt{\varepsilon}. \tag{14}$$

Proof of Lemma 4. Let us extend f to $F : S_2^d \to S_2^d$ where

$$F(x) := \|x\|_2 \left(f\left(\frac{x}{2\|x\|_2}\right) - f\left(-\frac{x}{2\|x\|_2}\right)\right).$$

This function is homogeneous, i.e., $F(\lambda x) = \lambda F(x)$ for all $\lambda \in \mathbb{R}$, and continuous as f and $\|\cdot\|$ are also continuous. Using Equation (12) and the triangle inequality, we obtain the new almost-symmetry condition

$$|\langle F(x), F(y)\rangle - \langle x, y\rangle| \le 4\varepsilon \|x\|_2 \|y\|_2 . \tag{15}$$

Hence, for any $z \in S_2^d$,

$$\begin{aligned}\left|\left\langle \sum_{j=1}^{n} F(x_j) - F\left(\sum_{j=1}^{n} x_j\right), F(z)\right\rangle\right| &= \left|\left\langle \sum_{j=1}^{n} F(x_j) - F\left(\sum_{j=1}^{n} x_j\right), F(z)\right\rangle - \left\langle \sum_{j=1}^{n} x_j - \sum_{j=1}^{n} x_j, z\right\rangle\right| \\ &\le 8\varepsilon \sum_{j=1}^{n} \|x_j\|_2 \|z\|_2 .\end{aligned}$$

Therefore, from the linearity of the inner product, we obtain Equation (13). Finally, we show that f and F are $\sqrt{\varepsilon}$-close. From Equation (12),

$$\begin{aligned}&|\langle F(x), f(x)\rangle - \langle x, x\rangle| \\ &= |\, \|x\|_2 \left(\langle f\left(\frac{x}{2\|x\|}\right), f(x)\rangle - \frac{\|x\|_2}{2}\right) - \|x\|_2 \left(\langle f\left(\frac{-x}{2\|x\|}\right), f(x)\rangle + \frac{\|x\|_2}{2}\right) | \\ &\le 2\varepsilon \|x\|_2 .\end{aligned}$$

Thus, with Equation (15), we have

$$\begin{aligned}\|F(x) - f(x)\|_2^2 &= \|F(x)\|_2^2 - \|x\|_2^2 - 2\,\mathrm{Re}\,(\langle F(x), f(x)\rangle - \langle x, x\rangle) + \|f(x)\|_2^2 - \|x\|_2^2 \\ &\le 4\varepsilon \|x\|_2^2 + 4\varepsilon \|x\|_2 + \varepsilon,\end{aligned}$$

which is less than 9ε for all $x \in B_{S_2^d}$. □

We consider now the twisted sum $Z = S_2^d \oplus_F S_2^d$ generated by the almost-linear map F. Before applying Theorem 1, we estimate the type 2 constant of Z. Since S_2^d is a Hilbert space, it has a type 2 constant equal to one and we obtain from Lemma 2 that

$$T_{2,n^2}(Z) \le 1 + 2T_{2,n}(Z) \qquad \text{for all } n. \tag{16}$$

As in the proof of Theorem 3, all two-dimensional subspaces of Z have a type less than $\sqrt{2}$ and $T_{2,2}(Z) \leq \sqrt{2}$. It follows from induction that, for $n \geq 3$,

$$T_{2,n}(Z) \leq 2(1+\sqrt{2}) \log_2 n.$$

Hence, from Lemma 3 with $n = 4d^2$,

$$T_2(Z) \leq 4(1+\sqrt{2}) \log_2 2d.$$

Accordingly, from $C_2(S_2^d) = 1$ and Theorem 1, there exists a linear map $H : S_2^d \to S_2^d$ such that, for all $x \in B_{S_2^d}$,

$$\|F(x) - H(x)\|_2 \leq 32(1+\sqrt{2}) \log_2(2d)\sqrt{\varepsilon}. \tag{17}$$

Finally, from Equation (14) and the triangle inequality, we obtain

$$\begin{aligned}\sup_{x \in B_{S_2^d}} \|f(x) - H(x)\|_2 &\leq \sup_{x \in B_{S_2^d}} \|f(x) - F(x)\|_2 + \|F(x) - H(x)\|_2 \\ &\leq 79\sqrt{\varepsilon}\,(1+\log_2 d).\end{aligned}$$

□

5. Discussion and Perspectives

Using Theorem 1, we are able to improve—up to some logarithmic factors—the upper bound on the dimension dependence of the linear stability of Wigner's theorem from d to $\sqrt{d}$. There seems to be room for an exponential improvement in the dimension as the lower bound is of order $\log d$ (see the discussion section of [8]). The method developed here allows us to study systematically the limitations of considering other types of almost-linear extensions to solve this problem. Even if we were able to extend the almost-symmetry f to an almost-linear map $F : S_1^d \to S_1^d$ with δ independent of d, we would still get from Theorem 1 an upper bound of order $\sqrt{d}$. This is just a consequence of how the type and cotype constants of S_1^d and S_∞^d behave. There is a trade-off in Theorem 1 between the type constant for individual spaces and the type constant of their twisted sum.

It can be seen from Table 1 and Lemma 2 that the best bound that can be obtained from Theorem 1 is in the case that X and Y are Hilbert spaces. This is the case of Proposition 1 and a logarithmic dependence is obtained there. However, the almost-symmetry condition holds there for the entire Hilbert–Schmidt unit ball, while, in Wigner's theorem, the almost-symmetry condition is required to hold only for the non-convex space of normalized hermitian rank-one projections.

Funding: This research received no external funding.

Acknowledgments: The author would like to thank Marius Junge, Willian Corrêa and Cambyse Rouzé for valuable discussions. Furthermore, the author wishes to thank the Institut Henri Poincaré in Paris and the organizers of the trimester on "Analysis in Quantum Information Theory" (IHP17). This work was supported by the German Research Foundation (DFG) and the Technical University of Munich (TUM) in the framework of the Open Access Publishing Program.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kalton, N.J.; Peck, N.T. Twisted sums of sequence spaces and the three space problem. *Trans. Amer. Math. Soc.* **1979**, *255*, 1–30. [CrossRef]
2. Brudnyi, Y.; Kalton, N.J. Polynomial approximation on convex subsets of $\mathbb{R}^n$. *Constr. Approx.* **2000**, *16*, 161–200. [CrossRef]
3. Kalton, N.J. A remark on quasi-isometries. *Proc. Am. Math. Soc.* **2003**, *131*, 1225–1231. [CrossRef]

4. Jung, S.-M. *Hyers-Ulam-Rassias Stability of Functional Equations in Nonlinear Analysis*; Springer: New York, NY, USA, 2011; Volume 48.
5. Wigner, E.P. *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*; Fredrik Vieweg und Sohn: Braunschweig, Germany, 1931; pp. 233–236.
6. Molnár, L.; Nagy, G.; Szokol, P. Maps on density operators preserving quantum f-divergences. *Quantum Inf. Process.* **2013**, *12*, 2309–2323. [CrossRef]
7. Aubrun, G.; Szarek, S.J. *Alice and Bob Meet Banach: The Interface of Asymptotic Geometric Analysis and Quantum Information Theory*; American Mathematical Society: Providence, RI, USA, 2017; Volume 223.
8. Cuesta, J.; Wolf, M.M. Are almost-symmetries almost linear? *J. Math. Phys.* **2019**, *60*, 082101. [CrossRef]
9. Tomczak-Jaegermann, N. *Banach–Mazur Distances and Finite-Dimensional Operator Ideals*; Longman Scientific & Technical: Harlow, UK, 1989.
10. Castillo, J.M.F.; González, M. *Three-Space Problems in Banach Space Theory*; Springer: Berlin/Heidelberg, Germany, 1997.
11. Kalton, N.J.; Peck, N.T.; Roberts, J. *An F-Space Sampler*; Cambridge University Press: Cambridge, UK, 1984; Volume 89.
12. Enflo, P.; Lindenstrauss, J.; Pisier, G. On the three space problem. *Math. Scand.* **1975**, *36*, 199–210. [CrossRef]
13. Maurey, B. Un théorème de prolongement. *C. R. Acad. Sc. Paris* **1974**, *279*, 329–332.
14. Albiac, F.; Kalton, N.J. *Topics in Banach Space Theory*, 1st ed.; Springer: New York, NY, USA, 2006.
15. Cabello Sánchez, F.; Castillo, J.M.F. Duality and twisted sums of Banach spaces. *J. Funct. Anal.* **2000**, *175*, 1–16. [CrossRef]
16. Benyamini, Y.; Lindenstrauss, J. *Geometric Nonlinear Functional Analysis*; American Mathematical Society: Providence, RI, USA, 2000; Volume 1.
17. Figiel, T.; Lindenstrauss, J.; Milman, V.D. The dimension of almost spherical sections of convex bodies. *Acta Math.* **1977**, *139*, 53–94. [CrossRef]

B Further articles

B.1 A stable quantum Darmois-Skitovich theorem

A stable quantum Darmois-Skitovich theorem

Javier Cuesta

Gaussian bosonic states are quantum states that model, in a good approximation, systems with quadratic bosonic Hamiltonian. They are a non-commutative analogue of the normal distribution in classical probability and therefore have a distinguish role in continuous variable quantum Information. There are many abstract characterizations of Gaussian bosonic states (GBS): assuming that the state is pure, GBS are the only states with positive Wigner function; GBS maximize the von Neumann entropy among states with the same second moments; GBS are the states whose quantum characteristic function is a Gaussian. In this work, we study a characterization of GBS in terms of a simple symmetry and show that this characterization is stable.

We start our work with a description of a non-commutative analogue of a classical theorem of Darmois and Skitovich. The latter is a generalization to more than two random variables of the following: if a pair of random variables is independent in two different coordinate systems, then the random variables are necessarily normally distributed. This can be understood as a symmetry where the independence of the random variables is preserved after a linear transformation. It turns out that we have a similar characterization in the quantum case. If a pair of independent quantum states remains independent after the action of a beam splitter transformation, then the states are necessarily Gaussian and with equal second moments. This in turn implies that GBS with equal covariance matrix are the only fixed points of a beam splitter transformation. We provide a new short and rigorous proof of this previously known fact. The main contribution of our work is the stability of this symmetry characterization of GBS. Namely, that if the output of a beam splitter of two incoming independent states is almost independent, then these input states can be approximated by Gaussian states. Moreover, their respective covariance matrix have to be close to each other. We give our bounds in terms of the Hilbert-Schmidt norm and present a first estimate of the stability constants of this problem which have a physical interpretation. Our stability result in terms of a p-norm was not known before, even in the classical scenario where the stability is known to hold in only in a weak sense.

I am the single author of this article and was thus solely involved in all parts of it. The idea for this project was the result of many discussions with my doctoral supervisor Prof. Dr. Michael M. Wolf.

Permission to include:

Javier Cuesta
A stable quantum Darmois-Skitovich theorem
Submitted to J. Math. Phys. (2019).

Permissions - AIP Publishing LLC

Reusing AIP Publishing Content

Permission from AIP Publishing is required to:

- republish content (e.g., excerpts, figures, tables) if you are not the author
- modify, adapt, or redraw materials for another publication
- systematically reproduce content
- store or distribute content electronically
- copy content for promotional purposes

To request permission to reuse AIP Publishing content, use RightsLink® for the fastest response or contact AIP Publishing directly at rights@aip.org and we will respond within one week:

For RightsLink, use Scitation to access the article you wish to license, and click on the Reprints and Permissions link under the TOOLS tab. (For assistance click the "Help" button in the top right corner of the RightsLink page.)

To send a permission request to rights@aip.org, please include the following:

- Citation information for the article containing the material you wish to reuse
- A description of the material you wish to reuse, including figure and/or table numbers
- The title, authors, name of the publisher, and expected publication date of the new work
- The format(s) the new work will appear in (e.g., print, electronic, CD-ROM)
- How the new work will be distributed and whether it will be offered for sale

Authors do **not** need permission from AIP Publishing to:

- quote from a publication (please include the material in quotation marks and provide the customary acknowledgment of the source)
- reuse any materials that are licensed under a Creative Commons CC BY license (please format your credit line: "Author names, Journal Titles, Vol.#, Article ID#, Year of Publication; licensed under a Creative Commons Attribution (CC BY) license.")
- reuse your own AIP Publishing article in your thesis or dissertation (please format your credit line: "Reproduced from [FULL CITATION], with the permission of AIP Publishing")
- reuse content that appears in an AIP Publishing journal for republication in another AIP Publishing journal (please format your credit line: "Reproduced from [FULL CITATION], with the permission of AIP Publishing")
- make multiple copies of articles–although you must contact the Copyright Clearance Center (CCC) at www.copyright.com to do this

Reusing Content Published by Others

To request another publisher's permission to reuse material in AIP Publishing articles, please use our Reuse of Previously Published Material form. (We require documented permission for all reused content.)

Reuse of Previously Published Material Form (pdf)

Unless the publisher requires a specific credit line, please format yours like this:

Reproduced with permission from J. Org. Chem. 63, 99 (1998). Copyright 1998, American Chemical Society.

You do not need permission to reuse material in the public domain, but you should still include an appropriate credit line which cites the original source.

Gmail **Javier Cuesta <jav.cuesta@gmail.com>**

JMP: Receipt of Manuscript MS #JMP19-AR-01166

jmp-edoffice@aip.org <jmp-edoffice@aip.org> Thu, Aug 1, 2019 at 5:55 PM
Reply-To: jmp-edoffice@aip.org
To: ga52giv@mytum.de
Cc: jav.cuesta@gmail.com

Dear Mr. Cuesta,

On 01-Aug-2019, we received the manuscript entitled: "A stable quantum Darmois-Skitovich theorem".
Your manuscript is assigned the Manuscript #JMP19-AR-01166.

The order of the authors on your submitted manuscript is as follows: Javier Cuesta.
Please note that if this order is changed or if there are any deletions or additions of authors, each author must provide approval for this change.

Your manuscript is being processed by the editorial office. You will receive another email soon indicating whether your manuscript has passed the submission check or if modifications are needed before it can enter the editorial process.

Currently, our records indicate that the ORCID ID for your account is 0000-0001-7730-1213. You may use the link below if you need to update your ID:
https://jmp.peerx-press.org/cgi-bin/main.plex?el=A5G6GMTB3A7Hehj1Bh4B9ftdHOSDo6IuYHI1wL4ZfaKUAY
You may check on the status of this manuscript by selecting the "Check Manuscript Status" link at the following URL:

https://jmp.peerx-press.org/cgi-bin/main.plex?el=A2G4GMTB7A3Hehj6F6A9ftdtKqwuR1cTLbOqYXgpUJkgY
Thank you for submitting your work to Journal of Mathematical Physics.

