A $W^\star$-algebraic formalism for parametric models in Classical and Quantum Information Geometry

Von der Fakultät für Mathematik und Informatik
der Universität Leipzig
eingenommene

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM
(Dr.rer.nat.)

im Fachgebiet

Mathematik

Vorgelegt

von M.Sc. Fabio Di Nocera

geboren am 19.09.1992 in Gragnano (Italien)

Die Annahme der Dissertation wurde empfohlen von:

1. Prof. Dr. Paolo Gibilisco Università di Roma "Tor Vergata"
2. Prof. Dr. Jürgen Jost MPI für Mathematik in den Naturwissenschaften

Die Verleihung des akademischen Grades erfolgt mit Bestehen

der Verteidigung am 17.04.2024 mit dem Gesamtprädikat *magna cum laude*

Abstract

Title of the thesis: A $W^{\star}$-algebraic formalism for parametric models in Classical and Quantum Information Geometry
Author: Fabio Di Nocera

The aim of this work is to lay down a formalism for parametric models that encapsulates both Classical and Quantum Information Geometry.

This will be done introducing parametric models on spaces of normal positive linear functionals on $W^{\star}$-algebras and providing a way of defining a Riemannian structure on this models that comes from the Jordan product of the $W^{\star}$-algebra. This Riemannian structure will have some features that are appealing from the viewpoint of Information Geometry.

After introducing this $W^{\star}$-algebraic framework, we will move to Estimation Theory. We will see how and to what extent it is possible to formulate in this framework two well-known statistical bounds: the Cramér-Rao bound and the Helstrom bound.

Finally, we will explicitly construct some examples that show how it is possible to reduce this general framework to obtain well-known structures in Classical and Quantum Information Geometry.

Authorship

Most of this thesis is based on the work [49] with Florio Maria Ciaglia, my supervisor Jürgen Jost and Lorenz Schwachhöfer. This work contains the general approach of the thesis as well as most of the results. Some collateral aspects of this work are already present in [42], which is a work with Florio Maria Ciaglia, Fabio Di Cosmo and Patrizia Vitale, and in the single-author conference paper [71]. Moreover, some of the results presented here have not been published yet.

In particular:

Chapter 2: the chapter contains no original results. The formulation presented here is originally written by me.

Chapter 3: all the results contained in this chapter are present in [49]. Me and Florio Maria Ciaglia contributed equally to the proof of Propositions from 23 to 31, while the other two authors, Jürgen Jost and Lorenz Schwachhöfer, had the role of giving the general direction of the work and of revision of the content of the work.

Chapter 4: this chapter contains material that has not been published yet. Propositions 33, 34 and 35 have been proven by me, stimulated by discussions with Florio Maria Ciaglia and my supervisor Jürgen Jost.

Chapter 5: Sections 5.1 and Subsection 5.2.2 are examples already present in [49] and that were developed by me and Florio Maria Ciaglia. Subsection 5.2.1 uses an approach presented in [42] and in [71], but the example itself has not been published yet and has been developed by me. Finally, Section 5.3 contains material that has not been published yet and it has been developed by me, stimulated by discussions with Florio Maria Ciaglia and Jürgen Jost.

A mio padre
Per la spinta alla bici
A mia madre
Per le cure dopo la caduta

Contents

Chapter 1

Introduction

In this work, we will introduce a formalism for parametric models on $W^\star$-algebras with the aim to give a unified framework for Classical and Quantum Information Geometry.

Roughly speaking, Information Geometry may be understood as the application of (differential) geometric methods to the investigation of Information Theory, Statistics, Machine Learning, and related branches. As such, it is an interdisciplinary field of research in which different mathematical tools may find new life both in terms of applications and of theoretical development.

A crucial cornerstone of Classical Information Theory is the work of Claude E. Shannon [163], where he developed, based on the previous work of Harry Nyquist [142] and Ralph Hartley [101], a mathematical theory to describe transmission of information over a possibly noisy channel.

Quantum Information Theory, in a nutshell, is the branch of Information Theory that investigates what happens if one chooses not to encode information in a classical physical system, but rather in a quantum system. This gives rise to the opportunity of exploiting all the typical phenomena of Quantum Mechanics, such as *quantum interference*, *entanglement* and *quantum teleportation* for typical tasks of Information Theory, like *signal processing*, *information encoding* and *error correction* [139]. The field was initiated, among others, by Gordon in [93] and Helstrom in [105]. It was then Holevo who systematically tackled the problem of transmission, already tackled by Shannon in the classical case, for a quantum channel [107, 108].

The foundation of Classical Information Geometry can be traced back to the work of Rao [156], who first recognized, building on previous works by Fisher [78] and Mahalanobis [130], the possibility of endowing spaces of probability vectors with a Riemannian structure. In [34], Čencov proved that this Riemannian structure is actually the only one, up to rescaling of the metric tensor, that is invariant under sufficient statistics. It was then with the work of Amari [5, 6, 7, 8] that it became common to refer to this field as Information Geometry. Among the contributions of Amari there is the study of a dually flat structure arising in Classical Information Geometry and that is given in terms of a structure now known as ***Amari-Čencov structure***.

Regarding Quantum Information Geometry, the first attempts to categorize the metric structures that can be used is made by Čencov and Morozova in [133]. This problem was then fully solved, in the finite-dimensional case, by Petz in [149] and Petz and Sudar in [151].

The idea of building a framework that encapsulates Classical and Quantum Information Geometry comes mainly from one reason: if both theories are written in the same language it becomes easier to compare classical and quantum structures. In particular, we will see that choosing to work with the classical case amounts to choose to work with an Abelian $W^{\star}$-algebra, thus giving strength to the idea that the typical phenomena of the quantum case arise somehow from the non-commutativity of the underlying $W^{\star}$-algebra structure. Then one can try to investigate what is the role of the non-commutativity of the product in the structures that one introduces and this would hopefully give an insight on the differences and analogies between the classical and quantum case.

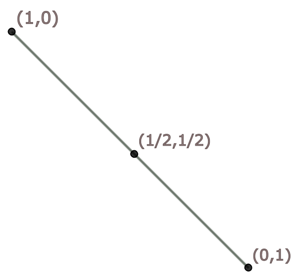
Now we will describe some aspects of both Classical and Quantum Information Geometry and formulate some questions, with the idea that an approach that claims to encapsulate both theories should be able to answer these questions. Of course, this work has to be intended as the first steps in the construction of such a framework.

In order to make any kind of geometrical analysis, we need in the first place a topological space that serves as background space for any additional structure we are going to introduce on it.

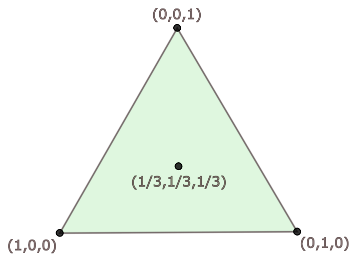
Let us first introduce such a background space for what we will refer in the following as the *classical, finite-dimensional case*. This means that we are considering probability distributions over a space of events X of finite cardinality n, i.e. probability vectors with n components. These objects can be put in a one-to-one correspondence with the point of the $(n-1)$-dimensional simplex, as it can be easily seen by choosing the so-called *barycentric coordinates* on such space [61]. Hence we can conclude that the appropriate background space for this case is the $(n-1)$-simplex, see Figure 1.1.

If the space of events X is not a finite set, we are in what we refer to the *classical, infinite-dimensional case*. Here, we have that the object of our analysis are sets of probability distributions over X. This case has been formulated in a general and elegant way in [12, 13], where the analysis of parametric models is carried out without referring to any reference measure. In this same work, an alternative approach due to Giovanni Pistone and Carlo Sempi [152, 153], is also discussed. We will say more about this last approach in the following.