Sincerely,

Editorial Office
Journal of Mathematical Physics

AIP Publishing
1305 Walt Whitman Road
Suite 300
Melville, NY 11747-4300 USA
phone: +1-516-576-2329
e-mail: jmp-edoffice@aip.org

A stable quantum Darmois-Skitovich theorem

Javier Cuesta*
Department of Mathematics, Technische Universität München, 85748 Garching, Germany and
Munich Center for Quantum Science and Technology (MCQST), München, Germany
(Dated: July 12, 2019)

The Darmois-Skitovich theorem is a simple characterization of the normal distribution in terms of the independence of linear forms. We present here a non-commutative version of this theorem in the context of Gaussian bosonic states and show that this theorem is stable under small errors in its underlying conditions. An explicit estimate of the stability constants which depend on the physical parameters of the problem is given.

CONTENTS

I. INTRODUCTION AND SUMMARY OF RESULTS

Among all characterizations of the normal distribution, the ones concerning the independence of linear forms stand out because of their simplicity. The landmark result of such classical characterizations is due to Darmois [1] and Skitovich [2]. Their theorem is a generalization to $n-$random variables and arbitrary coefficients of the following fact: if X, Y are independent real-valued random variables with $X+Y$ and $X-Y$ independent, then X and Y are normally distributed with the same variance (see Theorem 4). We will be interested in studying a quantum (read as non-commutative) version of the Darmois-Skitovich theorem, which we now write shortly as DS theorem. In this case, the role of the normal distribution is taken by Gaussian bosonic states: quantum states whose statistics is completely determined by the knowledge of the first and second moments and whose canonical observables obey the bosonic commutation relations.

Particularly noteworthy is the fact that the quantum DS theorem has a clear physical realization. Consider an arbitrary product state that passes through a beam-splitter as in figure 1. If the output state of the beam-splitter is also a product state, then the input states are Gaussian bosonic states with the *same* second moments. This is the content of the quantum DS-theorem. Mathematically, the content of the quantum DS theorem is given on theorem 7.

* j.cuesta@tum.de

That there does not exist any two copies of identical non-Gaussian states fulfilling this is by no means trivial, since the action of a beam splitter does not create second-moment cross-correlations for an identical product of quantum states (Gaussian and non-Gaussian). This operational characterization of Gaussian states was already known [3], however without a direct reference to the DS theorem. We show this characterization for a general $n-$mode Gaussian bosonic state by means of the DS theorem. This has the advantage of a much clear statistical interpretation and a simpler proof. Additionally, we show that a beam splitter is the only non-trivial linear operation that can have a factorizable output for *all* identical input states (Lemma 6). The latter places the beam splitter as the basic element for detecting non-Gaussianity.

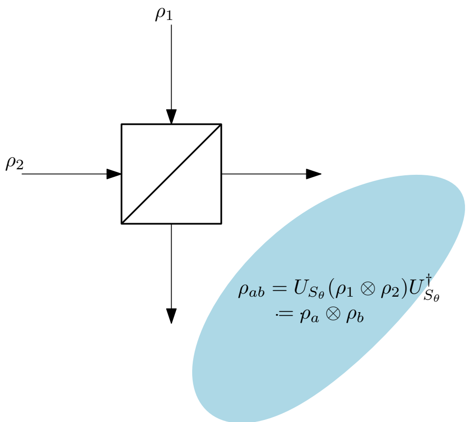


FIG. 1. Quantum Darmois-Skitovich theorem: let U_S be the unitary operator corresponding to the action of a non-trivial beam-splitter transformation. The output state $\rho_{ab} := U_S(\rho_1 \otimes \rho_2)U_S^*$ is a product state if and only if ρ_1 and ρ_2 are Gaussian bosonic states with the same moments.

Of course in real life we cannot completely guarantee that two states are totally independent. Therefore it is crucial to study how stable the DS theorem is. This means, how does the conclusion of the quantum DS theorem changes, when we assume that the output state is not exactly a product, but is *approximately close* to a product state. Our main result is a proof of the stability of the quantum DS theorem for quantum states whose all statistical moments in position and momentum, including mixed moments, exist and are finite. Such states are described by the set of Schwartz density operators [4]. For independent input states whose output from a beam splitter is close in trace norm to a product state, we show that they are close in Hilbert-Schmidt norm to their respective Gaussian counterpart (i.e. the Gaussian state which has the same first and second moments). Moreover, the corresponding second moments of the input states have to be approximately close as well. A precise mathematical statement of this result is given in theorem 9. We make an effort to present explicit stability constants which reflect the effect of the the physical parameters of the problem. These explicit constants contain information of how the problem can become unstable, and have not been estimated before in either the classical or quantum case. The robustness of the quantum DS theorem depends on the transmitivity of the beam-splitter, the number of modes and the largest fourth moment of the output state.

The layout of the paper is as follows. In the next subsection we give the basic definitions and results in continuous-variable quantum information that will be used. Section II contains the main result. We give a simple proof for the characterization of Gaussian bosonic states using the DS theorem and then proceed to state the stability of the DS theorem. Section II B introduces and summarizes the main properties of Schwartz operators. In section III C we give a full proof of the stability of the DS theorem. Finally, in section III D we show some auxiliary lemmas and in the appendix V some explicit bounds are rigorously computed using the properties of Schwartz operators. These bounds will be use for the estimate of the constants appearing in the stability of the DS theorem 9.

A. Notation and preliminaries

We will be entirely concerned with continuous-variable systems with a discrete number n of modes. We denote by

$$R := (Q_1, P_1, \ldots, Q_n, P_n), \tag{1}$$

the vector of canonical operators for a quantum system and $R_k, k = 1, \dots 2n$ its components. Here $Q_l, P_l, l = 1, \dots n$ act on the $l-$tensor factor of the Fock space $\mathcal{H} = \bigotimes_{k=1}^n L^2(\mathbb{R})$ where $L^2(\mathbb{R})$ denotes the space of Lebesgue square integrable functions on $\mathbb{R}$. The *canonical commutation relations* (CCR) are defined by

$$[R_k, R_l] = i\sigma_{kl}, \tag{2}$$

where σ_{ij} are the entries of the symplectic matrix

$$\sigma = \bigoplus_{i=1}^{n} \omega \quad \text{with} \quad \omega := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3}$$

We frequently use the shorthand notation $R_\xi := \xi \cdot \sigma R, \qquad \xi \in \mathbb{R}^{2n}$. The phase-space description of a quantum state ρ is determined by the *characteristic function* $\chi : \mathbb{R}^{2n} \to \mathbb{C}$ defined by

$$\chi(\xi) := \mathrm{Tr}[W_\xi \rho], \tag{4}$$

where $W_\xi = e^{i\xi\cdot\sigma R}$ is the so-called Weyl operator. The CCR are encoded in the Weyl relation

$$W_\xi W_\eta = e^{-\frac{i}{2}\xi\cdot\sigma\eta}\, W_{\xi+\eta}, \qquad \xi, \eta \in \mathbb{R}^{2n}. \tag{5}$$

The name of characteristic function for the map in Eq. (4) comes from an analogy with the classical characteristic function which is the Fourier transform of a probability distribution. In fact by taking a fixed direction in phase space we recover the classical characteristic function and from there, we can "import" all the known results of the classical world. This indeed, will be used in order to give a simple proof of the characterization of Gaussian bosonic states. The condition for a function $\chi : \mathbb{R}^{2n} \to \mathbb{C}$ to be a bona-fide quantum characteristic function is the property of sigma-positiveness. For clarity we state these results and refer the reader to Ref. [5, section 5.4] for a proof.

Theorem 1 (Quantum Bochner-Khinchin)**.** *For $\chi : \mathbb{R}^{2n} \to \mathbb{C}$ to be a characteristic function of a quantum state, the following conditions are necessary and sufficient:*

1. *$\chi(0) = 1$ and χ is continuous at $\xi = 0$,*
2. *χ is σ-positive definite, i.e. for any $m \in \mathbb{N}$, any set $\{\xi_1, \xi_2, \dots, \xi_m\}$ of vectors in $\mathbb{R}^{2n}$, and any set $\{c_1, c_2, \dots, c_m\}$ of complex numbers*

$$\sum_{k,l=1}^{m} c_k \overline{c_l}\; \chi(\xi_k - \xi_l)\; e^{\frac{i}{2}\xi_k\cdot\sigma\xi_l} \geq 0 \tag{6}$$

Corollary 2 (Classical Marginals)**.** *Let $\chi(\xi)$ be the characteristic function of a quantum state. Then for every fixed $\xi \in \mathbb{R}^{2n}$ the function*

$$\mathbb{R} \ni t \mapsto \chi(t\xi),$$

is a classical characteristic function, i.e. the Fourier transform of a classical probability distribution.

As in the classical case, the characteristic function is a moment generating function. The *displacement vector* is defined by the entries $d_k := \mathrm{Tr}[\rho R_k]$ and we say that the state is *centered* if $d = 0$. The *covariance matrix* (CM) is defined by the matrix entries $\varGamma_{kl} = \mathrm{Tr}[\rho\{R_k - d_k, R_l - d_l\}]$. In order that $\varGamma$ corresponds to a genuine quantum CM the CCR impose the further condition [6, 7] $\varGamma + i\sigma \geq 0$, which is nothing but the uncertainty principle expressed in a coordinate-free form.

A *Gaussian bosonic state* is defined as a state with a Gaussian characteristic function

$$\chi(\xi) = \exp[-\frac{\xi\cdot\varGamma\xi}{4} + i\xi\cdot d]. \tag{7}$$

We write χ_ρ to emphasize that χ is the characteristic function of the state ρ. We denote by $M(2n, \mathbb{R})$ and $Sp(4n, \mathbb{R})$ the set of $2n \times 2n$ matrices with real entries and the group of $4n \times 4n$ symplectic matrices with real entries, respectively. The latter is defined as the group of matrices $S \in M(2n, \mathbb{R})$ such that $S\sigma S^T = \sigma$.

Unitary Gaussian operations, i.e. unitary evolutions coming from quadratic Hamiltonians in P and Q, are described by symplectic transformations [8]. These operations have the property that

$$\chi_{(U_S \rho U_S^*)}(\xi) = \chi_\rho(S^T \xi), \tag{8}$$

where U_S is a unitary operation associated to the symplectic transformation S (strictly speaking U_S is determined up to a phase, however this ambiguity disappears in the conjugation $U_S \cdot U_S^*$). The unitary evolution of a Gaussian state is completely determined by the new displacement vector $d' = Sd$ and CM, $\Gamma' = S\Gamma S^T$.
A one mode *non-trivial beam splitter* transformation is the one corresponding to the symplectic transformation

$$S = \begin{pmatrix} \cos\theta\, \mathbb{1}_2 & \sin\theta\, \mathbb{1}_2 \\ -\sin\theta\, \mathbb{1}_2 & \cos\theta\, \mathbb{1}_2 \end{pmatrix}, \qquad \theta \neq m\pi/2, \quad m = 0, 1, 2, \ldots$$

It corresponds to a unitary evolution where the Hamiltonian is

$$H = \frac{\theta}{4}(a_1^* a_2 + a_2^* a_1),$$

with $a_j = (Q_j + iP_j)/\sqrt{2}, a_j^* = (Q_j - iP_j)/\sqrt{2}$, $j = 1, 2$ the creation and annihilation operators. A *local transformation* acts in each separated mode and corresponds therefore to transformations that can be written as $S = \bigoplus_{k=1}^n S_k$. In the context of quantum optics, examples of local transformations are phase-shifts and one-mode squeezing transformations.

The *Wigner phase space distribution* is defined to be the (symplectic) Fourier transform of the characteristic function

$$\mathcal{W}(\eta) = \frac{1}{(2\pi)^n} \int e^{i\eta\cdot\sigma\xi}\, \chi(\xi)\, d\xi. \tag{9}$$

Its importance lies in the fact that, due to corollary 2, all one-dimensional marginals are positive distributions in phase space, which can be associated to the usual probability distributions e.g. on position and momentum of a state ρ.

We write A^* for the adjoint operator of A and $\|\cdot\|$ for the uniform norm. The trace norm is defined as $\|A\|_1 = \operatorname{Tr}\sqrt{A^*A}$ and the Hilbert-Schmidt (HS) norm $\|A\|_2 = (\operatorname{Tr}[A^*A])^{1/2}$. We have the order $\|\cdot\| \leq \|\cdot\|_2 \leq \|\cdot\|_1$. These norms are in fact unitarily invariant and $\|A^*\|_p = \|A\|_p$ for $p = 1, 2$. The usual norm in $L^2(\mathbb{R}^{2n})$ will be denoted by $\|\cdot\|_{L^2(\mathbb{R}^{2n})}$. We use sometimes the Dirac notation for a vector $|\phi\rangle \in \mathcal{H}$ and the inner product notation $\langle\phi|\varphi\rangle$. The commutator and anticommutator are written as $[\cdot,\cdot]$ and $\{\cdot,\cdot\}$ respectively. The space of bounded operators on the Hilbert space $\mathcal{H}$ is denoted by $\mathfrak{B}(\mathcal{H})$.

The inverse relation of Eq. (4) is called the *Weyl transform*

$$T = \frac{1}{(2\pi)^n} \int \operatorname{Tr}[W_\xi T] W_{-\xi} d\xi, \tag{10}$$

where the integral converges weakly for any Hilbert-Schmidt operator T. This is a consequence of the *quantum Parseval theorem* [5] which due to its importance we state here.

Theorem 3 (Quantum Parseval relation)**.** *Let $\{W_\xi\}$ be a strongly continuous and irreducible Weyl systems acting on the Hilbert space $\mathcal{H}$ with respective phase space $X \simeq \mathbb{R}^{2n} \ni \xi$. Then $T \mapsto \operatorname{Tr}[W_\xi T]$ extends uniquely to an isometric map from the Hilbert space of Hilbert-Schmidt class operators on $\mathcal{H}$ onto $L^2(X)$, such that*

$$\operatorname{Tr} T_1^* T_2 = \frac{1}{(2\pi)^n} \int \overline{\operatorname{Tr}[W_\xi T_1]}\, \operatorname{Tr}[W_\xi T_2] d\xi. \tag{11}$$

This theorem also implies that Eq. (4) is also valid for T Hilbert-Schmidt. The map $\xi \mapsto \operatorname{Tr} W_\xi T$ is called the *inverse Weyl transform* of T; being the characteristic function the special case T a density operator.
We will be using repeatedly the following trace inequalities. If B is a bounded operator and T a trace-class operator, then a particular case of Hölder's inequality states

$$\operatorname{Tr} BT \leq \|B\|\, \|T\|_1\, .$$

Let T_1, T_2 be two Hilbert-Schmidt operators. The trace operator version of the Cauchy-Schwarz inequality is

$$\operatorname{Tr} T_1 T_2 \leq \|T_1 T_2\|_1 \leq \|T_1\|_2 \|T_2\|_2 .$$

While it is true that the Hilbert-Schmidt norm is often used out of convenience, it also has operational interpretation, making it preferable for some tasks. This includes equality testing and state discrimination with fixed or random measurements [9]. Furthermore, the Hilbert-Schmidt norm can be a good measure to quantify the difference between two quantum states in quantum optics. There, the Wigner functions of infinite dimensional quantum states are accessible by means of Tomography. The difference between two quantum states is quantified by the Hilbert-Schmidt distance of the respective Wigner functions.

II. MAIN RESULT

In the next subsections we present the quantum version of the DS theorem and our main stability result. The detailed proof of the stability of the DS theorem is presented in section III C.

A. Quantum Darmois-Skitovich theorem

We are interested in a quantum analogue of the following theorem.

Theorem 4 (*Darmois-Skitovich*)**.** *Let $X_1, \ldots, X_n$ ($n \geq 2$) be independent random variables and $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R} \setminus \{0\}$. If the two linear forms*

$$Y_1 = \sum_i a_i X_i \quad \text{and} \quad Y_2 = \sum_i b_i X_i \quad \text{are independent,} \tag{12}$$

then X_i is normally distributed.

Different proofs and the history of the classical DS theorem can be found in p. 78 in Ref. 10 and in Ref. 11. Our setup for the quantum version is the following. We consider two $n-$mode quantum states $\rho_1, \rho_2 \in \mathfrak{B}(\mathcal{H})$ with respective canonical operators

$$\begin{aligned} R_1 &= (Q_1, P_1, \ldots, Q_n, P_n) \\ R_2 &= (Q_{n+1}, P_{n+1}, \ldots, Q_{2n}, P_{2n}) \end{aligned} \tag{13}$$

and write $R = (R_1, R_2)$. We assume that ρ_1 and ρ_2 are independent so that their state in $\mathfrak{B}(\mathcal{H} \otimes \mathcal{H})$ is a product state $\rho_1 \otimes \rho_2$. We refer to ρ_1 and ρ_2 as input states.

The action of producing linear forms of random variables can be mimiced by (Gaussian) unitary evolutions U_S. These unitary evolutions are generated by Hamiltonians that are quadratic expressions in the canonical operators. Moreover[8], the unitary evolution $U_S \in \mathfrak{B}(\mathcal{H} \otimes \mathcal{H})$ is associated with a symplectic transformation $S \in Sp(4n, \mathbb{R})$. In other words, the linear trasformation

$$R \mapsto SR \qquad\qquad S \in Sp(4n, \mathbb{R}), \tag{14}$$

corresponds to a unitary evolution $\rho \mapsto U_S \rho U_S^*$.

In order to obtain an analogue of Eq. (12), we classify the set of unitaries U_S which produce a bipartite independent output, i.e. such that

$$U_S(\rho_1 \otimes \rho_2) U_S^* = \rho_a \otimes \rho_b, \tag{15}$$

where $\rho_a, \rho_b \in \mathfrak{B}(\mathcal{H})$ are $n-$mode quantum states. This is equivalent to classifying the respective set of symplectic transformation for which Eq. (15) holds. If we are to expect that U_S preserves independence, the transformation S should at least preserve uncorrelated inputs (a generally weaker condition than independence which only deals with the second moments). Furthermore, acting locally on each state or swapping them are trivial operations which preserve independence for arbitrary states. So we need to consider other operations in order to obtain a meaningful statement for the quantum DS theorem. The following two lemmas show in fact that there is only one non-trivial symplectic transformation for our setup.

Lemma 5. *Let $S \in Sp(4n,\mathbb{R})$ be such that*

$$S\begin{pmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{pmatrix} S^T = \begin{pmatrix} \cdot & 0 \\ 0 & * \end{pmatrix} \quad \text{for all CM} \quad \Gamma_1, \Gamma_2 \in \mathbb{R}^{2n\times 2n}.$$

Here $*, \cdot$ denote any CM $\in \mathbb{R}^{2n\times 2n}$. Then S is either of the form $\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$ or $\begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix}$ with $A, B, C, D \in Sp(2n,\mathbb{R})$.

If we consider identical inputs we obtain:

Lemma 6. *Let $S \in Sp(4n,\mathbb{R})$ be such that*

$$S\begin{pmatrix} \Gamma & 0 \\ 0 & \Gamma \end{pmatrix} S^T = \begin{pmatrix} \cdot & 0 \\ 0 & * \end{pmatrix} \quad \text{for all CM} \quad \Gamma \in \mathbb{R}^{2n\times 2n}. \tag{16}$$

Here $*, \cdot$ denote any CM $\in \mathbb{R}^{2n\times 2n}$. Then

$$\begin{aligned} S &= \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \frac{1}{\sqrt{1+\alpha^2}} \begin{pmatrix} \mathbb{1}_{2n} & \alpha\,\mathbb{1}_{2n} \\ -\alpha\,\mathbb{1}_{2n} & \mathbb{1}_{2n} \end{pmatrix}, \\ &= \begin{pmatrix} 0 & X \\ Y & 0 \end{pmatrix} \frac{1}{\sqrt{1+\gamma^2}} \begin{pmatrix} \mathbb{1}_{2n} & -\gamma\,\mathbb{1}_{2n} \\ \gamma\,\mathbb{1}_{2n} & \mathbb{1}_{2n} \end{pmatrix}, \end{aligned} \tag{17}$$

where $X, Y \in Sp(2n,\mathbb{R})$ and $\alpha, \gamma \in \mathbb{R} \cup \{\pm\infty\}$.

Thus the only non-trivial linear transformation in Eq. (14) is of the form of Eq. (17). We discard the trivial operations and set $\alpha = \tan\theta$ in Eq. (17)

$$S_\theta = \begin{pmatrix} \cos\theta\,\mathbb{1}_{2n} & -\sin\theta\,\mathbb{1}_{2n} \\ \sin\theta\,\mathbb{1}_{2n} & \cos\theta\,\mathbb{1}_{2n} \end{pmatrix}, \qquad \theta \neq m\pi/2, \quad m = 0,1,2,\ldots, \tag{18}$$

which is the symplectic transformation associated to a ($n-$mode) beam splitter operation. We refer to the latter operation as a *non-trivial beam splitter* transformation.

Although for every covariance matrix Γ, $S_\theta(\Gamma \oplus \Gamma)S_\theta^T = (\Gamma \oplus \Gamma)$, it turns out from the DS theorem that there does not exist a non-Gaussian state ρ such that $U_{S_\theta}(\rho \otimes \rho)U_{S_\theta}^* = \rho \otimes \rho$. See figure 1.

Theorem 7 (Quantum Darmois-Skitovich)**.** *Let U_S be the unitary operation corresponding to a non-trivial beam splitter Eq. 18. Consider the state $\rho_{ab} = U_S(\rho_1 \otimes \rho_2)U_S^*$ obtained after the unitary evolution of an arbitrary product state. If the output state is a product state $\rho_{ab} = \rho_a \otimes \rho_b$, then ρ_1 and ρ_2 are Gaussian bosonic states with the same CM but not necessarily same displacement vector.*

Due to the 1–1 correspondence between quantum states and their characteristic function we have the following consequence.

Corollary 8. *Let χ_1 and χ_2 be the characteristic function of the quantum states ρ_1 and ρ_2 respectively, which have finite second moments. If for a fixed $\theta \neq m\pi/2, \quad m = 0,1,2,\ldots$ the characteristic functions satisfy the functional equation*

$$\chi_1(\cos\theta\xi_1 + \sin\theta\xi_2)\chi_2(\cos\theta\xi_2 - \sin\theta\xi_1) = \chi_1(\cos\theta\xi_1)\chi_1(\sin\theta\xi_2)\chi_2(\cos\theta\xi_2)\chi_2(-\sin\theta\xi_1), \tag{19}$$

for all $\xi_1, \xi_2 \in \mathbb{R}^{2n}$, then ρ_1 and ρ_2 are Gaussian bosonic states with the same CM but not necessarily same displacement vector.

An immediate proof of the quantum DS theorem can be obtained from its corresponding classical result and corollary 2. We recall that the latter corollary tells us that we always obtain a positive Wigner function (a classical probability distribution) from the quantum characteristic function whenever we move through a fixed direction in phase space.

Proof Theorem 7. Using Eq. (8) the evolution of the input states can be expressed in terms of characteristic functions as Eq. (19). We fixed the direction $\xi_1 = \xi_2 = \xi$ and parametrize $\xi_1 = t\xi$, $\xi_2 = s\xi$ with $t, s \in \mathbb{R}$. Moreover, we introduce the classical characteristic functions $\chi_j(u) := \chi_j(u\xi)$, $j = 1,2$ with $u \in \mathbb{R}$ so that Eq. (19) reads

$$\chi_1(\cos\theta t + \sin\theta s)\chi_2(\cos\theta s - \sin\theta t) = \chi_1(\cos\theta t)\chi_1(\sin\theta s)\chi_2(\cos\theta s)\chi_2(-\sin\theta t).$$

This last equation is the functional version of the classical DS theorem (c.f. Eq. 8.7 of section XV.9 in Ref. 10). From this classical result it follows that χ_1 and χ_2 are one dimensional Gaussian characteristic functions with the same variance. We can compute the moments by taking derivates of the characteristic function to obtain that

$$\chi_j(t\xi) = \exp[-\frac{t^2}{4}(\xi \cdot \Gamma\xi) + it\xi \cdot d_j] \qquad j = 1, 2.$$

The result follows with $t = 1$. □

B. Stability

In the last section we presented an exact version of the DS theorem which brings naturally an operational characterization of Gaussian bosonic states. There, exact factorizability of the output state is assumed. In practice, it is impossible to assure such thesis since there are always errors in the measurements and therefore the validity of the result is not completely clear in real life. Moreover, any practical application can immediately be ruled out if the conclusion is not robust against small changes in the defined conditions.

We are interested in finding to which extent the results of theorem 7 are affected if the main assumption is not exact but approximately satisfied. It is not always the case that characterizations of the normal distribution are generally stable. Cramer's characterization of the normal distribution states that if the sum $X + Y$ of two random independent random variables X and Y has a normal distribution, then necessarily both X and Y are normal. It turns out that the classical theorem of Cramer is only stable in a weak sense and that it fails to be robust for stronger notions of distance such as the entropic distance or the total variation norm [12].

Before specifying the conditions of the stability of the quantum DS theorem, we briefly comment on the respective classical stability problem [13]. The stability of the classical DS theorem is due to Yu. R. Gabovich [17]. In his work, the word "approximately" is quantified in terms of the closeness of the cumulative distribution functions of the random variables. It is shown (Theorem 3 in Ref. 17) that for approximate independency the considered classical probability distributions are $c(\log\log(1/\varepsilon))^{-1/8}$-close to a normal distribution in the Levy metric. In Gabovich's estimate there is no explicit value of the constant c and therefore it is not known how it depends on the coefficients of the linear forms, neither on the number of random variables involved. We adress this in our stability proof as it contains relevant information of the physical problem such as where instability could raise.

At the level of density operators it was proven [18] that weak operator topology is equivalent to trace-norm topology. Therefore Gabovich result and corollary 2 should imply at least in a qualitative manner that the quantum DS theorem is stable. However, we can obtain a better concrete estimate of the stability of the DS theorem by considering not just the marginals and the classical result, but rather using the entire phase space and the natural restrictions on the quantum characteristic functions. Namely, due to Heisenberg's uncertainty principle there cannot be characteristic functions which are highly concentrated in some regions of phase space. In particular there are no quantum characteristic functions with compact support [20]. This naturally rules out the construction of ill-behaved distributions that can appear in the classical case. With no more preambules, we give a precise statement of our result.

We say that a quantum state ρ_{ab} is an $\varepsilon-$*approximate product state* if there is a product state $\rho_a \otimes \rho_b$ such that

$$\|\rho_{ab} - \rho_a \otimes \rho_b\|_1 \leq \varepsilon, \tag{20}$$

Suppose that two independent states evolve according to the action of a beam splitter, but this time the output state is an $\varepsilon-$approximate product state (see Fig. 2). The following theorem describes the robustness of the quantum DS theorem.

Theorem 9 (Stability DS)**.** *Let U_S be the unitary operation corresponding to a non-trivial beam splitter characterized in Eq. (18) and ρ_1, ρ_2 density operators of two n-mode systems with finite moments of all orders and whose CMs are Γ_1 and Γ_2. Consider the output state $\rho_{ab} = U_S(\rho_1 \otimes \rho_2)U_S^*$ and define by ρ_1', ρ_2' Gaussian bosonic states that have the same CMs and displacement vectors as ρ_1, ρ_2. If for sufficiently small $\varepsilon \in (0,1)$ the output state ρ_{ab} is ε-close to a product state in trace norm, then*

$$\left\|\rho_j - \rho_j'\right\|_2 \leq c_1 \varepsilon^{1/3} + \frac{c_2}{\sqrt{\log(1/\varepsilon)}}, \qquad j = 1, 2, \tag{21}$$

$$\|\Gamma_1 - \Gamma_2\|_2 \leq c_3 \varepsilon^{1/2}, \tag{22}$$

where the constants c_1, c_2 and c_3 (which are made explicit in the proof) depend on the transmitivity θ of the beam-splitter, the number of modes n and the second and fourth absolute moments of the output state ρ_{ab}.

Note that under the conditions of theorem 9, the bound of Eq. (21) also applies to the distance between the Wigner functions of ρ and ρ'. Furthermore, the stability result could in principle be extended to a larger set of quantum states since Schwartz operators are dense in the set of trace class operators (see Lemma 2.5 in [4]).

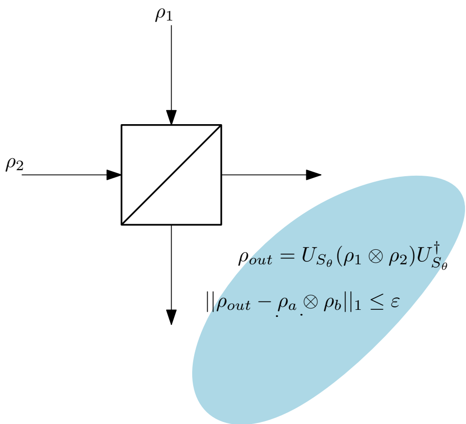


FIG. 2. Stability of the Darmois-Skitovich theorem. In this case we consider the output state to be an $\varepsilon-$approximate product state.