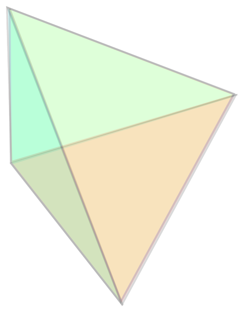
As mentioned before, in Quantum Information Theory, one wants to encode information in a quantum system $\mathcal{S}$, the description of a quantum system is usually given in terms of a Hilbert space $\mathcal{H}$, and observables are described as elements of the space $\mathcal{B}(\mathcal{H})_{sa}$ of self-adjoint operators acting on $\mathcal{H}$. Then the idea behind the description (or, better said, one of the possible descriptions) of quantum states is to see them as some map that associate to the observable A the ***expected value*** of A with respect to the state of the system. One then usually assumes that these maps are linear, i.e. they have to be elements of the dual of $\mathcal{B}(\mathcal{H})$. Then, for reasons



(a) 1-simplex.



(b) 2-simplex.



(c) 3-simplex.

Figure 1.1: Pictorial representations of the 1,2 and 3-simplex.

that are related to the so-called Copenhagen interpretation of Quantum Mechanics and in particular to Born's rule for obtaining the probabilities of the outcomes of measurements [25], these maps have to satisfy two requirements that are analogous to the fact that probability distributions are non-negative and sum to 1.

Let us start from considering a finite-dimensional Hilbert space $\mathcal{H}$ and let us denote the complex dimension of $\mathcal{H}$ with n, thus introducing the *quantum, finite-dimensional case*. Because of the finite-dimensional setting, operators and functionals on the space of operators can be identified and quantum states are described as semi-positive definite operators of trace 1, so the appropriate background space for this case is the intersection of the cone $\mathcal{P}(\mathcal{H})$ of semi-positive definite operators with the hyperplane Π_1 given by the constraint on the trace. The space $\mathscr{S}(\mathcal{H}) = \mathcal{P}(\mathcal{H}) \cap \Pi_1$ is called the ***space of quantum states*** of the quantum system $\mathcal{S}$ [19].

The case $n = 2$ is usually referred to as the ***qubit*** and has been extensively studied for its applications to Quantum Information Theory and Quantum Computing [139, 150, 172]. The reason for this is that, in analogy with classical bits, qubits are two level quantum systems and can be efficiently used to encode information. In this case the space of quantum states can be represented as a sphere, usually referred to as the ***Bloch sphere***.

Going from the *quantum, finite-dimensional case* to the *quantum, infinite-dimensional case* means allowing the Hilbert space to be infinite-dimensional, or more generally considering any infinite-dimensional non-commutative $C^\star$-algebra or $W^\star$-algebra.

We are now ready to formulate our first question.

Question 1. *Does the proposed approach allow to reconstruct the appropriate background spaces of Classical and Quantum Information Geometry in both the finite and infinite-dimensional case?*

No theory of information can disregard the idea of mapping information from one system to another, and for this reason, some particular kinds of maps are considered in each of the cases. The idea behind all of this maps is that they have to preserve the nature of the object they act on, i.e. sending probability distributions in probability distributions and sending quantum states in quantum states.

In the classical, finite-dimensional case, one can easily see that a matrix whose entries are non-negative and whose rows and columns sum to 1 sends probability vectors in probability vectors, a matrix with these properties is called a ***Markov matrix*** (or ***stochastic matrix***). These are used extensively in the study of Markov processes, i.e. processes that don't exhibit memory, meaning that the way the system changes from one step to another only depends on the state of the system at that given step [86].

The extension of this concept in the case that the space of events is continuous is called a ***Markov kernel*** (or ***stochastic kernel***) [121].

Definition 1 (Markov kernel)**.** *Let X and Y be sets, Σ_X and Σ_Y be σ-algebras respectively on X and Y, so that (X, Σ_X) and (Y, Σ_Y) are measurable spaces. Then a map k from $\Sigma_Y \times X$ to $[0, 1]$ such that*

- *For any fixed $B \in \Sigma_Y$ and for all $x \in X$ the map $x \mapsto k(B, x)$ is measurable on (X, Σ_X),*
- *For any fixed $x \in X$ and for all $B \in \Sigma_Y$ the map $B \mapsto k(B, x)$ defines a probability measure on (Y, Σ_Y),*

is called a Markov kernel with source (X, Σ_X) and target (Y, Σ_Y).

This is the relevant class of maps in the classical, infinite-dimensional case, and it can be easily seen that if X and Y are space of events of finite cardinality and the measures used are the counting measures on X and Y this definition reduces to the one of stochastic matrix.

When we want to refer to classical maps without specifying whether the underlying setting is that of a finite-dimensional sample space, an infinite-dimensional discrete sample space, or a continuum sample space, we will just use the name ***Markov maps***. An elegant wat of introducing these maps, making use of a category-theoretical framework, can be found in [81, 82, 83].

In the context of Quantum Information Theory, the appropriate class of maps are ***completely positive, trace-preserving*** (***CPTP***) ***maps***, sometimes also called ***quantum channels***. The trace-preserving requirement is an analogue of the preservation of normalization of probabilities in the classical case. A positive map in this context is a map that sends positive operators in positive operators, complete positivity is a stronger requirement. It is used for ensuring the positivity of the map even when the quantum system is coupled with an *ancilla*. Let us denote this map with Φ, not only the map Φ has to be positive, but also every map of the kind $\mathbb{I}_A \otimes \Phi$, where $\mathbb{I}_A$ is the identity of the ancillary system has to be positive, whatever the dimension of the ancillary system [139, 183]. Also these class of maps can be introduced with the aid of category theory, this has been done in [144].

Question 2. *In the proposed approach, can a class of maps that encapsulates both Markov maps and quantum channels be defined?*

Central objects in both Classical and Quantum Information Geometry are the metric tensors that can be defined on the background spaces referred to in Question 1.

Let us start as usual from probability vectors. In his seminal work [78], Ronald Fisher introduced the ***Fisher information matrix*** as the variance of the score of a random variable. Then Rao, in [156], came to realize that this matrix actually defines a Riemannian structure on spaces of probability vectors. This goes under the name of ***Fisher-Rao metric*** and its expression on the interior of the $(n-1)$-simplex is given by

$$G_p^{FR} = \sum_{j=1}^{n} p^j d\log p^j \otimes d\log p^j = \sum_{j=1}^{n} \frac{1}{p^j} dp^j \otimes dp^j, \tag{1.1}$$

so that its action on two vectors A and B tangent to the open interior of the $(n-1)$-simplex is given by

$$G_p^{FR}(A, B) = \sum_{j=1}^{n} \frac{1}{p^j} A^j B^j. \tag{1.2}$$

This same metric can be extended to the classical, infinite-dimensional case, let (X, Σ, μ) be a measure space, $S(X)$ be the space of signed measures on (X, Σ), and $\nu \in S(X)$. Let also σ_1 and σ_2 be two measures that are absolutely continuous with respect to ν, one can then define

$$G_\nu^{FR}(\sigma_1, \sigma_2) = \int_X \frac{d\sigma_1}{d\nu} \frac{d\sigma_2}{d\nu} \, d\nu, \tag{1.3}$$

where we denoted with $d\sigma_1/d\nu$ the Radon-Nikodym derivative of σ_1 with respect to ν. Of course then we have to be able to perform the integral and this puts more constraints on the measures one can use, see [13, Chapter 3] for a detailed discussion on how can this metric be defined in a more precise and organic way with the use of *parametric measure models*.