Schwartz operators

We begin this section by reviewing some important facts about Schwartz operators which are the non-commutative analogue of Schwartz functions. The latter are infinite differentiable functions whose derivatives decay faster than any polynomial at infinity. The introduction of Schwartz operators allows us to handle differentiability and boundness problems in an elegant manner and it is for this reason that they play an important technical role in this paper. This class of operators was first introduced in Ref. 4, and the reader may find there a detailed exposition.

In our proof of the stability of the DS theorem, we deal with terms of the form $\mathrm{Tr}[R_k R_l \rho R_s R_r]$ which are a priori not necessarily well-defined on any dense domain of ρ; it may happen that ρ maps outside the domain of R_l. This is a common issue among others when dealing with unbounded operators (see Section 17.2.1 in [19] for some missleading formal manipulations with unbounded operators). There is an immense advantage when working with Schwartz operators since many regular properties for bounded operators become available for unbounded ones (*e.g.* "cycling under the trace" property holds). Indeed, Theorem 13 below plays a decisive role in the calculation of the constants appearing in our proof of the stability of DS theorem.

Definition 10 (Schwartz Operators)**.** *An operator $T \in \mathcal{B}(\mathcal{H})$ is called a Schwartz operator if*

$$\left\|P^\alpha Q^\beta T P^{\alpha'} Q^{\beta'}\right\|_1 < \infty \qquad \text{for all } \alpha, \alpha', \beta, \beta' \in I_n,$$

where $I_n := \{\alpha = (\alpha_1, \dots, \alpha_n) | \alpha_i \in \mathbb{N} \cup \{0\}$ *for all* $i = 1, \dots, n\}$ *is the set of multi-indices, and*

$$Q^\alpha = Q_1^{\alpha_1} \cdots Q_n^{\alpha_n}, \qquad P^\alpha = P_1^{\alpha_1} \cdots P_n^{\alpha_n}. \tag{23}$$

The set of all Schwartz operators will be denoted by $\mathfrak{S}(\mathcal{H})$.

So for Schwartz-density operators all the statistical moments in Q and P exist and are finite. We denote by

$$\mathscr{S}(\mathcal{H}) := \{\rho \in \mathfrak{S}(\mathcal{H}) \mid \rho \text{ is a density operator } \}$$

the space of density operators which are also Schwartz operators. For $\rho \in \mathscr{S}(\mathcal{H})$ we have the following neat characterization

Proposition 11. *Let T be a Hilbert-Schmidt operator. Then T is a Schwartz operator if and only if the respective Weyl transform is a Schwartz function.*

Corollary 12. *A density operator ρ is a Schwartz operator if and only if its characteristic function χ, or Wigner function $\mathcal{W}$, is a Schwartz function. Moreover, the partial trace of a Schwartz-density operator is a Schwartz operator.*

The following Theorem contains the basic properties of Schwartz operators that we use.

Theorem 13. *Let* $\mathcal{H} = L^2(\mathbb{R}^{2n})$ *and* $T \in \mathcal{B}(\mathcal{H})$. *Then*

(i) Let f be a polynomial function on the entries of the vector $R = (Q_1, P_1, \dots, Q_n, P_n)$ and $\{W_\xi\}$ a Weyl system. If $T \in \mathfrak{S}(\mathcal{H})$, then $\operatorname{Tr}[f(R)T] = \operatorname{Tr}[Tf(R)]$. Moreover, $f(R)T \in \mathfrak{S}(\mathcal{H})$ and $\operatorname{Tr}[W_\xi f(R)T] = \operatorname{Tr}[f(R)TW_\xi] = \operatorname{Tr}[TW_\xi f(R)]$.

(ii) If $T \in \mathfrak{S}(\mathcal{H})$, then T is trace-class.

(iii) If $T \in \mathfrak{S}(\mathcal{H})$, then $|T| \in \mathfrak{S}(\mathcal{H})$.

(iv) If $0 < T \in \mathfrak{S}(\mathcal{H})$, then $\sqrt{T} \in \mathfrak{S}(\mathcal{H})$.

(v) If $T \in \mathfrak{S}(\mathcal{H})$ then $T^* \in \mathfrak{S}(\mathcal{H})$.

For Schwartz-density operators we can write explicit formulas for the gradient and Hessian of the characteristic function in terms of a trace: Consider $\{\xi_k\}_{k=1}^{2n}$ any basis in $\mathbb{R}^{2n}$, then the gradient of $\chi(\xi)$, denoted by $\nabla\chi(\xi)$, is defined by the entries

$$\frac{\partial\chi(\xi)}{\partial\xi_k} = \frac{d}{dt}\chi(\xi + t\xi_k)\Big|_{t=0}.$$

The following Lemma generalizes Lemmas 5.4.2 and 5.4.3 in [5].

Lemma 14 (Gradient of the Weyl Operator). *Let T be a Schwartz operator and $\nabla_\eta := \eta \cdot \nabla$. Then the following identities hold*

$$\chi_{R_{\xi_k}T}(\xi) = \left(\frac{1}{2}\xi_k \cdot \sigma\xi - i\frac{\partial}{\partial\xi_k}\right)\chi_T(\xi) \tag{24}$$

$$\chi_{TR_{\xi_k}}(\xi) = \left(-\frac{1}{2}\xi_k \cdot \sigma\xi - i\frac{\partial}{\partial\xi_k}\right)\chi_T(\xi) \tag{25}$$

$$(\nabla_\eta\chi_T)(\xi) = \frac{i}{2}\operatorname{Tr}(\{W_\xi, R_\eta\}T) = \frac{i}{2}\operatorname{Tr}(W_\xi\{R_\eta, T\}) \tag{26}$$

$$(\xi_k \cdot \sigma\xi)\chi_T(\xi) = \operatorname{Tr}([W_\xi, R_\eta]T) = \operatorname{Tr}(W_\xi[R_\eta, T]) \tag{27}$$

Proof of lemma 14. First note that Eq. (26) and Eq. (27) follow from adding and substracting Eq. (24) and Eq. (25) together with Theorem 13(i). We show first that

$$\frac{d}{dt}(\operatorname{Tr} W_{t\eta}T)\Big|_{t=0} = i\operatorname{Tr} R_\eta T. \tag{28}$$

Since T is trace-class (Theorem 13(ii)) we can decompose $T = T_1 + iT_2$ with T_1, T_2 self-adjoint trace-class operators. Moreover, we can write T_1, T_2 as a finite linear combination of positive, trace-class operators and thus from the linearity of the trace we can assume without loss of generality that T is a positive, trace-class operator. From the spectral decomposition $R_\eta = \int x dE(x)$ and the functional calculus we obtain

$$\begin{aligned}\left|\operatorname{Tr}\left(\frac{W_{t\eta}-\mathbb{1}}{it}-R_\eta\right)T\right| &= \left|\int\left(\frac{e^{itx}-1}{it}-x\right)\operatorname{Tr} dE(x)T\right|,\\ &\leq \int\left|\frac{e^{itx}-1}{it}-x\right|\operatorname{Tr} dE(x)T.\end{aligned}$$

Using $\left|\frac{e^{itx}-1}{it}-x\right| \leq 2|x|$, the Cauchy-Schwarz inequality and Theorem 13(iv)

$$\begin{aligned}\int\left|\frac{e^{itx}-1}{it}-x\right|\operatorname{Tr} dE(x)T &\leq 2\operatorname{Tr}|R_\eta|T,\\ &\leq 2\left\|R_\eta^2\sqrt{T}\right\|_2\left\|\sqrt{T}\right\|_2 < \infty.\end{aligned}$$

Hence from the dominated convergence theorem we proved what is required. Now we proceed to prove Eq. (24). From Theorem 13(i,ii) we have that $R_{\xi_k}T$ is trace-class and therefore the Weyl transform exists. Moreover, $\chi_{R_{\xi_k}T}(\xi)$ is Schwartz, hence continuous, as it is the Weyl transform of a Schwartz operator (Proposition 11). So we just need to verify the relation of Eq. (24) as Eq. (25) is similar. This follows directly from Eq. (28) and (5)

$$\begin{aligned}\chi_{R_{\xi_k}T}(\xi) &= -i\frac{d}{dt}\operatorname{Tr} W_\xi W_{t\xi_k}T\Big|_{t=0},\\ &= -i\frac{d}{dt}\left(e^{it\xi_k\cdot\sigma\xi/2}\chi(\xi+t\xi_k)\right)\Big|_{t=0}.\end{aligned}$$

□

We remark that since T is Schwartz, higher order derivatives of $\chi_T(\xi)$ can be written explicitely by using theorem 13(i) and Eq. (26). For instance, the Hessian of $\chi_\rho(\xi)$ where $\rho \in \mathscr{S}(\mathcal{H})$ has entries given by

$$\frac{\partial^2\chi(\xi)}{\partial\xi_k\partial\xi_l} = -\frac{1}{4}\operatorname{Tr}[W_\xi\{R_{\xi_k},\{R_{\xi_l},\rho\}\}]. \tag{29}$$

In particular, if the state ρ has covariance matrix Γ and displacement vector d

$$\nabla\chi(0) = i\sigma d \qquad \text{and} \qquad (\text{Hessian }\chi)(0) = -\sigma\left(\frac{\Gamma}{2}+dd^T\right)\sigma^T. \tag{30}$$

III. PROOFS

A. Proof of lemma 5

We write $S = \begin{pmatrix} A & B \\ C & D\end{pmatrix}$ and express the condition of the vanishing off-diagonal terms

$$A\Gamma_1 C^T + B\Gamma_2 D = 0 \quad \text{for all} \quad \Gamma_1, \Gamma_2 \geq i\sigma. \tag{31}$$

If we fix A, B, C, and D all different from zero such that Eq.(31) is satisfied, we can always find Γ_1 and Γ_2 such that for these choice of submatrices $A\Gamma_1 C^T + B\Gamma_2 D^T \neq 0$. Then each summand in Eq.(31) must be zero. For instance if $A\Gamma_1 C^T = -B\Gamma_2 D^T$ take now $\Gamma_1 \to 2\Gamma_1$ then $A(2\Gamma_1)C^T + B\Gamma_2 D^T = A\Gamma_1 C^T = 0$.
W.l.o.g let us consider $A\Gamma_1 C^T = 0$, the other case will be analogous. We take the singular value decomposition of $A = UDV$ and $C = W\Sigma Z$ where U, V, W, Z are unitaries. From here we choose $\Gamma_1 = V^{-1}PZ^{-T}$ where P is a positive definite matrix (recall that the sigma positive condition $\Gamma_1 \geq i\sigma$ can be always obtained from rescaling a positive matrix). Then $A\Gamma_1 C^T = UDP\Sigma W^T \neq 0$ everytime we choose the proper Γ_1. Consequently $A = 0$, $C = 0$ or both.

Likewise for B and D.
The only matrices that fulfill the symplectic conditions are

$$\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}, \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix}, \tag{32}$$

provided A, B, C and D are symplectic.

B. Proof of lemma 6

First, it should be noted that the given assumptions immediately imply $S \begin{pmatrix} \Gamma & 0 \\ 0 & \Gamma \end{pmatrix} S^T = \begin{pmatrix} \cdot & 0 \\ 0 & * \end{pmatrix}$ for all $\Gamma > 0$ and by continuity for all $\Gamma \geq 0$ and consequently for all $\Gamma = \Gamma^T$. The latter is due to the fact that every symmetric matrix can be decomposed in a semi-definite positive and negative part.
From Lemma 16 we know which is the block structure of S. We consider the case where A, B, C and D are invertible, the other cases will be contained here as we will see. Using part (ii) of Lemma 16 we have

$$A \otimes C = -B \otimes D. \tag{33}$$

We multiply Eq. (33) by $(A^{-1} \otimes \mathbb{1})$ from the left and take the trace in the first component, likewise we multiply Eq. (33) by $(\mathbb{1} \otimes D^{-1})$ and take the trace in the second component to obtain

$$C = \alpha D \quad \text{where} \quad \alpha := \frac{-\operatorname{Tr} A^{-1} B}{2n} \in \mathbb{R}, \tag{34}$$

and

$$B = \beta A \quad \text{where} \quad \beta := \frac{-\operatorname{Tr} D^{-1} C}{2n} \in \mathbb{R}. \tag{35}$$

From equation Eq. (33) we obtain that $\alpha = -\beta$ and that

$$S = \begin{pmatrix} A & \alpha A \\ -\alpha D & D \end{pmatrix}.$$

Moreover, the symplectic constraints on S give us that $A, D \in \frac{1}{\sqrt{1+\alpha^2}} Sp(2n, \mathbb{R})$. We write $A = \frac{1}{\sqrt{1+\alpha^2}} X$ and $D = \frac{1}{\sqrt{1+\alpha^2}} Y$ with $X, Y \in Sp(2n, \mathbb{R})$ to obtain Eq. (17). Finally for $\alpha \neq 0$ we define $\gamma = \frac{1}{\alpha}$ and obtain the remaining equation. The case $\gamma = 0$ is covered by $\alpha \to \pm\infty$ and it gives the swap operation. Clearly $\alpha = 0$ gives the local transformation.

C. Proof of the stability of DS theorem 9

The proof involves a series of steps. We first use Parseval's theorem to express the distance of the quantum states in terms of the L_2-distance of their respective characteristic functions. Next, we show that there is a ball $\mathfrak{B}_r$ around the origin of phase space where the characteristic functions do not vanish. The radius of this ball scales inversely proportional to the largest variance of the input states and proportional to $\log(1/\varepsilon)$. Thus the smaller the error parameter ε the bigger this region is. We then proceed to bound the distance separately on $\mathfrak{B}_r$ and its complement $\mathfrak{B}_r^c$. For the latter region, we exploit the relation between the tails of a distribution and the finiteness of its moments. Inside $\mathfrak{B}_r$, the problem is equivalent to the stability of the Gaussian functional Eq. (19) with restricted convex domain. For that matter, the stability of the Gaussian functional equation is reduced to the stability of the fundamental functional equation, namely the Cauchy functional equation.

We may assumed without lost of generality that the input states (and therefore the output states) are centered. This is justified by the fact that the trace and HS norms are unitarily invariant and the operation of "Gaussification" (the operation on bosonic quantum states which produces Gaussian states with the same first and second moments of the input state) commutes with the displacement operation $\rho \mapsto W_d \rho W_d^*$.

We denote by χ_{ab}, χ_a and χ_b the characteristic functions of ρ_{ab}, ρ_a and ρ_b, respectively. From Lemma 17, we have $\|g\|_1 \leq 3\varepsilon$. Let $\eta_1, \eta_2 \in \mathbb{R}^{2n}$ and R_1, R_2 be vectors of canonical operators as in Eq. (13). We write

$$G(\eta_1,\eta_2) := \mathrm{Tr}[e^{i\eta_1\cdot\sigma R_1} \otimes e^{i\eta_2\cdot\sigma R_2} g] = \chi_{ab}(\eta_1,\eta_2) - \chi_a(\eta_1)\chi_b(\eta_2). \tag{36}$$

Now, using the covariant property of Gaussian unitary operations Eq. (8)

$$\chi_{\rho_{ab}}(\xi) = \chi_{\rho_1\otimes\rho_2}(S_\theta^T \xi),$$

and the fact that $G(\eta_1, 0) = G(0, \eta_2) = 0$, we write for all $\eta_1, \eta_2 \in \mathbb{R}^{2n}$ the dynamical process in terms of characteristic functions as

$$\chi_1(\cos\theta\eta_1 + \sin\theta\eta_2)\chi_2(\cos\theta\eta_2 - \sin\theta\eta_1) = \chi_1(\cos\theta\eta_1)\chi_1(\sin\theta\eta_2)\chi_2(\cos\theta\eta_2)\chi_2(-\sin\theta\eta_1) + G(\eta_1,\eta_2). \tag{37}$$

This last equation resembles the ideal functional equation Eq. (19) plus a new remainder term $G(\eta_1, \eta_2)$. From Hölder's inequality and the definition of G we note that $\|G\| \leq 3\varepsilon$.

We state the following lemma with proof in section III D.

Lemma 15. *Let χ_1 and χ_2 be two characteristic functions with respective density operators ρ_1, ρ_2 and covariance matrices Γ_1, Γ_2. Define for any $\theta \notin \{\mathbb{Z}\frac{\pi}{2}\}$*

$$G(\eta_1,\eta_2) := \chi_1(\cos\theta\eta_1)\chi_1(\sin\theta\eta_2)\chi_2(\cos\theta\eta_2)\chi_2(-\sin\theta\eta_1) - \chi_1(\cos\theta\eta_1 + \sin\theta\eta_2)\chi_2(\cos\theta\eta_2 - \sin\theta\eta_1).$$

Assume $|G(\eta_1, \eta_2)| \leq 3\epsilon$ for all $\eta_1, \eta_2 \in \mathbb{R}^{2n}$, let $\lambda := \frac{1}{2}\max\{\|\Gamma_1\|, \|\Gamma_2\|\}$ be the largest variance of the states ρ_1 and ρ_2 and define

$$r := \sqrt{\frac{1}{\lambda}\log_2\frac{1}{\varepsilon^{1/12}}}. \tag{38}$$

Then for $\eta \in \mathfrak{B}_r := \{\xi \mid \|\xi\|_2 \leq r\}$

$$|\chi_i(\eta)| \geq 12\varepsilon^{1/12} \qquad \textit{i=1,2}. \tag{39}$$

The choice of exponent $1/12$ for ε in Eq. (38) will become clear in the estimates done in section V.

We divide phase space in two separating regions. One where the characteristic functions do not vanish, namely inside the ball $\mathfrak{B}_r := \left\{\xi \mid \|\xi\|_2^2 \leq r^2\right\}$, and its complement which we denote by $\mathfrak{B}_r^c$. So that with the help of Parseval's theorem, we express the distance between the two states as

$$\begin{aligned}(2\pi)^n \|\rho_1 - \rho_1'\|_2^2 &= \|\chi_1 - \Phi\|_2^2 \\ &= \int\limits_{\xi\in\mathfrak{B}_{r/2}} |\chi_1(\xi) - \Phi(\xi)|^2 d\xi + \int\limits_{\xi\in\mathfrak{B}_{r/2}^c} |\chi_1(\xi) - \Phi(\xi)|^2 d\xi.\end{aligned} \tag{40}$$

We compute the bound for $\mathfrak{B}_{r/2}$ and $\mathfrak{B}_{r/2}^c$ separately.

Bound on the region where the characteristic function might vanish

We use $|z_1 - z_2|^2 \leq (|z_1|^2 + |z_2|^2)/2$ for $z_1, z_2 \in \mathbb{C}$ to express the bound as

$$\int\limits_{\xi\in\mathfrak{B}_{r/2}^c} |\chi_1(\xi) - \Phi(\xi)|^2 d\xi \leq \frac{1}{2}\int\limits_{\xi\in\mathfrak{B}_{r/2}^c} |\chi_1(\xi)|^2 d\xi + \frac{1}{2}\int\limits_{\xi\in\mathfrak{B}_{r/2}^c} |\Phi(\xi)|^2 d\xi. \tag{41}$$

For the first term of the RHS of Eq. (41) we use that $1 < 2\left\|\xi\right\|_2/r$ for $\xi \in \mathfrak{B}^c_{r/2}$ so that

$$\int\limits_{\xi\in\mathfrak{B}^c_{r/2}} |\chi_1(\xi)|^2 d\xi \le \frac{4}{r^2} \int\limits_{\xi\in\mathbb{R}^{2n}} \left\|\xi\right\|_2^2 |\chi_1(\xi)|^2 d\xi. \tag{42}$$

Let us denote by $\mathcal{W}(\eta)$ the Wigner function of $|\chi_1(\xi)|^2$ (the product of two characteristic functions is a characteristic function),

$$\mathcal{W}(\eta) = \frac{1}{(2\pi)^{2n}} \int\limits_{\xi\in\mathbb{R}^{2n}} e^{i\eta\cdot\sigma\xi} |\chi_1(\xi)|^2 d\xi. \tag{43}$$

It can be easily verified by direct computation, that the characteristic function $|\chi_1(\xi)|^2$ is centered and has CM $2\varGamma_1$. Moreover, we have

$$\int\limits_{\xi\in\mathbb{R}^{2n}} \left\|\xi\right\|_2^2 |\chi_1(\xi)|^2 d\xi = -(2\pi)^n \sum_{k=1}^{2n} \frac{\partial^2\mathcal{W}(0)}{\partial\eta_k^2}, \tag{44}$$

where η_k are the components of the vector η in an arbitrary but fixed basis.
We use now the representation of the Wigner function in terms of the expectation values of the parity operator[24] $\mathcal{P}$

$$\mathcal{W}(\eta) = \frac{1}{\pi^n}\,\mathrm{Tr}[\rho W_\eta \mathcal{P} W_{-\eta}], \tag{45}$$

where ρ is the density operator corresponding to the characteristic function $|\chi_1(\xi)|^2$. The operator ρ is clearly Schwartz as its characteristic function is a Schwartz function (corollary 12). The parity operator $\mathcal{P}$ is the n-fold tensor product of the parity operators for a single degree of freedom and is the unitary operator that satisfies

$$\begin{aligned}
&\mathcal{P} W_\xi \mathcal{P}^* = W_{-\xi},\\
&\mathcal{P} R_k \mathcal{P}^* = -R_k,\\
&\mathcal{P} = \mathcal{P}^* = \mathcal{P}^{-1}.
\end{aligned}$$

Using Eq. (45) and Eq. (26) we compute

$$\frac{\partial^2\mathcal{W}(0)}{\partial\eta_k\partial\eta_l} = -\frac{2}{\pi^n}\,\mathrm{Tr}[\mathcal{P}\rho\{R_{\eta_l}, R_{\eta_k}\}]. \tag{46}$$

Thus from Eq. (42), (44) and Eq. (46) we find that

$$\int\limits_{\xi\in\mathfrak{B}^c_{r/2}} |\chi_1(\xi)|^2 d\xi \le \frac{2^{n+4}}{r^2} \sum_{k=1}^{2n} \mathrm{Tr}[\mathcal{P}\rho R_k^2]. \tag{47}$$

Hence, we just need to bound the terms $\mathrm{Tr}[\mathcal{P}\rho R_k^2]$. In order to do this, we notice that $\mathcal{P}\rho = \rho\mathcal{P}$. Indeed,

$$\begin{aligned}
\mathcal{P}\rho = \mathcal{P}\frac{1}{(2\pi)^n}\int |\chi(\xi)|^2 W_{-\xi} d\xi &= \frac{1}{(2\pi)^n}\int |\chi(\xi)|^2 W_{\xi} d\xi\ \mathcal{P},\\
&= \frac{1}{(2\pi)^n}\int |\chi(\xi)|^2 W_{-\xi} d\xi\ \mathcal{P} = \rho\mathcal{P}.
\end{aligned}$$

Moreover, from the spectral decomposition of ρ we have $\mathcal{P}\sqrt{\rho} = \sqrt{\rho}\,\mathcal{P}$. Accordingly,

$$\begin{aligned}\mathrm{Tr}[\mathcal{P}\rho R_k^2] &= \mathrm{Tr}[\mathcal{P}\sqrt{\rho}R_k^2\sqrt{\rho}],\\ &\leq \left\|\sqrt{\rho}R_k^2\sqrt{\rho}\right\|_1,\\ &= \mathrm{Tr}[\rho R_k^2].\end{aligned}$$

Here we have used the Cauchy-Schwarz inequality and the cyclicity of the trace that comes from the properties of the Schwartz operator ρ; see Theorem 13 $(i), (ii), (iv)$.

In summary, we have the following bound for the tails of our characteristic function

$$\begin{aligned}\int_{\xi\in\mathfrak{B}^c_{r/2}} |\chi_1(\xi)|^2 d\xi &\leq \frac{2^{n+4}}{r^2}\,\mathrm{Tr}\,\mathit{\Gamma}_1,\\ &= \left(2^{n+4}12\lambda\,\mathrm{Tr}\,\mathit{\Gamma}_1\right)\frac{1}{\log\frac{1}{\varepsilon}}.\end{aligned}$$

Since $\Phi(\xi)$ has the same CM as $\chi(\xi)$ we can use the same bound to obtain

$$\begin{aligned}\frac{1}{(2\pi)^n}\int_{\xi\in\mathfrak{B}^c_{r/2}} |\chi_1(\xi)-\Phi(\xi)|^2 d\xi &\leq \left(\frac{192\lambda\,\mathrm{Tr}\,\mathit{\Gamma}_1}{\pi^n}\right)\frac{1}{\log\frac{1}{\varepsilon}},\\ &\leq \frac{c_2^2}{\log\frac{1}{\varepsilon}},\end{aligned} \tag{48}$$

where

$$c_2 := 8\sqrt{\frac{3\lambda\,\mathrm{Tr}\,\mathit{\Gamma}_{ab}}{\pi^n}}, \tag{49}$$

and $\mathit{\Gamma}_{ab}$ is the CM of the output state ρ_{ab}. Note that since the BS is a passive transformation, the trace of the input CM, $\mathit{\Gamma}_1\oplus\mathit{\Gamma}_2$, is the same as the trace of the output CM $\mathit{\Gamma}_{ab}$. Thus it is clear that $\mathrm{Tr}\,\mathit{\Gamma}_1 \leq \mathrm{Tr}\,\mathit{\Gamma}_{ab}$.

Bound inside the region where χ does not vanish:

We proceed to compute the bound for the first term of the RHS of Eq. (40) following the ideas of Ref. 15–17. For that matter, let $\eta_1,\eta_2\in\mathbb{R}^{2n}$ with $\|\eta_j\|_2 < r/2$ so that $\cos\theta\eta_1+\sin\theta\eta_2, \cos\theta\eta_2-\sin\theta\eta_1\in\mathfrak{B}_r$. Let us take the logarithm (principal branch) on both sides of Eq. (37). We write

$$\Psi_1(\cos\theta\eta_1+\sin\theta\eta_2)+\Psi_2(\cos\theta\eta_2-\sin\theta\eta_1) = \Psi_1(\cos\theta\eta_1)+\Psi_1(\sin\theta\eta_2)+\Psi_2(\cos\theta\eta_2)+\Psi_2(-\sin\theta\eta_1)+Q(\eta_1,\eta_2), \tag{50}$$

where $\Psi_j(\eta) := -\log\chi_j(\eta)$ for $j=1,2$ and

$$Q(\eta_1,\eta_2) := -\log\left(1+\frac{G(\eta_1,\eta_2)}{\chi_1(\cos\theta\eta_1)\chi_1(\sin\theta\eta_2)\chi_2(\cos\theta\eta_2)\chi_2(-\sin\theta\eta_1)}\right). \tag{51}$$

Since $\rho_j, j=1,2$ is Schwartz, we can define continuous vector-valued functions $\phi_j(\xi):\mathfrak{B}_{r/2}\to\mathbb{C}^{2n}$ by

$$\phi_j(\xi) := \nabla\Psi_j(\xi).$$

Note that $\phi_j(\xi), j=1,2$, are in fact conservative vector fields and that $\phi_j(0)=0$ as ρ_j is centered. The gradient of ϕ_j is the Hessian of χ_j (see Lemma 14) and $2\nabla\phi_j(0)=\sigma\mathit{\Gamma}_j\sigma^T$ for $j=1,2$.

Inhomogeneous Cauchy Functional Equation

Next, we want to obtain a functional equation only depending on χ_1 or χ_2. In order to do so, we differentiate Eq. (50) in the direction of η_1 to find

$$\cos\theta\phi_1(\cos\theta\eta_1+\sin\theta\eta_2)-\sin\theta\phi_2(\cos\theta\eta_2-\sin\theta\eta_1) = \cos\theta\phi_1(\cos\theta\eta_1)-\sin\theta\phi_2(-\sin\theta\eta_1)+Q_1(\eta_1,\eta_2), \tag{52}$$

where $Q_1(\eta_1,\eta_2) := \frac{dQ(\eta_1+t\eta_1,\eta_2)}{dt}\Big|_{t=0}$. We evaluate in Eq. (52) $\eta_1 = 0$ to get

$$\cos\theta\phi_1(\sin\theta\eta_2) - \sin\theta\phi_2(\cos\theta\eta_2) = Q_1(0,\eta_2), \tag{53}$$

In a similar fashion we differentiate Eq. (50) in the direction of η_2 and set to zero to obtain

$$\sin\theta\phi_1(\cos\theta\eta_1+\sin\theta\eta_2)+\cos\theta\phi_2(\cos\theta\eta_2-\sin\theta\eta_1) = \sin\theta\phi_1(\sin\theta\eta_2)+\cos\theta\phi_2(\cos\theta\eta_2)+Q_2(\eta_1,\eta_2), \tag{54}$$

$$\sin\theta\phi_1(\cos\theta\eta_1)+\cos\theta\phi_2(-\sin\theta\eta_1) = Q_2(\eta_1,0), \tag{55}$$

where $Q_2(\eta_1,\eta_2) := \frac{dQ(\eta_1,\eta_2+t\eta_2)}{dt}\Big|_{t=0}$.