As already mentioned, in the classical case we have a crucial theorem, proved by Čencov in [34], that singles out Fisher-Rao metric as the only Riemannian metric on a statistical manifold that is invariant under sufficient statistics, which are a particular class of stochastic maps in the classical, finite-dimensional case. This is the reason why the Fisher-Rao metric is the only one that is used in Classical Information Geometry. Moreover, classical information satisfies a property called ***Data processing inequality***, whose meaning, in a nutshell, is that one can not gain information via stochastic maps [16, 29, 59].

In the quantum case, there is no analogous of Čencov's theorem. A *plethora* of metric tensors can be defined on the space of quantum states, and many of them are used by the community of Quantum Information Theory. Without claiming to be exhaustive, let us just mention the ***Wigner-Yanase metric tensor*** [91, 92, 102], the ***Bogolubov-Kubo-Mori metric tensor*** [138, 148] and the ***quantum Tsallis metric tensors*** [132]. Nonetheless, a principle analogous to the data processing inequality can be stated [128].

In fact Dénes Petz, following what done by Morozowa and Čencov in [133], characterized all metric tensors on the space of quantum states obeying monotonicity under completely positive, trace-preserving maps [149, 151]. For this reason, we will sometimes refer to the metrics used in Quantum Information Geometry as ***monotone metrics***. In particular, he showed that for every ***operator monotone function*** [20] f, one can define a superoperator depending on the state ρ as

$$\mathbf{K}_\rho^f = f(\mathbf{L}_\rho \mathbf{R}_{\rho^{-1}})\mathbf{R}_\rho, \tag{1.4}$$

where $\mathbf{L}_\rho$ and $\mathbf{R}_\rho$ are respectively the left and right multiplications by means of ρ, namely

$$\mathbf{L}_\rho(B) = \rho B, \quad \mathbf{R}_\rho(B) = B\rho. \tag{1.5}$$

Then the action of any monotone metric on two vectors $A, B \in T_\rho \mathscr{S}(\mathcal{H})$ can be written as

$$G^f|_\rho(A, B) = Tr\left(A\left(\mathbf{K}_\rho^f\right)^{-1}(B)\right) \tag{1.6}$$

for some operator monotone function f. Moreover, monotone metrics are in a one-to-one correspondence with operator monotone functions such that

$$\begin{aligned} f(x) &= xf(x^{-1}), \\ f(1) &= 1. \end{aligned} \tag{1.7}$$

Question 3. *Does the proposed approach allow to reconstruct the metric tensors that are used in Classical and Quantum Information Geometry in both the finite and infinite-dimensional case?*

We introduced a relevant class of maps and the relevant metrics for both Classical and Quantum Information Geometry. Also, we saw that they are related, meaning that the relevant metrics are selected in both theories precisely because of some property that they satisfy with respect to the relevant class of maps in the respective theory. Meaning that the data processing inequality transfers into a monotonicity property of the Riemannian structures of Information Geometry.

Question 4. *Is there an equivalent of the data processing inequality in this framework? More specifically, do the metric tensors obtained in this framework (see Question 3) satisfy some monotonicity property with respect to the maps considered in Questions 2?*

The study of the geometric structures underlying space of probabilities has brought to formulate a result that is crucial in Estimation Theory and goes under the name of ***Cramér-Rao bound***.

The Cramér-Rao bound is a lower bound in the attainable precision of an unbiased estimator in terms of the Fisher-Rao metric. Even if it is named after Harald Cramér and Calyampudi Radhakrishna Rao [62, 156], it has been noted in [164] that the idea was already present in the work [3] by Alexander Aitken and Harold Silverstone. Other authors that have stated the inequality independently are Maurice Fréchet [79] and Georges Darmois [64].

It is possible to formulate a bound on accuracy of estimators that can be seen as the quantum counterpart of the Cramér-Rao bound and that goes under the name of ***Helstrom bound*** [15, 103, 104, 105, 109, 110, 143]. On a quantum system one can perform different measurements to obtain different *classical shadows* of the considered quantum system, and while the Cramér-Rao bound is a lower bound for the accuracy of any unbiased estimator of these single measurements, the Helstrom bound is universal in the sense that it is related to the quantum system and does not depend on the particular measurement one chooses to perform on such system.

Question 5. *Can some result related to Estimation Theory, such as the Cramér-Rao bound and the Helstrom bound, be obtained in this framework?*

Another essential element of our framework, apart from the use of $W^\star$-algebras, is the use of parametric models.

First, let us stress that what are usually referred to as *non-parametric* approaches still put some constraints on the probability distributions or quantum states considered. In the finite-dimensional classical case it is typical to restrict to the interior of the simplex, so that the Fisher-Rao metric, see Equation (1.1), is always well defined, something analogous happens in the quantum finite-dimensional case, where one restricts to faithful states. Another example of this instance is the approach for infinite-dimensional Classical Information Geometry introduced in [153], where they use measures that are *compatible with a fixed measure* μ_0, which means that they have the same null sets. In Section 5.3 a more detailed discussion of this aspect will be given, and this will bring us to define as ***singular models*** those models that do not satisfy this kind of regularity.

Our notion of parametric model is based on the possibility of describing the class of probability distributions or quantum states as parametrized by a possibly infinite dimensional Banach manifold. The idea is then that some of the non-parametric approaches could still be recovered, if the Banach manifold considered encodes the constraints we mentioned before.

Question 6. *Can we produce meaningful examples of parametric models? Does this approach allow to describe models that can not be described in non-parametric approaches? Are we able to reconstruct non-parametric approaches as some particular models?*

Now we give the structure of this work, indicating how the various parts of the work give an answer to the questions we formulated.

- In Chapter 2, the mathematical background for our discussion is settled. In particular, we will see that simplices, space of probability distributions, and spaces of positive and normalized functionals on the space of bounded operators can all be seen as space of states of a $W^\star$-algebra. This represents a positive answer to Question 1. Also, we will introduce maps that have the property of being the dual of a completely positive, unital map between $W^\star$-algebras, and we will see that this class of maps encapsulates both the Markov maps of Classical Information Theory and the CPTP maps of Quantum Information Theory, answering in a positive way also to Question 2. In Section 2.3, we will tackle, in the finite-dimensional case, Question 3. We will see how it is possible, following what done in [53], to follow some sort of symmetric analogous of the Kirillov-Kostant-Souriau construction [119, 120, 123, 165] to obtain a metric tensor defined on some orbits of some group actions on the space of states of a finite-dimensional $C^\star$-algebra. This will allow us to reconstruct the Fisher-Rao metric tensor on the interior of the simplex, the Fubini-Study metric on the space of pure states of a quantum system and the Bures-Helstrom metric on the space of faithful states of a quantum system.

- In Chapter 3, the central idea of this work is introduced, i.e. ***parametric models of normal positive linear functionals*** on a $W^\star$-algebra, and we will see how, and to what extent, we can reproduce what presented in Section 2.3 for the infinite-dimensional case, this is done in Section 3.2. Of course, whether this answers positively to Question 3, depends on whether these metrics are the appropriate ones, more on this will be said in Chapter 5. In Section 3.3 the behaviour of parametric models and of the Riemannian structures defined on them as in Section 3.2 is investigated. In particular, we prove a monotonicity result for these Riemannian structures under the maps that have the property of being the dual of a ultra-weakly continuous, completely positive, unital map between $W^\star$-algebras, recall that these were pointed as the appropriate answer to Question 2. This result gives a positive answer to Question 4.

- In Chapter 4, with the intent to give an answer to Question 5, some concepts of Estimation Theory are introduced in our framework. This will allow us to formulate what will be called a ***Pre-Cramér-Rao bound*** and a ***Pre-Helstrom bound***, which represent preliminary results hinting to a positive answer to question 5.

- In Chapter 5 an answer to Question 6 is given, and different examples of models are presented. We will explicitly see how our approach enables to reproduce the approach described in [13], hence showing that this framework succeeds in restricting to Classical Information Geometry, both in the finite and infinite dimensional case, when the $W^{\star}$-algebra considered is Abelian. We will then move to describe some quantum models and will be able to obtain some well-known monotone metrics for both the finite dimensional and the infinite-dimensional case.

  In Section 5.3, the concept of singular model is introduced and some examples of singular models are presented.

- Finally, we will draw the conclusions of this work, to see to what extend, and how, the body of this work answers the questions presented in this Introduction, and outline the possible outlooks for future investigation.

Acknowledgements

First of all, I would like to thank my supervisor, Prof. Jürgen Jost, for the stimulating scientific discussions, the immense support and the precious advices, both academic and human. None of this would have been possible without you, and I am glad and honoured that I had the opportunity of working with you.

Another big thank you goes to Florio Ciaglia, bigger academic brother, scientific sparring partner, precious collaborator, and just a genuinely great human being.

Also I would like to thank my fellow colleagues at MPI MiS, my coauthors, all the scientists I had the chance to discuss with in these years, and the administrative staff of MPI MiS.

Apart from the academic aspects, these years have also been an extraordinary human experience. While going through this huge rollercoaster, I always had somebody to share it with, somebody to help and support me, somebody to listen to me. This is just unvaluable, and my gratitude goes to all of you: my family, the old friends, the new ones, the ones that stayed and the ones that left. *Caminante no hay camino, sino estelas en la mar*, our *estelas* crossed at some point, and that makes me feel blessed.

Chapter 2

Mathematical preliminaries

The purpose of this chapter is to review some well-known mathematical structures and results, together with some less-known ones, that are needed for the development of the unified formalism for classical and quantum information geometry alluded to in the introduction. The mathematical background that we will assume to be known to the reader are basic concepts of functional analysis, theory of linear operators on Hilbert space, differential geometry and group theory. This chapter is also enriched with some examples that will hopefully help the reader in familiarizing with the concept introduced.

Let us now describe the content of each Section.

- Our take on Information Geometry is based on the use of parametric models of normal positive linear functionals on $W^\star$-algebras. For this reason, we will devote Section 2.1 to introduce basic definitions and results about $C^\star$-algebras and $W^\star$-algebras. First, $C^\star$-algebras will be introduced, particular focus will be given to properties of ***positive functionals***, then we will move to maps between $C^\star$-algebras, in particular, a concept of maps from a $C^\star$-algebra to another that encapsulates both the concept of ***stochastic maps*** (or ***Markov maps***) of Classical Information Theory and of ***quantum channels*** (or ***quantum operations***) of Quantum Information Theory will be defined. After all of this, we will also discuss how $C^\star$-algebras can be represented on the space of bounded operators of an Hilbert space, following what is called the GNS construction (see [38] and [65, Theorem I.9.6]).

  After this, $W^\star$-algebras will be introduced, particular attention will be given to the properties of ***normal functionals***, thus giving us the tools necessary to introduce parametric models of normal linear functionals on $W^\star$-algebras in Chapter 3. Questions about the representability of $W^\star$-algebras on the space of bounded operators of some Hilbert space will also be tackled here. General references for this section are standard textbooks on the subject, such as [9, 23, 26, 65, 74, 116, 117, 159, 170].

- In Quantum Information Geometry and in Geometric Quantum Mechanics, it is a useful approach to look at the geometric structure of the space of quantum states by studying the orbits of some group actions [37, 40, 96]. In particular, the groups $GL(\mathcal{H})$ of invertible linear operators on a Hilbert space $\mathcal{H}$ and $\mathcal{U}(\mathcal{H})$ of unitary linear operators on $\mathcal{H}$ can act on the space of quantum states associated to $\mathcal{H}$, and the orbits of such actions have relevant properties with respect to the rank and the spectrum of the states involved. With this idea in mind, in Section 2.2 we will introduce the concept of ***Banach-Lie group***, which looks to us as the natural generalization of the concept of Lie group to use in this context. In fact, roughly speaking, Banach-Lie groups are Lie groups that can be modelled over a (possibly infinite-dimensional) Banach space.

After having introduced basic concepts and notion about Banach-Lie groups, the Banach-Lie groups that are the natural generalization in this context of $GL(\mathcal{H})$ and $\mathcal{U}(\mathcal{H})$ will be introduced. We will then proceed to define actions of these groups on $C^\star$-algebras that are the generalization of the ones used in Quantum Information Geometry. As we will see, some of the properties of the actions of $GL(\mathcal{H})$ and $\mathcal{U}(\mathcal{H})$ usually used in the finite-dimensional context are transported to the infinite-dimensional case.

General references for this section are the recent work [50] and textbooks on the subject, such as [17, 39, 178].

- Finally, in section 2.3, a Riemannian structure on finite-dimensional $C^\star$-algebras that stems from the Jordan product that can be canonically defined on $C^\star$-algebras will be introduced. The content of this section can be, to some extent, regarded as the finite-dimensional version of what is done in the central part of this work. In fact also the Riemannian structure that we introduce in Chapter 3 will be linked to the anti-commutator product on our possibly infinite-dimensional $W^\star$-algebras. Having introduced the main character of this discussion, i.e. the Riemannian structure coming from the Jordan product, the rest of the section will be devoted to see how this can be used to reproduce some examples in both Classical and Quantum Information Geometry.

  The main reference for this section is the recent work [53].

2.1 $C^\star$-algebras and $W^\star$-algebras

The study of algebras of operators was motivated by understanding formal aspects of the theory of Quantum Mechanics. In [179] and in the celebrated series of works [135, 136, 137, 180], Murray and Von Neumann began to study what they called *rings of operators*, i.e. †-closed subalgebras of $\mathcal{B}(\mathcal{H})$ that are closed with respect to the weak operator topology. Now these structure are called *Von Neumann algebras*. More or less in the same period, in [112] and [113], Jordan, Von Neumann and Wigner started considering as relevant structure for the formal description of Quantum Mechanics what are now called Jordan algebras. A ***Jordan algebra*** is a nonassociative algebra with a multiplication $\{\cdot,\cdot\}$ that is commutative and satisfies the Jordan identity, i.e.

$$\{\{a,b\},\{a,a\}\} = \{a,\{b,\{a,a\}\}\}, \tag{2.1}$$

for all a and b in the algebra.

Then in [89], Gelfand and Naimark began to study a different kind of algebras of operators, i.e. norm-closed and †-closed subspaces of the space $\mathcal{B}(\mathcal{H})$ of bounded operator over an Hilbert space $\mathcal{H}$. These structures are now called $C^\star$-algebras since the weak operator topology is weaker then the topology induced by the norm, it is clear that every Von Neumann algebra is also a $C^\star$-algebra. A strong push in the interest towards $C^\star$-algebras and their use for Quantum Mechanics was the work of Segal in [161], where he laid a precise axiomatic formulation of Quantum Mechanics in terms of $C^\star$-algebras.

In [179], the *double commutant theorem* is proved, this theorem on one hand gives a characterization of Von Neumann algebras, on the other hand is also a powerful tool for understanding whether an operator O belongs to a given Von Neumann algebra $\mathscr{A} \subseteq \mathcal{B}(\mathcal{H})$. Let us quickly recall the content of the theorem here.