Now we will be able to decouple ϕ_1 and ϕ_2. First, we substract Eq. (53) from Eq. (52) and Eq. (55) from Eq. (54) to obtain

$$\begin{aligned}[\phi_1(\cos\theta\eta_1+\sin\theta\eta_2)-\phi_1(\cos\theta\eta_1)-\phi_1(\sin\theta\eta_2)] &= \tan\theta[\phi_2(\cos\theta\eta_2-\sin\theta\eta_1)-\phi_2(\cos\theta\eta_2)-\phi_2(-\sin\theta\eta_1)] \\ &\quad+\frac{Q_1(\eta_1,\eta_2)-Q_1(0,\eta_2)}{\cos\theta},\end{aligned} \tag{56}$$

$$\begin{aligned}[\phi_2(\cos\theta\eta_2-\sin\theta\eta_1)-\phi_2(\cos\theta\eta_2)-\phi_2(-\sin\theta\eta_1)] &= -\tan\theta[\phi_1(\cos\theta\eta_1+\sin\theta\eta_2)-\phi_1(\cos\theta\eta_1)-\phi_1(\sin\theta\eta_2)] \\ &\quad+\frac{Q_2(\eta_1,\eta_2)-Q_2(\eta_1,0)}{\cos\theta}.\end{aligned} \tag{57}$$

Thus from Eq. (56) and (57) we find the following inhomogeneous Cauchy equations

$$\begin{aligned}[\phi_1(\cos\theta\eta_1+\sin\theta\eta_2)-\phi_1(\cos\theta\eta_1)-\phi_1(\sin\theta\eta_2)] &= \sin\theta[Q_2(\eta_1,\eta_2)-Q_2(\eta_1,0)] \\ &\quad+\cos\theta[Q_1(\eta_1,\eta_2)-Q_1(0,\eta_2)],\end{aligned} \tag{58}$$

$$\begin{aligned}[\phi_2(\cos\theta\eta_2-\sin\theta\eta_1)-\phi_2(\cos\theta\eta_2)-\phi_2(-\sin\theta\eta_1)] &= \sin\theta[Q_1(\eta_1,\eta_2)-Q_1(0,\eta_2)] \\ &\quad+\cos\theta[Q_2(\eta_1,\eta_2)-Q_2(\eta_1,0)].\end{aligned} \tag{59}$$

Bound on $\mathfrak{B}_{r/2}$

Now that ϕ_1 and ϕ_2 are decoupled, we continue only with ϕ_1 as with ϕ_2 is analogous and the same upper bound is obtained. We recall that the derivative of a vector with respect to a vector can be represented as a matrix. Thus when we differentiate Eq. (58) in the direction of η_2 and evaluate at $\eta_2 = 0$, we obtain the following matrix-valued equation

$$\nabla\phi_1(\cos\theta\eta_1) - \frac{\sigma\Gamma_1\sigma^T}{2} + \frac{Q_{12}(0,0)}{\tan\theta} = Q_{22}(\eta_1) + \frac{Q_{12}(\eta_1)}{\tan\theta}. \tag{60}$$

Here $Q_{12}(\eta_1,0) : \mathbb{R}^{2n} \to \mathbb{C}^{2n\times 2n}$ and $Q_{22}(\eta_1,0) : \mathbb{R}^{2n} \to \mathbb{C}^{2n\times 2n}$ are defined as

$$Q_{12}(\eta_1,0) := \frac{\partial^2 Q(\eta_1+t\eta_1, s\eta_2)}{\partial s\partial t}\Big|_{s=t=0} \qquad \text{and} \qquad Q_{22}(\eta_1) := \frac{\partial^2 Q(\eta_1, s\eta_2+t\eta_2)}{\partial s\partial t}\Big|_{s=t=0}. \tag{61}$$

Accordingly, we integrate the previous equation twice from zero to η, $\|\eta\|_2 \le r/2$. We obtain for $\xi \in \mathfrak{B}_{r/2}$

$$\chi_1(\xi) = \exp[-\xi\cdot\left(\frac{\sigma\Gamma_1\sigma^T}{4}-\frac{V}{2}\right)\xi - F\left(\frac{\xi}{\cos\theta}\right)], \tag{62}$$

with

$$V := -\frac{Q_{12}(0,0)}{\tan\theta} = \frac{\sigma(\operatorname{Tr} gR_1R_2^T)\sigma^T}{\tan\theta},$$

$$F(\xi) := \cos^2\theta \int_{\mathcal{C}(\xi)} \left(\int_{\mathcal{C}(\eta)} \left(Q_{22}(\eta_1) + \frac{Q_{12}(\eta_1)}{\tan\theta} \right) \cdot d\eta_1 \right) \cdot d\eta,$$

where $\mathcal{C}(\xi), \mathcal{C}(\eta)$ are curves in phase space connecting the origin with the vectors ξ and η. Moreover, these last terms can be upper bounded by (see Appendix on section V)

$$\|V\|_2 \leq \left(\frac{\sqrt{24n^2\kappa}}{|\tan\theta|} \right) \sqrt{\varepsilon}, \tag{63}$$

$$\left| F\left(\frac{\xi}{\cos\theta} \right) \right|^2 \leq \left(\frac{n^2\kappa \, \|\xi\|_2^4}{2\tan^2\theta} \right) \varepsilon^{2/3}, \tag{64}$$

where the largest absolute fourth moment of ρ_{ab} is defined as

$$\kappa := \max \left\{ \left\| \rho_{ab} R_\xi^2 R_\eta^2 \right\|_1 \;\middle|\; \|\xi\|_2 = \|\eta\|_2 = 1 \right\}. \tag{65}$$

At this point we can show that the CM of ρ_1 and ρ_2 are $\varepsilon-$close. From differentiating Eq. (53) with respect to η_2 and evaluating at zero, we get the relation between the CMs of ρ_1 and ρ_2,

$$\varGamma_1 - \varGamma_2 = \frac{2}{\cos^2\theta} V.$$

Hence, from Eq. (63)

$$\|\varGamma_1 - \varGamma_2\|_2 \leq \left(\frac{\sqrt{384n^2\kappa}}{|\sin 2\theta|} \right) \sqrt{\varepsilon}. \tag{66}$$

Now we can proceed to show that for $\xi \in \mathfrak{B}_{r/2}$, the characteristic function of the state ρ_1 is $\varepsilon-$close to the Gaussian characteristic function $\Phi = \exp[-\xi \cdot \varGamma_1 \xi/4]$.

We use Eq. (62), $|e^x - 1| \leq |x| \max\left\{1, e^{\operatorname{Re}[x]}\right\}$ and $|\chi_1(\xi)|^2 = |\Phi(\xi)|^2 e^{\xi \cdot V\xi - 2\operatorname{Re}[F(\xi/\cos\theta)]}$ to write the bound as

$$\begin{aligned}
\int\limits_{\xi \in \mathfrak{B}_{r/2}} |\chi_1(\xi) - \Phi(\xi)|^2 \, d\xi &= \int\limits_{\xi \in \mathfrak{B}_{r/2}} \left| \Phi(\xi)|^2 | \exp[\xi \cdot V\xi/2 - F(\xi/\cos\theta)] - 1 \right|^2 d\xi \\
&\leq \int\limits_{\xi \in \mathfrak{B}_{r/2}} |\chi_1(\xi)|^2 \left| \frac{\xi \cdot V\xi}{2} - F(\xi/\cos\theta) \right|^2 d\xi, \\
&\leq \frac{1}{2} \int\limits_{\xi \in \mathfrak{B}_{r/2}} |\chi_1(\xi)|^2 \left(\frac{|\xi \cdot V\xi|^2}{4} + |F(\xi/\cos\theta)|^2 \right) d\xi \\
&\leq \left(\frac{4n^2\kappa\varepsilon^{2/3}}{\tan^2\theta} \right) \int\limits_{\xi \in \mathfrak{B}_{r/2}} |\chi_1(\xi)|^2 \, \|\xi\|_2^4 \, d\xi,
\end{aligned} \tag{67}$$

Here in the second inequality we have used again $|z_1 - z_2|^2 \leq (|z_1|^2 + |z_2|^2)/2$ for $z_1, z_2 \in \mathbb{C}$ and in the last inequality Eq. (63) and Eq. (64). We show now that

$$\frac{1}{(2\pi)^n} \int\limits_{\xi \in \mathfrak{B}_{r/2}} |\chi_1(\xi)|^2 \, \|\xi\|_2^4 \, d\xi \leq \frac{512n^2\kappa}{\pi^n} \left(\frac{1 + 3\sin 2\theta}{\cos^4\theta} \right). \tag{68}$$

We used again the Wigner representation, Eq.(43), of the characteristic function $|\chi_1(\xi)|^2$ in order to compute the bound of Eq. (68)

$$\begin{aligned}\frac{1}{(2\pi)^n}\int\limits_{\xi\in\mathfrak{B}_{r/2}}|\chi_1(\xi)|^2\,\|\xi\|_2^4\,d\xi &\leq \frac{1}{(2\pi)^n}\int\limits_{\xi\in\mathbb{R}^{2n}}|\chi_1(\xi)|^2\,\|\xi\|_2^4\,d\xi,\\ &= \frac{1}{(2\pi)^n}\sum_{k,l=1}^{2n}\int\limits_{\xi\in\mathbb{R}^{2n}}|\chi_1(\xi)|^2\xi_k^2\xi_l^2\,d\xi,\\ &= \frac{\partial^4\mathcal{W}(0)}{\partial\eta_k^2\partial\eta_l^2}.\end{aligned}$$

Using Lemma 14 with the help of the CCR gives

$$\sum_{k,l=1}^{2n}\frac{\partial^4\mathcal{W}(0)}{\partial\eta_k^2\partial\eta_l^2} = \frac{8}{\pi^n}\sum_{k,l=1}^{2n}\mathrm{Tr}[\rho\mathcal{P}\{R_k^2,R_l^2\}].$$

Here ρ is again the density operator corresponding to the characteristic function $|\chi(\xi)|^2$. We use again that $\mathcal{P}\rho=\rho\mathcal{P}$ (see subsection III C) and Hölder's inequality to bound

$$\frac{8}{\pi^n}\sum_{k,l=1}^{2n}\mathrm{Tr}[\rho\mathcal{P}\{R_k^2,R_l^2\}] \leq \frac{8}{\pi^n}\sum_{k,l=1}^{2n}\left(\left\|\sqrt{\rho}R_k^2R_l^2\sqrt{\rho}\right\|_1+\left\|\sqrt{\rho}R_l^2R_k^2\sqrt{\rho}\right\|_1\right).$$

Now using Cauchy-Schwarz inequality and the cyclicity properties for Schwartz operators we find

$$\begin{aligned}\frac{8}{\pi^n}\sum_{k,l=1}^{2n}\mathrm{Tr}[\rho\mathcal{P}\{R_k^2,R_l^2\}] &\leq \frac{16}{\pi^n}\left(\sum_{k=1}^{2n}(\mathrm{Tr}\,\rho R_k^4)^{1/2}\right)^2,\\ &\leq \frac{32n}{\pi^n}\sum_{k=1}^{2n}\mathrm{Tr}\,\rho R_k^4.\end{aligned}\tag{69}$$

Since we want to specify all the constants in terms of the moments of the output state ρ_{ab} we need to do the following computations. Using the explicit form of the BS transformation Eq. (18) and the derivatives of $\chi_1(\xi)$, we obtain after a tedious, but straightforward calculation

$$\cos^4\theta\sum_{k=1}^{2n}\mathrm{Tr}\,\rho_1R_{1k}^4+\sin^4\theta\sum_{k=1}^{2n}\mathrm{Tr}\,\rho_2R_{2k}^4-6\sin^2\theta\cos^2\theta\sum_{k=1}^{2n}\mathrm{Tr}\,\rho_1R_{1k}^2\,\mathrm{Tr}\,\rho_2R_{2k}^2=\sum_{k=1}^{2n}\mathrm{Tr}\,\rho_{ab}R_k^4,$$

which together with the positivity of the fourth moments implies

$$\sum_{k=1}^{2n}\mathrm{Tr}\,\rho_1R_k^4\leq 6\tan^2\theta\sum_{k=1}^{2n}\mathrm{Tr}\,\rho_1R_{1k}^2\,\mathrm{Tr}\,\rho_2R_{2k}^2+\frac{1}{\cos^4\theta}\sum_{k=1}^{2n}\mathrm{Tr}\,\rho_{ab}R_k^4.\tag{70}$$

Moreover, the relation between the fourth moments of the symmetrized state ρ and ρ_1 is given by

$$\mathrm{Tr}\,\rho R_k^4 = 2\,\mathrm{Tr}\,\rho_1R_k^4+6(\mathrm{Tr}\,\rho_1R_k^2)^2,\qquad k=1,\ldots,2n.\tag{71}$$

Combining Eqs. (70) and (71) and using $(\mathrm{Tr}\,\rho_1R_k^2)^2\leq\mathrm{Tr}\,\rho_1R_k^4\leq\kappa$ we obtain

$$\sum_{k=1}^{2n}\mathrm{Tr}\,\rho R_k^4\leq 16n\kappa\left(\frac{1+3\sin 2\theta}{\cos^4\theta}\right),$$

and Eq. (69) gives the claimed bound in Eq. (68). Hence, inserting Eq. (68) in Eq. (67) we obtain the bound for the non-vanishing region $\mathfrak{B}_{r/2}$

$$\frac{1}{(2\pi)^n}\int\limits_{\xi\in\mathfrak{B}_{r/2}} |\chi_1(\xi)-\Phi(\xi)|^2\, d\xi \le c_1^2\varepsilon^{2/3}, \tag{72}$$

where

$$c_1 := 32\sqrt{\frac{2}{\pi^n}\left(\frac{1+3\sin 2\theta}{\sin 2\theta}\right)}n^2\kappa. \tag{73}$$

The result of the theorem follows from Eq. (48), (72) and the elementary inequality for non-negative scalars x_1, x_2, $\sqrt{x_1+x_2}\le\sqrt{x_1}+\sqrt{x_2}$.

D. Auxiliary Lemmas

Lemma 16. *Let $S\in GL(4n,\mathbb{R})$ such that*

$$S\begin{pmatrix}\Gamma & 0\\ 0 & \Gamma\end{pmatrix}S^T = \begin{pmatrix}\cdot & 0\\ 0 & *\end{pmatrix} \quad \textit{is } 2n\times 2n \textit{ block diagonal for all symmetric } \Gamma\in\mathbb{R}^{2n\times 2n}. \tag{74}$$

Then it follows that:

(i) S is either of the form

(a) $S=\begin{pmatrix}A & 0\\ 0 & D\end{pmatrix}$ or $S=\begin{pmatrix}0 & B\\ C & 0\end{pmatrix}$,

(b) $S=\begin{pmatrix}A & B\\ C & D\end{pmatrix}$,

with A,B,C and $D\in\mathbb{R}^{2n\times 2n}$ invertible.