Let $\mathcal{X}$ be a subset of $\mathcal{B}(\mathcal{H})$, the ***null set*** of $\mathcal{X}$ is the subspace of vectors $\phi \in \mathcal{H}$ such that $X(\phi) = 0$ for all $X \in \mathcal{X}$. The ***commutant*** of $\mathcal{X}$ is the set

$$\mathcal{X}' = \{T \in \mathcal{B}(\mathcal{H}) : TX = XT \quad \forall X \in \mathcal{X}\}, \tag{2.2}$$

the commutant $(\mathcal{X}')'$ of $\mathcal{X}'$ is called ***bicommutant*** of X and is denoted by $\mathcal{X}''$. Also, $\mathcal{X}$ is said to be self-adjoint when it is closed under the action of the adjoint.

Theorem 1 (Double Commutant Theorem)**.** *Let A be a self-adjoint subalgebra of operators which has trivial null space. Then A is dense in its bicommutant A'' in the strong and in the weak operator topologies.*

Proof. See [9, Th. 1.2.1] □

In [89], the authors also give an algebraic characterization of $C^\star$-algebras that does not refer to any underlying Hilbert space. The same can be done also for Von Neumann algebras. In fact, one can define a $W^\star$-algebra as a $C^\star$-algebra that admits a predual, i.e. a space whose dual is our algebra, and then show that every $W^\star$-algebra can be represented as a †-closed, weakly closed subalgebra of $\mathcal{B}(\mathcal{H})$, i.e. a Von Neumann algebra. This was proved by Shoichiro Sakai in [157].

Our approach will be to give as definitions what historically came as algebraic characterizations of functional-analytic objects, but it is still useful to keep in mind that there exists this alternative approach. In fact many times in our examples the relevant algebras will be exactly algebras of operators on an Hilbert space, since these are the objects that are most fit to describe Information Geometry.

2.1.1 $C^\star$-algebras and positive linear functionals

Having given some historical background and settled some nomenclature, we are now ready to give a brief account on $C^\star$-algebras and $W^\star$-algebras. Let us start from the definition of a Banach algebra.

Definition 2 (Banach algebra)**.** *Let A be a Banach space over the field $\mathbb{C}$ of complex numbers, with norm given by $\|\cdot\|$. Let A be also an associative algebra over $\mathbb{C}$.*

If the product and the norm satisfy

$$\|ab\| \leq \|a\|\|b\| \qquad \forall a, b \in A, \tag{2.3}$$

then A is said to be a ***Banach algebra****.*

Remark 1. *The validity of (2.3) ensures that the multiplication operation is continuous in the topology on A induced by the norm.*

An involution on an algebra is a map $*$ from A into itself such that, for all a and b in A and for all λ in $\mathbb{C}$,

$$\begin{aligned} (a^*)^* &= a, \\ (a+b)^* &= a^* + b^*, \\ (\lambda a)^* &= \overline{\lambda} a^*, \\ (ab)^* &= b^* a^*. \end{aligned} \tag{2.4}$$

Typical examples of involutions are complex conjugation, transposition of a square matrix or taking the adjoint of an operator. $C^\star$-algebras are Banach algebras that admit an involution satisfying particular requirements.

Definition 3 (Involutive Banach algebras and $C^\star$-algebras)**.** *Let A be a Banach algebra, and let $*$ be an involution on A such that*

$$\|a^*\| = \|a\|, \tag{2.5}$$

then A is said to be an ***involutive Banach algebra****. If also holds*

$$\|a^* a\| = \|a a^*\|, \tag{2.6}$$

then A is said to be a **$C^\star$*-algebra***.

Remark 2. *A $C^\star$-algebra may not contain an identity, a $C^\star$-algebra that contains an identity is called* ***unital****. Since we are interested in introducing $W^\star$-algebras and, as we will see, all $W^\star$-algebras contain the identity, we will always assume that our $C^\star$-algebras are unital and will denote the identity by $\mathbb{I}$.*

Definition 4. *An element a of a $C^\star$-algebra A is said to be self-adjoint (or hermitian) if $a^* = a$, normal if $a^* a = a a^*$, a projection if it is self-adjoint and $h^2 = h$.*

The space of self-adjoint elements of A will be denoted by A_{sa}, it is a Banach space over the field $\mathbb{R}$ of real numbers, and there is a direct sum decomposition

$$A = A_{sa} \oplus i A_{sa}, \tag{2.7}$$

where clearly i is the imaginary unit. So that for every element a of A exists a unique decomposition given by

$$a = h + ik, \tag{2.8}$$

with h and k in A_{sa}. Said differently, A is the complexification [168, p. 75] of A_{sa}.

Example 1. *Typical examples of $C^\star$-algebras are:*

- *Complex numbers endowed with the usual sum, product, modulus and complex conjugation. The self-adjoint elements are given by real numbers.*
- *Complex $n \times n$ matrices endowed with the usual sum, the row-column product, the operator norm and the adjoint.*
- *Continuous, complex-valued functions on a compact space endowed with the pointwise sum and product, the sup-norm, and complex conjugation.*
- *Essentially bounded, complex valued functions on some measure space, endowed with the pointwise sum and product, the sup-norm, and complex conjugation.*
- *Bounded operators on a Hilbert space endowed with the usual sum, product, operator norm and the adjoint.*

An element g in A is said to be *invertible* if there exists g' in A such that $gg' = g'g = \mathbb{I}$, g' is said to be the *inverse* of g, is denoted with g^{-1} and, if it exists, is unique. The set of all invertible elements of A is denoted with $GL(A)$.

If $u \in A$ is such that $u^{-1} = u^*$, then u is called a *unitary* element of A, the set of all unitary elements of A is denoted with $\mathcal{U}(A)$. Both $GL(A)$ and $\mathcal{U}(A)$ have a group structure, this will be discussed in more detail in Section 2.2.

As already said, $C^\star$-algebras were introduced for describing algebras of observables in the context of Quantum Mechanics. The conceptual idea behind this approach is that observables are described in terms of self-adjoint operators on an Hilbert space, and the information on the possible outcomes of measurements on an observables is contained in its spectrum. This notion allows, among other things, to compute expectation value of observables. Regarding the classical case, on the other hand, we have that observables are described in terms of bounded functions, with the possible outcomes of measurements given by the image of such function. In this regard, it should be mentioned that in [122] Koopman has applied successfully this framework to Hamiltonian mechanics.

There exist a notion of spectrum for an element of a $C^\star$-algebra that allows to recover both notions.

Definition 5 (Spectrum of an element of a $C^\star$-algebra)**.** *Let $a \in A$, the set*

$$Sp_A(a) = \{\lambda \in \mathbb{C} : a - \lambda\mathbb{I} \quad \textit{is not invertible}\} \tag{2.9}$$

is called the ***spectrum*** *of a in A.*

The spectra of self-adjoint and unitary elements of an algebra satisfy particular conditions.

Proposition 1. *Let $h \in A_{sa}$ and $u \in \mathcal{U}(A)$, then*

- *$Sp_A(h) \subseteq \mathbb{R}$;*

- *$\lambda \in Sp_A(u)$ implies $\|\lambda\| = 1$.*

Proof. See [74, Prop. 1.3.9]. □

An element $p \in A$ is said to be *positive* if there exists $a \in A$ such that $p = a^* a$. Clearly, every positive element is self-adjoint.

Theorem 2. *The following statements are equivalent:*

- *p is positive;*
- *The spectrum of p is contained in the interval $[0, \infty)$;*
- *There exist a self-adjoint element h in A such that $p = h^2$.*

Proof. See [170, Theorem 6.1]. □

It is possible to prove (see [26, Theorem 2.2.10]) that h in the third statement of Theorem 2 can be chosen to be positive, and in that case it is unique; it is called the *square root* of p. Self-adjoint elements of A admit a decomposition in terms ofpositive elements.

Proposition 2 (Orthogonal decomposition)**.** *Let $a \in A_{sa}$ and let*

$$a_{\pm} = \frac{|a| \pm a}{2}, \tag{2.10}$$

*with $|a|$ being the absolute value of a self-adjoint a, defined as $|a| := \sqrt{a^2}$, this is well defined, since a^2 is positive. Then we have that $a = a^+ + a^-$ and $a^+ a^- = 0$. This decomposition is called **orthogonal decomposition**, a^+ (resp. a^-) is called the **positive** (resp. **negative**) part of a.*

Proof. See [26, Proposition 2.2.11]. □

Remark 3. *To quickly give intuition about this last proposition, consider as a $C^\star$-algebra the space $\mathcal{C}(X)$ of continuous functions on a compact space X, it is clear that in this setting the positive and negative parts coincide with the usual definition of positive and negative part of a function and Proposition 2 holds trivially.*

Since A is a Banach space, we can consider the Banach dual A^* of A, i.e. the space of norm-continuous linear functionals η on A. Let $\eta \in A^*$, we can define a norm on A^* given by

$$\|\eta\| = \sup_{\|a\| \leq 1} \{|\eta(a)|\}. \tag{2.11}$$

When equipped with this norm, A^* is also a Banach space [28]. Notice that from the definition of this norm we can immediately see that

$$|\eta(a)| \leq \|\eta\| \|a\|. \tag{2.12}$$

Let us define the adjoint functional η^* of η as that element of A^* such that

$$\eta^*(a) = \overline{\eta(a^*)} \qquad \forall a \in A, \tag{2.13}$$

this allows to define *self-adjoint* functionals, i.e. functionals such that $\eta^* = \eta$. The set of self-adjoint linear functionals is a real Banach space and will be denoted by A^*_{sa}.

Definition 6. *A functional ω on $\mathscr{A}$ is called* ***positive*** *if*

$$\omega(a^*a) \geq 0 \qquad \forall a \in A. \tag{2.14}$$

Remark 4. *The definition of a positive functional allows to define a partial order on the space of linear self-adjoint functionals as follows: given two self-adjoint functionals η_1 and η_2, we say that $\eta_1 \geq \eta_2$ if $\eta_1 - \eta_2$ is a positive functional.*

Positive functionals satisfy the general Cauchy-Schwarz inequality.

Lemma 1 (Cauchy-Schwarz inequality)**.** *Let ω be a positive linear functional over A. We have that*

$$\begin{aligned} &\overline{\omega(a^*b)} = \omega(b^*a) \\ &\omega(a^*b) = \overline{\omega(b^*a)}|\omega(a^*b)| \leq \omega(a^*a)\omega(b^*b) \end{aligned} \tag{2.15}$$

for all $a, b \in A$.

Proof. See [26, Lemma 2.3.10] □

This result allows to prove several properties regarding the relations between positivity, continuity and normalization of functionals.

Proposition 3. *Let ω be a positive linear functional over A. We have that:*

1. *ω is norm-continuous;*
2. *If A is unital, we have $\|\omega\| = \omega(\mathbb{I})$;*
3. *$|\omega(a)|^2 \leq \omega(aa^*)\|\omega\| \quad \forall a \in A$.*
4. *$|\omega(a^*ba)|^2 \leq \omega(a^*a)\|b\| \quad \forall a, b \in A$.*

Proof. For a proof of points 1 and 2 see [170, Lemma 9.9], while a proof of points 3 and 4 can be found in [26, Prop. 2.3.11]. □

Let a be an element of the $C^\star$-algebra A, it is possible to define the action of a holomorphic function f defined on a neighbourhood U_f of the spectrum $\mathrm{Sp}_A(a)$ of a on the element a of A. We set

$$f(a) := \frac{1}{2\pi i}\int_C f(\lambda)(\lambda - a)^{-1} d\lambda, \tag{2.16}$$

where C is a smooth simple closed curve in U_f enclosing $\mathrm{Sp}_A(a)$, notice that, with an abuse of notation, we keep denoting with f the function acting on A. Notice also that this definition does not depend on C because of Cauchy's theorem for contour integrals, for more details on this construction see [170, p. 9] or [23, p.18].

The relevance of the map defined in (2.16) is given by the following theorem.

Theorem 3 (Spectral mapping theorem)**.** *Let a be an element of a $C^\star$-algebra A, f be an holomorphic function defined on a neighbourhood of $Sp_A(a)$ and $f(a)$ defined as in (2.16). We have*

$$Sp_A(f(a)) = f(Sp_A(a)). \tag{2.17}$$

Moreover, if g is a function defined on a neighbourhood of $Sp_A(f(a))$, we have

$$(g \circ f)(a) = g(f(a)). \tag{2.18}$$

Proof. See [170, Prop. 2.8]. □

We will denote the space of positive linear functionals by $\mathcal{P}(A)$, from the geometrical point of view, this is a convex cone. A positive linear functional is said to be faithful if $\omega(aa^*) = 0$ implies $a = 0$. The space of faithful positive linear functionals is denoted by $\mathcal{P}_+(A)$.

A trace f on a $C^\star$-algebra is a positive linear functional acting on A such that

$$f(ab) = f(ba) \quad \forall a, b \in A. \tag{2.19}$$

A trace is said to be *finite* if $f(a) < \infty$ for all $a \in A$.

In Quantum Mechanics, states are described in terms of positive operators whose norm is 1. The natural generalization of these objects to the $C^\star$-algebraic setting are positive functionals whose norm (as defined in (2.11)) is 1. These are called *states* of the $C^\star$-algebra A. We will denote the space of states of the algebra A by $\mathcal{S}(A)$, and the space of faithful states of A by $\mathcal{S}_+(A)$.

If ω_1 and ω_2 are two states, we clearly have that a convex combination of the two is a state itself,

$$\omega = \lambda\omega_1 + (1-\lambda)\omega_2 \in \mathcal{S}(A) \quad \text{with} \quad 0 \leq \lambda \leq 1. \tag{2.20}$$

Also, we have that

$$\begin{aligned} \omega &\geq \lambda\omega_1, \\ \omega &\geq (1-\lambda)\omega_2, \end{aligned} \tag{2.21}$$

with respect to the partial order on functionals introduced in Remark 4.

Definition 7 (Pure states)**.** *A state is said to be* ***pure*** *if it cannot be written as a convex combination of other states. Or, taking advantage of the previous remark on majorization, a state ω is pure if the only positive linear functionals majorized by ω are of the form $\lambda\omega$ with $0 < \lambda < 1$.*

Let us conclude this subsection with some topological remarks.

There are two topologies that can be introduced on the dual A^* of a $C^\star$-algebra A, one is the already mentioned norm (or ***uniform***) topology, meaning that the open neighbourhood are given by the open balls in the norm. Another is the weak* topology, recall here that the weak* topology is the initial topology with respect to the family of maps

$$\phi_a : A^* \ni \eta \mapsto \phi_a(\eta) = \eta(a) \in \mathbb{C}. \tag{2.22}$$

I.e. it is the weakest topology that makes this family of maps continuous.

We now state a theorem that establishes some topological properties of the space of states and of pure states.

Theorem 4. *The space of states of a $C^\star$-algebra A is convex. Moreover it is weakly* compact if and only if the $C^\star$-algebra A is unital. In the case that A is unital, then it also holds that the extremal points of the space of states coincide with the pure states.*

Proof. See [26, Th. 2.3.15] □

We now end this Subsection with two simple examples, we will show how the space of probability distributions for a space of events with finite cardinality and the space of quantum states of a finite-dimensional quantum system can be seen as the space of states of some $C^\star$-algebra. These examples do not show the full power of the formalism, since we are dealing with a finite-dimensional setting, but will hopefully convince the reader that the notions given so far can be used in Classical and Quantum Information Geometry. To treat the infinite-dimensional case more care is needed, due to the technical difficulties of the case, hence such examples will be discussed in detail in Chapter 5.