(ii)

$$S\begin{pmatrix}X & 0\\ 0 & X\end{pmatrix}S^T = \begin{pmatrix}\cdot & 0\\ 0 & *\end{pmatrix} \quad \textit{is } 2n\times 2n \textit{ block diagonal for all } X\in\mathbb{R}^{2n\times 2n}. \tag{75}$$

Proof of Lemma 16. We decompose $S=\begin{pmatrix}A & B\\ C & D\end{pmatrix}$ into blocks A,B,C and $D\in\mathbb{R}^{n\times n}$ and observe that equation (74) is equivalent to

$$A\Gamma C^T + B\Gamma D^T = 0 \quad \text{for all } \Gamma=\Gamma^T. \tag{76}$$

Using tensor notation this equation can be written[25] as

$$\begin{aligned} &(A\otimes C + B\otimes D)|\Gamma\rangle = 0 \qquad \text{for all } |\Gamma\rangle\in\mathbb{R}^{4n}\otimes\mathbb{R}^{4n} \text{ symmetric.}\\ \iff &(A\otimes C + B\otimes D)P_+|X\rangle = 0 \qquad \text{for all } |X\rangle\in\mathbb{R}^{4n}\otimes\mathbb{R}^{4n}.\\ \iff &(A\otimes C + B\otimes D)P_+ = 0. \end{aligned}$$

$$\iff P_+(A^T\otimes C^T + B^T\otimes D^T) = 0, \tag{77}$$

where P_+ denotes the projector onto the symmetric subspace of $\mathbb{R}^{4n}\otimes\mathbb{R}^{4n}$. We make the following remark that will be used frequently during this proof: the symmetrization or symmetric component of a non-zero product state does not vanish. Suppose it does, then $P_+(|v\rangle\otimes|w\rangle)=0$ and $|v\rangle\otimes|w\rangle+|w\rangle\otimes|v\rangle=0$. Performing the scalar product with $\langle v|\otimes\langle w|$ leads to $\|v\|^2\|w\|^2+|\langle v|w\rangle|^2=0$ which is only zero if and only if $|v\rangle=|w\rangle=0$.

On (i): We now prove the first part of the Lemma by considering the two cases:
(a) One of the submatrices A,B,C or D is zero: we only treat the case $A = 0$, the others are similar. Then $A^T \otimes C^T = 0$ and B^T is invertible (otherwise S would not have full rank). Since any non-zero product $B^T v \otimes D^T w$ (for some $v, w \in \mathbb{R}^{2n}$) would contain a non-vanishing symmetric component, equation (77) implies that $D^T = 0$.
(b) We prove by contradiction that in this case the four submatrices are invertible. For instance, assume that A is not invertible. Then there exists a vector $0 \neq |a\rangle \in \mathrm{Ker} A^T$. We choose $|e\rangle \notin \mathrm{Ker} D^T$ (recall that $D \neq 0$) and with (77) we then find

$$0 = P_+(A^T \otimes C^T + B^T \otimes D^T)|a\rangle \otimes |e\rangle = P_+(B^T|a\rangle \otimes D^T|e\rangle). \tag{78}$$

By the same argument as in (a) we now conclude $|a\rangle \in \mathrm{Ker} B^T$. Moreover,

$$S^T \begin{pmatrix} a \\ 0 \end{pmatrix} = \begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix} \begin{pmatrix} a \\ 0 \end{pmatrix} = 0.$$

The latter is in contradiction to the invertibility of S (if S is an invertible matrix then the kernel is trivial), hence A –and due to analogous reasoning– B, C and D are invertible.

On (ii): We now proof the second part of the Lemma, Eq. (75), by showing that the equivalent expression $AXC^T + BXD^T = 0$ for all $X \in \mathbb{R}^{2n\times 2n}$ is true.
This is trivially satisfied if S is given in form of case (i,a). Therefore we are left with the case (i,b) where in particular B and C are invertible. W.l.o.g. we choose $B = C = \mathbb{1}$ (this can be done by redefining $A \to A^{-1}B$ and $D \to D^{-1}C$ in (76)) and equation (77) reads

$$P_+(A^T \otimes \mathbb{1} + \mathbb{1} \otimes D^T) = 0. \tag{79}$$

We now show in three steps that $A^T \otimes \mathbb{1} + \mathbb{1} \otimes D^T = 0$, which then concludes the proof. First we show that there exists $\lambda \in \mathbb{C}$ such that $Spec(A^T) = \{\lambda\}$ and $Spec(D^T) = \{-\lambda\}$. To this purpose we choose eigenvectors $|e\rangle$ of A^T and $|f\rangle$ of and D^T with eigenvalues λ and ω respectively. Then

$$0 = P_+(A^T \otimes \mathbb{1} + \mathbb{1} \otimes D^T)|e\rangle \otimes |f\rangle = (\lambda + \omega)P_+(|e\rangle \otimes |f\rangle).$$

Again, since the symmetrization of a non-zero product state is different from zero, we find $\lambda = -\omega$. Note that this holds for arbitrary eigenvalues λ of A^T and ω of D^T.
Second, using the Jordan normal form decomposition, we decompose A^T (and D^T) into a diagonalizable $\lambda\,\mathbb{1}$ and nilpotent part $N_A(N_D)$ and observe that $(A^T \otimes \mathbb{1} + \mathbb{1} \otimes D^T) = (\lambda\,\mathbb{1} \otimes \mathbb{1} + N_A \otimes \mathbb{1}) + \mathbb{1} \otimes (-\lambda\,\mathbb{1} + N_D) = (N_A \otimes \mathbb{1} + \mathbb{1} \otimes N_D)$.
Finally, equation (79) reads

$$P_+(N_A \otimes \mathbb{1} + \mathbb{1} \otimes N_D) = 0, \tag{80}$$

and we can conclude the proof by deriving that this implies $(N_A \otimes \mathbb{1} + \mathbb{1} \otimes N_D) = 0$. This is the third step.

Assume $N_A \otimes \mathbb{1} + \mathbb{1} \otimes N_D \neq 0$. Using the symmetry argument about non-zero product states we find $N_A \neq 0$ and $N_D \neq 0$. Let s be such that $N_A^s = 0$ and $N_A^{s-1} \neq 0$. Then we multiply (80) from the right by $N_A^{s-1} \otimes \mathbb{1}$ to get $P_+(N_A^{s-1} \otimes N_D) = 0$. But this, in turn, implies $N_A^{s-1} \otimes N_D = 0$ and leads to a contradiction. Therefore

$$(N_A \otimes \mathbb{1} + \mathbb{1} \otimes N_D) = 0. \tag{81}$$

□

Lemma 17. *Let ρ_{12} be a density operator of a bipartite system with reduced states ρ_1 and ρ_2. If $\tilde{\rho}_1 \otimes \tilde{\rho}_2$ describe an arbitrary product state and $\|\rho_{12} - \tilde{\rho}_1 \otimes \tilde{\rho}_2\|_1 \leq \varepsilon$, then $\|\rho_{12} - \rho_1 \otimes \rho_2\|_1 \leq 3\varepsilon$.*

Proof. Using the triangle inequality twice, we find

$$\begin{aligned}
\|\rho_{12} - \rho_1 \otimes \rho_2\|_1 &\leq \|\rho_{12} - \tilde{\rho}_1 \otimes \tilde{\rho}_2\|_1 + \|\tilde{\rho}_1 \otimes \tilde{\rho}_2 - \rho_1 \otimes \rho_2\|_1 \,, \\
&\leq \varepsilon + \|\tilde{\rho}_1 \otimes \tilde{\rho}_2 - \tilde{\rho}_1 \otimes \rho_2\|_1 + \|\tilde{\rho}_1 \otimes \rho_2 - \rho_1 \otimes \rho_2\|_1 \,, \\
&= \varepsilon + \|\tilde{\rho}_2 - \rho_2\|_1 + \|\tilde{\rho}_1 - \rho_1\|_1 \,.
\end{aligned}$$

Exploiting that $\|X\|_1 = \sup_{Y:\|Y\|\leq 1} \operatorname{Tr}[YX]$ we can bound

$$\begin{aligned}\|\tilde{\rho_1} - \rho_1\|_1 &= \sup_{Y:\|Y\|\leq 1} \operatorname{Tr}[(Y\otimes \mathbb{1})(\tilde{\rho_1}\otimes\tilde{\rho_2} - \rho_{12})],\\ &\leq \sup_{\hat{Y}:\|\hat{Y}\|\leq 1} \operatorname{Tr}[\hat{Y}(\tilde{\rho_1}\otimes\tilde{\rho_2} - \rho_{12})],\\ &= \|\tilde{\rho_1}\otimes\tilde{\rho_2} - \rho_{12}\|_1 \leq \epsilon,\end{aligned}$$

and similar for the other term. □

Proof of Lemma (15). We follow the proof idea of Lemma 1 from Ref. 17. W.l.o.g assume $0 < \theta \leq \pi/4$ and set $\eta_2 = \tan\theta\eta_1, \eta_1 = \xi$ in Eq. (37). In case $\theta > \pi/4$, set $\eta_1 = \eta_2/\tan\theta$ in Eq. (37) and proceed likewise. Then Eq. (37) becomes

$$\chi_1\left(\left(1+\tan^2\theta\right)\cos\theta\xi\right) = \chi_1\left(\cos\theta\xi\right)\chi_1(\sin\theta\tan\theta\xi)\chi_2(\cos\theta\tan\theta\xi)\chi_2(-\sin\theta\xi) + G(\xi,\tan\theta\xi),$$

for all $\xi\in\mathbb{R}^{2n}$. Replace $\xi\mapsto(\xi/\cos\theta)$ in the previous equation to obtain

$$\chi_1\left(\left(1+\tan^2\theta\right)\xi\right) = \chi_1\left(\xi\right)\chi_1(\tan^2\theta\xi)|\chi_2(\tan\theta\xi)|^2 + G\left(\frac{\xi}{\cos\theta}, \frac{\tan\theta\xi}{\cos\theta}\right).$$

Since $\|G\|\leq 3\varepsilon$, we have for all $\xi\in\mathbb{R}^{2n}$:

$$\left|\chi_1\left(\left(1+\tan^2\theta\right)\xi\right)\right| \geq |\chi_1(\xi)||\chi_1(\tan^2\theta\xi)||\chi_2(\tan\theta\xi)|^2 - 3\varepsilon.$$

Similarly, for $0<\theta\leq\pi/4$ and $\eta_1 = -\tan\theta\eta_2, \eta_2=\xi$ we arrive to

$$\left|\chi_2\left(\left(1+\tan^2\theta\right)\xi\right)\right| \geq |\chi_2(\xi)||\chi_2(\tan^2\theta\xi)||\chi_1(\tan\theta\xi)|^2 - 3\varepsilon.$$

With $\gamma(\xi) := \min_j \min_{\|\eta\|<\|\xi\|} |\chi_j(\eta)|$ we obtain $\gamma\left(\left(1+\tan^2\theta\right)\xi\right) \geq \gamma^4\left(\xi\right) - 3\varepsilon$. Replacing ξ by $\left(1+\tan^2\theta\right)^k\xi$ with $k\in\mathbb{N}$ in the previous equation gives

$$\gamma\left(\left(1+\tan^2\theta\right)^{k+1}\xi\right) \geq \gamma^4\left(\left(1+\tan^2\theta\right)^k\xi\right) - 3\varepsilon \qquad \text{for all } \xi\in\mathbb{R}^{2n}, k\in\mathbb{N}. \tag{82}$$

It is a fact that for any classical characteristic function $\phi(t)$ with variance λ the following inequality holds (see for instance Ref. 21, p. 89)

$$|\phi(t)| \geq 1 - \frac{1}{2}\lambda t^2. \tag{83}$$

Let us fix ξ in the direction of phase space in which we obtain the largest variance λ of ρ_1 and ρ_2 and consider the region where $\|\xi\|_2 \leq \sqrt{\frac{1}{\lambda}}$. Then from Eq. (83) for $\|\xi\|_2\leq\sqrt{\frac{1}{\lambda}}$ we have $|\chi(\xi)|\geq 1/2$.
Using Eq. (82) and the inequality $(1-a)^n \geq 1-na, \forall n\in\mathbb{N}, \forall a\in[0,1]$, we can show by induction that

$$\gamma\left(\left(1+\tan^2\theta\right)^{k+1}\xi\right) \geq \left(\frac{1}{2}\right)^{4^k} - 4\varepsilon \qquad \text{for } k\in\mathbb{N}, \|\xi\|_2 \leq \sqrt{\frac{1}{\lambda}}.$$

Moreover, for $\varepsilon < 1$ we clearly have

$$\gamma\left(\left(1+\tan^2\theta\right)^{k+1}\xi\right) \geq \left(\frac{1}{2}\right)^{4^k} - 4\varepsilon^{1/12} \qquad \text{for } k\in\mathbb{N}, \|\xi\|_2 \leq \sqrt{\frac{1}{\lambda}}. \tag{84}$$

Finally, we take k_0 such that $2^{k_0} = \sqrt{\log_2\frac{1}{\varepsilon^{1/12}}}$, to obtain with the help of Eq. (84)

$$\begin{aligned}\gamma\left(\left(1+\tan^2\theta\right)^{k_0}\xi\right) &\geq \left(\frac{1}{2}\right)^{4^{k_0-1}} - 4\varepsilon^{1/12},\\ &= 12\varepsilon^{1/12}.\end{aligned}$$

We thus have $\gamma\left(\left(1+\tan^2\theta\right)^{k_0}\xi\right) > \varepsilon^{1/12}$ and $\left(1+\tan^2\theta\right)^{k_0}\xi \in \mathfrak{B}_r$ as claimed. □

IV. DISCUSSION

The DS theorem can be understood as the statement that Gaussian bosonic states with same covariance matrix are the only fixed point states of a non-trivial beam splitter transformation. In constrast with other characterizations of Gaussian states such as the one from Hudson[14], the DS theorem does not require any constraint on the purity of the state. The stability result of Theorem 9 provides an explicit estimate of the robustness of a characterization of Gaussian states through linear independence. In particular, we have obtained an estimate of the constants which reflects the fact that the quantum DS theorem is unstable when the beam-splitter is close to being transparent ($\theta = 0$) or a mirror ($\theta = \pi/2$). Throughout this work, we have made an effort to present explicit constants as well as to improve the order of the error parameter; however this does not mean that they are anywhere close to optimal. In fact, it is not known to us if the optimal constant must necessarily depend on the number of modes n or whether the $\log(1/\varepsilon)^{-1/2}$ dependence can be lifted to a polynomial dependence.

The Darmois-Skitovich theorem is not only interesting as a neat characterization problem, but also because of its practical applications: it is the main theoretical concept behind the signal reconstruction method known as *blind source separation* [23] and *independent component analysis* [22] which are actively studied in the field of communication and signal processing. These classical applications only work because the result is sufficiently robust. It is conceivable that there will be analogous quantum applications where robustness plays a similar role. We hope with this study of the stability of the quantum DS theorem to stimulate a further investigation of this theorem and its extensions in the quantum information community.