Example 2 (Probability distributions for a space of events of finite cardinality)**.** *Let us consider the space $\mathbb{C}^n$ with the usual vector space structure and norm. Let us also equip this space with the Hadamard (component-wise) product and component-wise conjugation, with this structures this space is a $C^\star$-algebra, let us denote it with A_c^n, where the superscript n denotes the complex dimension of the algebra, while the subscript c stands for classical. In fact, as we will see, this is the appropriate setting for describing classical probability distributions for a space of events of finite cardinality. The choice of the product may seem quite odd, but it is actually quite natural if one looks at this algebra as an algebra of diagonal matrices, so that the Hadamard product coincides with the usual row-column product. More specifically, the algebra A_c^n can always be immersed as a subalgebra of $M_n(\mathbb{C})$ as follows*

$$i : A_c^n \ni v = (v_1, v_2, \dots, v_n) \mapsto \begin{pmatrix} v_1 & 0 & \dots & 0 \\ 0 & v_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & v_n \end{pmatrix} \in M_n(\mathbb{C}). \tag{2.23}$$

This will be used in Subsection 5.2.1 to build models starting from an unfolding of the space of quantum states [42, 46, 132].

The space of self.adjoint elements of this algebra is given by vectors with real components, and the cone of positive elements is given by vectors with real and non-negative components. Since A_c^n is a finite-dimensional $C^\star$-algebra, it is isomorphic to its dual.

Considering states means considering norm-1 positive functionals, taking into account Point 2 of Proposition 3, amounts to considering vectors with non-negative components that sum to 1. A brief moment of reflection is enough to see that these can be seen as the barycentric coordinates on the $(n-1)$-simplex, which is the appropriate background for Classical Information Geometry in the case of a finite sample space.

Example 3 (Space of Quantum States of a finite-dimensional quantum system)**.** *Let us consider the $C^\star$-algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on the finite dimensional Hilbert space $\mathcal{H}$, let also $\dim_{\mathbb{C}} \mathcal{H} = n$. We will assume that $\mathcal{H} = \mathbb{C}^n$ and thus $\mathcal{B}(\mathcal{H})$ will be given just by the space $M_n(\mathbb{C})$ of $n \times n$ matrices with complex entries. This can be done without loss of generality since every n-dimensional complex Hilbert space is isomorphic and isometric to $\mathbb{C}^n$. Actually an even stronger result holds, i.e. that if $\mathcal{H}$ is infinite-dimensional but separable, then it is isomorphic and isometric to the Hilbert space $l^2(\mathbb{C})$ [28, p.144, Remark 10]. The space A_{sa} of self-adjoint elements of A is clearly given just by self-adjoint operators.*

Again because of the finite-dimensional setting, $\mathcal{B}(\mathcal{H})$ is isomorphic to its dual. Thus functionals on $\mathcal{B}(\mathcal{H})$ can be seen again just as matrices and positive functionals acting on $\mathcal{B}(\mathcal{H})$ are just positive matrices. The norm of a positive functional ω is then just given by the trace of the matrix associated to it, see Point 2 of Proposition 3. This means that the space of states of the $C^\star$-algebra $\mathcal{B}(\mathcal{H})$ is given the space of positive matrices with trace 1, this is usually referred to as the ***space of quantum states*** *of an n-levels quantum system and denoted with $\mathscr{S}_n$. In [95], the authors show that the space $\mathscr{S}_n^k$ of states with rank k is a smooth and connected manifold of (real) dimension $2nk - k^2 - 1$, in particular the space $\mathscr{S}_n^n$ of maximal rank (i.e. faithful) states is (n^2-1)-dimensional and coincides with the open interior of $\mathscr{S}_n$. The spaces $\mathscr{S}_n^k$ are often referred to as the* ***strata*** *of the space of quantum states $\mathscr{S}_n$. As a consequence, the boundary of $\mathscr{S}_n$ is given by the disjoint union of all the strata of rank k with $1 \leq k < n$. The space $\mathscr{S}_n^1$ of pure (i.e. rank 1) states is contained in this boundary and represent the extremal points of the space of quantum states, in agreement with Theorem 4. In general, the boundary of $\mathscr{S}_n$ can not be a manifold, since it is the union of manifolds of different dimensions. Of course an exception to this is the case $n = 2$, usually referred to as the* ***qubit*** *and has been extensively studied for its applications to Quantum Information Theory and Quantum Computing [139, 150, 172]. Here the boundary needs to coincide with the space $\mathscr{S}_2^1$ of pure states. More specifically the space of quantum states is a closed sphere, usually referred to as* ***Bloch sphere****, and the pure states of the qubit lie on the spherical surface that is the boundary of the Bloch sphere.*

The space of pure states of the $C^\star$-algebra $\mathcal{B}(\mathcal{H})$ can be represented as complex rays on the Hilbert space $\mathcal{H}$, which in turn can be identified with rank-one projectors. This means that every pure state ρ on $\mathcal{B}(\mathcal{H})$ can be written, using the bra-ket notation, *as*

$$\rho = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}. \tag{2.24}$$

More details on this construction can be found, for example, in [1, 10, 38]. Considering Theorem 4, then any state can be then written as a convex combination of at most n pure states.

2.1.2 Maps between $C^\star$-algebras and the GNS construction

Until now we always worked with a single $C^\star$-algebra, of course, in order to develop a framework in which these concepts are used in Information Geometry, we need to define maps that let us move from a $C^\star$-algebra to another.

As we already mentioned in the introduction of this chapter, we want to introduce a notion of positivity of maps between $C^\star$-algebras that reproduces both the stochastic maps of Classical Information Theory and the quantum channels of Quantum Information Theory, this will then be used in Section 3.3 to prove a result on the monotonicity of our metric tensors that can be related to both Čencov theorem and Petz characterization.

Let now A and B be two unitals $C^\star$-algebras, a map $\phi : A \to B$ is said to be ***unital*** if it maps the identity of A in the identity of B.

A map $\phi : A \to B$ is said to be ***positive*** if it maps positive elements in positive elements. As an example, a positive functional on the $C^\star$-algebra A is a positive map from A to $\mathbb{C}$.

Proposition 4. *Let $\phi : A \to B$, with A and B two $C^\star$-algebras, be a positive map, then its dual map ϕ^* sends positive functionals to positive functionals and preserves the norm of positive functionals.*

Proof. First, let a be a positive element of A, and η a positive functional on B, we have

$$\eta(\phi(a)) \geq 0 \quad \forall a \in A_+, \tag{2.25}$$

since ϕ is a positive map. Then, using the definition of dual map,

$$\eta(\phi(a)) = (\phi^*\eta)(a). \tag{2.26}$$

Implying that $(\phi^*\eta)(a)$ is non-negative for all $a \in A_+$, meaning that $\phi^*\eta$ is a positive functional on A.

Let now $\mathbb{I}_A$ and $\mathbb{I}_B$ denote respectively the identity of the $C^\star$-algebras A and B, we have

$$\|\phi^*(\eta)\| = \phi^*(\eta)(\mathbb{I}_A) = \eta(\phi(\mathbb{I}_A)) = \eta(\mathbb{I}_B) = \|\eta\|. \tag{2.27}$$

Where we used the definition of dual map, point 2 of Proposition 3 and the fact that ϕ^* is positive. □

We will now show a small example that clarifies the fact a map ϕ^* which is the dual of a positive, unital map ϕ reproduces Markov matrices and Markov kernels between probability spaces.