Acknowledgments: I would like to thank Michael M. Wolf for his continuous interest and many insightful discussions throughout this work. In particular, I thank him for suggesting the problem considered in this paper. In addition, I would like to thank A. Michelangeli for an insightful discussion of Lemma 1 from Ref. 17, and M. Keyl for clarifications concerning Schwartz operators. Furthermore, I wish to thank M. Christandl and J. P. Solovej for their hospitality during my visit at the QMATH Center as well as the financial support from the Danish Council for Independent Research (Sapere Aude). This work was partly written during this visit.

V. APPENDIX

A. Upper bound of Eq. (63)

Let us write $R_{1k}, k = 1, \dots 2n$, and $R_{2l}, l = 1, \dots 2n$, for the entries of the vector $R_1 = (Q_1, P_1, \dots, Q_n, P_n)$ and $R_2 = (Q_{n+1}, P_{n+1}, \dots, Q_{2n}, P_{2n})$ respectively. First, from the orthogonality of the symplectic matrix σ

$$\begin{aligned} \|V\|_2 &= \frac{\left\|\sigma(\operatorname{Tr} gR_1R_2^T)\sigma^T\right\|_2}{|\tan\theta|} \\ &= \frac{\left\|\operatorname{Tr} gR_1R_2^T\right\|_2}{|\tan\theta|} = \frac{\left(\sum_{kl} |\operatorname{Tr} gR_{1k}R_{2l}|^2\right)^{1/2}}{|\tan\theta|}. \end{aligned}$$

Our plan is to bound each entry $|\operatorname{Tr} gR_{1k}R_{2l}|^2$. Since the operator g is the difference of two Schwartz operators it is also Schwartz. Thus we have from Theorem 13 $(iii) - (iv)$ that $|g|^{1/2}$ is a Schwartz operator. If $g = \sum \lambda_i |i\rangle\langle i|$ is the spectral decomposition of g, we consider the following factorization

$$g = |g|Q, \qquad \text{with} \qquad Q := \sum_i \operatorname{sgn}(\lambda_i)|i\rangle\langle i|.$$

Clearly, $Q = Q^{-1}$ commutes with $|g|$ and $|g|^{1/2}$. Moreover, from Theorem 13 (i) we will be able to use the trace cyclicity for our following computation. Using twice the Cauchy-Schwarz inequality and $\|g\|_1 \leq 3\varepsilon$, we find

$$\begin{aligned} |\operatorname{Tr} gR_{1k}R_{2l}|^2 = |\operatorname{Tr} |g|^{1/2}|g|^{1/2}QR_{1k}R_{2l}|^2 &\leq (\operatorname{Tr}|g|)\left(\operatorname{Tr}|g|R_{1k}^2R_{2l}^2\right), \\ &\leq 3\varepsilon\sqrt{\operatorname{Tr}|g|R_{1k}^4 \operatorname{Tr}|g|R_{2l}^4}, \\ &\leq (3\varepsilon)\max\{\operatorname{Tr}|g|R_{1k}^4, \operatorname{Tr}|g|R_{2l}^4\}. \end{aligned}$$

Using again the decomposition of g we obtain for $j = 1, 2$ that

$$|\operatorname{Tr}|g|R_{jk}^4| = |\operatorname{Tr} QgR^4| \leq \|Q\| \left\|gR_{jk}^4\right\|_1 = \left\|(\rho_{ab} - \rho_a \otimes \rho_b)R_{jk}^4\right\|_1 \\ \leq 2\left\|\rho_{ab}R_{jk}^4\right\|_1,$$

since R_{jk}^4 is a local operator on one part of the output. Consequently, $|\operatorname{Tr} gR_{1k}R_{2l}|^2 \leq 6\varepsilon \max\{\operatorname{Tr}\rho_{ab}R_{1k}^4, \operatorname{Tr}\rho_{ab}R_{2l}^4\} \leq 6\varepsilon\kappa$ where $\kappa := \max\left\{\left\|\rho_{ab}R_\xi^2R_\eta^2\right\|_1 \mid \|\xi\|_2 = \|\eta\|_2 = 1\right\}$ is the largest generalized fourth moment of ρ_{ab}. Note that since ρ_{ab} is a Schwartz operator $\kappa < \infty$. Thus

$$\|V\|_2 \leq \frac{\sqrt{24n^2\kappa\varepsilon}}{|\tan\theta|}.$$

B. Upper bound of Eq. (64)

The line integral of a matrix $A \in \mathbb{C}^{2n\times 2n}$ is defined in terms of the line integrals of the rows of A. Namely, if $A_k, k = 1, \ldots, 2n$ are the rows of A, then

$$\int A \cdot d\eta := \begin{pmatrix} \int A_1 \cdot d\eta \\ \vdots \\ \int A_{2n} \cdot d\eta \end{pmatrix}.$$

Thus the line integral of a matrix-valued function is a vector, and the line integral of a vector field is a scalar. Let us denote by $M : \mathbb{R}^{2n} \to \mathbb{C}^{2n\times 2n}$ a matrix-valued function. The upper bound for Eq. (64) is equivalent to bound

$$F(\xi) := \int_{\mathcal{C}(\xi)} \left(\int_{\mathcal{C}(\eta)} M(z)\cdot dz\right)\cdot d\eta,$$

with $\xi, \eta \in \mathfrak{B}_{r/2}$ and

$$M(z) = \cos^2\theta\left(Q_{22}(z) + \frac{Q_{12}(z)}{\tan\theta}\right).$$

Let us parametrize the curve $\mathcal{C}(\xi)$ via $\vartheta : [0,1] \to \mathcal{C}$, $t \mapsto t\xi$ and write $M_{kl}(z)$ for the entries of the matrix $M(z) \in \mathbb{C}^{2n\times 2n}$. Define the matrix-valued function $\mathbb{R}^{2n} \ni z \mapsto Y(z) \in \mathbb{C}^{2n\times 2n}$ to have the entries $Y_{kl}(z) := \int_0^1 M_{kl}(sz)ds$. From the explicit parametrization of the line integrals and the Cauchy-Schwarz inequality we obtain that

$$\begin{aligned} |F(\xi)| &\leq \max_{t\in[0,1]} |t\xi\cdot Y(t\xi)\xi|, \\ &\leq \|\xi\|_2^2 \max_{t\in[0,1]}\left(\sum_{k,l=1}^{2n} |Y(t\xi)_{kl}|^2\right)^{1/2}. \end{aligned} \tag{85}$$

In order to bound $|Y(t\xi)_{kl}|$ we differentiate Eq. (51) with the help of Lemma 14 to find

$$Q_{22}(z) = \frac{G_2(z)G_2^T(z)}{\chi_1^2(\cos\theta z)\chi_2^2(-\sin\theta z)} - \frac{G_{22}(z)}{\chi_1(\cos\theta z)\chi_2(-\sin\theta z)} \in \mathbb{C}^{2n\times 2n},$$

$$Q_{12}(z) = -\frac{G_{12}(z)}{\chi_1(\cos\theta z)\chi_2(-\sin\theta z)} - \frac{\nabla\chi_{\rho_a}(z)G_2^T(z)}{\chi_1^2(\cos\theta z)\chi_2^2(-\sin\theta z)} \in \mathbb{C}^{2n\times 2n},$$

for $z \in \mathbb{R}^{2n}$, $\|z\|_2 \leq r/2$ where

$$\begin{aligned} G_2(z) &= -\frac{i}{2}\sigma\operatorname{Tr}[e^{iz\cdot\sigma R_1}\{R_2, g\}] \in \mathbb{C}^{2n}, \\ G_{22}(z) &= -\frac{1}{4}\sigma\operatorname{Tr}[e^{iz\cdot\sigma R_1}\{\{R_2, g\}, R_2^T\}]\sigma^T \in \mathbb{C}^{2n\times 2n}, \\ G_{12}(z) &= -\frac{1}{4}\sigma\operatorname{Tr}[e^{iz\cdot\sigma R_1}\{\{R_1, g\}, R_2^T\}] \in \mathbb{C}^{2n\times 2n}, \\ \nabla\chi_{\rho_a}(z) &= -\frac{i}{2}\operatorname{Tr}[e^{iz\cdot\sigma R_1}\{R_1, \rho_a\}] \in \mathbb{C}^{2n}. \end{aligned}$$

Let us write $R_{1k}, k = 1, \ldots 2n$, and $R_{2l}, l = 1, \ldots 2n$, for the entries of the vector $R_1 = (Q_1, P_1, \ldots, Q_n, P_n)$ and $R_2 = (Q_{n+1}, P_{n+1}, \ldots, Q_{2n}, P_{2n})$ respectively. From Lemma 15 we know that for $z \in \mathfrak{B}_r$, $\chi(z) > 12\varepsilon^{1/12}$ and therefore we can upper bound each entry of $Y(t\xi)_{kl}$ by

$$|Y(t\xi)_{kl}| \leq \cos^2\theta \left(\frac{\|\{\{R_{2k}, g\}, R_{2l}\}\|_1}{4(12\varepsilon^{1/12})^2} + \frac{\|\{R_{2k}, g\}\|_1 \, \|\{R_{2l}, g\}\|_1}{4(12\varepsilon^{1/12})^4} + \frac{\|\{\{R_{1k}, g\}, R_{2l}\}\|_1}{4(12\varepsilon^{1/12})^2 \tan\theta} + \frac{\|\{R_{1k}, \rho_a\}\|_1 \, \|\{R_{2l}, g\}\|_1}{4(12\varepsilon^{1/12})^4 \tan\theta} \right).$$

Following a similar procedure as for the bound of $\|V\|_2$ (see Apendix V A) we obtain

$$|Y(t\xi)_{kl}|^2 \leq \left(\frac{\kappa \cos^2\theta}{8\tan^2\theta} \right) \varepsilon^{2/3},$$

so from Eq. (85)

$$\left| F\left(\frac{\xi}{\cos\theta} \right) \right|^2 \leq \left(\frac{n^2 \kappa \, \|\xi\|_2^4}{2\tan^2\theta} \right) \varepsilon^{2/3}.$$

VI. BIBLIOGRAPHY

[1] G. Darmois, "Analyse generale des liaisons stochastiques. Etude partieuiere de l'analyse factorielle lineaire," Rev. Inst. Int. Statist. **21**, (1953).

[2] V. P. Skitovich, "Linear forms of independent random variables and the normal distribution law," Izv. Akad. Nauk SSSR, Ser. Mat. **18**, (1954) (Russian). Engl. Translation in: Selected Translations in Math.Stat. and Prob. **2** (1962) 211-218.

[3] A. Mann, M Revzen, F. C. Khanna and Y. Takahashi, "Bifactorizable wavefunctions," J. Phys. A: Math. Gen. **24** 425 (1991); S. C. Springer, J. Lee, M. Bellini, and M. S. Kim, "Conditions for factorizable output from a beam splitter," Phys. Rev. A **79**, 062303, (2009).

[4] M. Keyl, J. Kiukas and R. Werner, "Schwartz Operators," Rev. Math. Phys. 28, 1630001 (2016)

[5] A. Holevo, *Probabilistic and statistical aspects of quantum theory* (Publications of the Scuola Normale Superiore, 2011).

[6] R. Simon, N. Mukunda and B. Dutta, "Quantum noise matrix for multimode systems: $U(n)$ invariance, squeezing, and normal forms," Phys. Rev. A **49**, 1567 (1994).

[7] R. Simon, G. Sudarshan and N. Mukunda, "Gaussian Wigner distributions in quantum mechanics and optics," Phys. Rev. A **36** (1987).

[8] J. Arvind, B. Dutta, N. Mukunda and R. Simon, "The real symplectic groups in quantum mechanics and optics," Pramana, 471 (1995).

[9] A. W. Harrow, A. Montanaro, A. J. Short, "*Limitations on Quantum Dimensionality Reduction*," Automata, Languages and Programming ICALP 2011, 86-97, Springer.

[10] W. Feller, *An introduction to probability theory and its applications: Vol II,* 2nd. ed. (Wiley, New York 1971).

[11] A. Kagan, J. Linnik, and C. Rao, "*Characterization problems in mathematical statistics,*" Wiley, (1973); M. P. Quine and E. Seneta, "The generalization of Kac-Bernstein theorem," Probability and Mathematical Statistics **19**, 441, (1999).

[12] S. G. Bobkov, G. P. Chistyakov and F. Götze, "Entropic instability of Cramers characterization of the normal law". In: van de Geer S., Wegkamp M. (eds) Selected Works of Willem van Zwet, Chapter 15. Springer (2012)

[13] E. Lukacs, "Stability Theorems," Adv. Appl. Prob. 9, 336 (1977).

[14] R. Hudson, "When is the wigner quasi-probability density non-negative," Rep. Math. Phys. **61**, 249, (1974)

[15] P. R. Fisk, "A Note on a Characterization of the Multivariate Normal Distribution," The Annals of Mathematical Statistics **41**, 486 (1970).

[16] L. Klebanov and R. Yanushkevichius, "ϵ-dependence of $X_1 + X_2$ and $X_1 - X_2$," Lietuvos Matematikos Rinkinys **25** (1985).

[17] Y. Gabovich, "On the stability of certain characteristic properties of the normal distribution," Theor. Prob. Appl. **19**, 365 (1974).

[18] G. F. Dell'Antonio. "On the limits of sequences of normal states," Commun. Pure Appl. Math. **20**, 413 (1967); E. B. Davies, "Quantum stochastic processes," Commun. Math. Phys. **15** 277 (1969).

[19] R. Honegger, A. Rieckers, "*Photons in Fock Space and Beyond*", Volume 1, World Scientific (2015).

[20] A.J.E.M. Janssen, "Proof of a Conjecture on the Supports of Wigner Distributions," Journal of Fourier Analysis and Applications, 4, 723-726 (1998); P. Jaming, "Principe d'incertitude qualitatif et reconstruction de phase pour la transformée de Wigner," C. R. Acad. Sci. Paris, 327, 249-254 (1998).

[21] N. G. Ushakov, *Selected Topics in Characteristic Functions* (De Gruyter, 1999).

[22] P. Comon, "Independent component analysis, A new concept?" Signal Processing **36**, 287 (1994)
[23] X. Cao and R. Liu, "General Approach to Blind Source Separation," IEEE Transactions on signal processing **44**, 562 (1996); J. Cardoso, "Blind signal separation: statistical principles," Proceedings of the IEEE **86**, 2009 (1998).
[24] A. Grossmann, "Parity Operator and Quantization of $\delta-$functions," Commun. math. Phys. **48**, 191 (1976).
[25] R. Horn and C. Johnson, *Topics in Matrix Analysis* (Cambridge University Press, 2007)