Example 4. *Let us consider again Example 2, we have that the dual ϕ^* of a positive, unital map ϕ from A_c^m to A_c^n has to send positive functionals in positive functionals and preserve the norm of functionals, it is clear, recalling Example 2, that it can be represented as a matrix with non-negative entries and whose rows and columns sum to 1.*

Consider now a less trivial example, let $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ be two σ-finite measure spaces. Consider then the spaces $L^\infty(\mathbb{C}, \Omega_1, \mu_1)$ and $L^\infty(\mathbb{C}, \Omega_2, \mu_2)$ of complex-valued functions, respectively from Ω_1 and from Ω_2 that are respectively μ_1-essentially bounded and μ_2-essentially bounded. As already pointed out in Example 1, these are $C^\star$-algebras when equipped with the pointwise sum, product, and complex conjugation and with the sup-norm. The duals of these spaces are given by the spaces of all finitely additive, finite signed measures that are absolutely continuous with respect to respectively μ_1 and μ_2, equipped with the total variation norm [75, Theorem IV.8.16]. A map ϕ^* that associates positive functionals to positive functionals while preserving the norm of positive functionals in this case has to take a positive normalized measure on Ω_2 and send it into a positive normalized measure on Ω_1. A comparison between this and Definition 1 allows to immediately see that in this case ϕ^* defines a Markov kernel.

As already mentioned in the Introduction of this work, even if the concept of positive map in the $C^\star$-algebraic setting is enough to recover the stochastic maps of the classical case, the CPTP maps used in Quantum Information Theory, the notion of positivity introduced now is not enough to reproduce the behaviour with respect to ancillary systems of quantum channels, in order to reproduce this property of these maps we need a stronger requirement, i.e. complete positivity.

Given a $C^\star$-algebra A, we can construct the space $M_n(A)$ of $n \times n$ matrices with entries in A. It is quite immediate to check that this space can be given the structure of an involutive algebra, the product being the usual matrix product and the * operation being given by the transpose on the matrix together with the involution of A, i.e.

$$a^* = \begin{pmatrix} a_{11}^* & a_{21}^* & \dots & a_{n1}^* \\ a_{12}^* & a_{22}^* & \dots & a_{n2}^* \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n}^* & a_{2n}^* & \dots & a_{nn}^* \end{pmatrix} \tag{2.28}$$

with $a \in M_n(A)$. The involutive algebra $M_n(A)$ is also a $C^\star$-algebra, this can be seen by extending a faithful representation of A and extending it to $M_n(A)$ and showing that its image is a uniformly closed subalgebra of the space of bounded operators on some Hilbert space. For the details of this construction see [170, p. 192].

Definition 8 (n-positive and completely positive (CP) maps)**.** *Let A and B be two $C^\star$-algebras and ϕ be a linear map from A to B. We can construct the linear map*

ϕ_n from $M_n(A)$ to $M_n(B)$ defined by

$$\phi_n(a) = \begin{pmatrix} \phi_n(a_{11}) & \phi_n(a_{12}) & \dots & \phi_n(a_{1n}) \\ \phi_n(a_{21}) & \phi_n(a_{22}) & \dots & \phi_n(a_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_n(a_{n1}) & \phi_n(a_{n2}) & \dots & \phi_n(a_{nn}) \end{pmatrix}. \tag{2.29}$$

*The map ϕ is then said to be **n-positive** if ϕ_n is positive, and **completely positive** (CP) if it is n-positive for all n.*

We refer the reader to [166], where the notion of completely positive map is introduced and the fundamental result known as *Stinespring's dilation theorem* (or *Stinespring factorization theorem*) is proved, and to [36] where the notion is further investigated and *Choi's theorem on completely positive maps* is proved.

The next proposition clarifies what we mean when we say that the notion of complete-positivity is only relevant in the quantum case and not in the classical case.

Proposition 5. *A positive linear map from the $C^\star$-algebra A to the $C^\star$-algebra B is completely positive if either A or B is Abelian. In particular, any positive linear functional is a CP map.*

Proof. See [26, Corollary 3.5] □

This in turn implies that the notion of completely positive map we introduced not only is the appropriate kind of map to use in the quantum setting, but also reduces to positive maps (i.e. stochastic maps, see Example 4) in the classical case.

Remark 5. *In the following, and in particular when proving Proposition 30, we will use maps such that they are the dual of a completely positive, unital map. Notice here, that because of Propositions 4 and 5, and taking into account Example 4, these are the analogue of both the Markov mappings of Classical Information Geometry and of CPTP maps (or quantum channels) of Quantum Information Geometry.*

Finally, completely positive maps satisfy an important inequality.

Proposition 6. *Let A and B be two unital $C^\star$-algebras, let also $\phi : A \to B$ be a unital completely positive map. Then*

$$(\phi(a))^2 \leq \phi(a^2) \tag{2.30}$$

for all $a \in A_{sa}$.

Proof. The original proof in the context of operator algebras was given by Richard Kadison in [115]. While a stronger result, namely that the same inequality holds already for 2-positive maps, was obtained by Man-Duen Choi in [35]. □

We now want to move to the issue of representations of $C^\star$-algebras, this will bring us to discuss also what is usually referred to as ***GNS construction***. The first requirement that a representation should have is for sure the fact of behaving nicely with respect to the involution.

Definition 9 (*-homomorphisms)**.** *If a homomorphism of an involutive Banach algebra into another preserves the* * *operation, then it is called a* ****-homomorphism****.*

Proposition 7. *The following statements hold.*

- *A* **-isomorphism of a* $C^\star$*-algebra into another* $C^\star$*-algebra is an isometry.*
- *An isomorphism (not necessarily* $*$*-preserving) of* $C^\star$*-algebras is continuous.*

Proof. This statements follow directly from Propositions 5.2, 5.3 and 5.5 of [170]. □

We are now ready to give the definition of a *-representation.

Definition 10 (*-representation of a $C^\star$-algebra)**.** *A* **-homomorphism* π *of* A *into the space* $\mathcal{B}(\mathcal{H})$ *of bounded linear operator on some Hilbert space* $\mathcal{H}$ *is called a* ****-representation*** *and is denoted by* $\{\pi, \mathcal{H}\}$*. The Hilbert space* $\mathcal{H}$ *is called the* ***representation space*** *of* π*. A representation is said to be* ***faithful*** *if* $\pi(a) = 0$ *is equivalent to* $a = 0$*.*

Let ω be a positive linear functional on A and let us define the set

$$N_\omega := \{I \in A: \quad \omega(I^* I) = 0\}. \tag{2.31}$$

This is a left ideal, to see this, let $a \in A$ and $I \in N_\omega$, we have that

$$0 \leq \omega\left((aI)^* aI\right) = \omega(I^* a^* a I) \leq \|a\|^2 \omega(I^* I) = 0, \tag{2.32}$$

where we used point 4 of Proposition 3, showing that N_ω is closed under left-multiplication of elements of the algebra, the other requirements for a left ideal are immediately checked. It is called the ***Gelfand ideal*** of the functional ω. It is then possible to define a quotient space A/N_ω, the equivalence classes in this quotient are defined by

$$a_\omega = \{\tilde{a} \in A : \tilde{a} = a + I, \quad I \in N_\omega\}. \tag{2.33}$$

The set of equivalence classes has a complex vector spaces structure if one equips it with the addition and scalar multiplication they inherit from A,

$$\begin{aligned} a_\omega + b_\omega &= (a+b)_\omega, \\ \lambda a_\omega &= (\lambda a)_\omega. \end{aligned} \tag{2.34}$$

It is also possible to define a bilinear functional on the quotient space A/N_ω as

$$(a_\omega, b_\omega) := \omega(b^* a), \tag{2.35}$$

with $a \in a_\omega$ and $b \in b_\omega$. This product is well defined on the quotient space, i.e. it does not depend on the choice of the elements in the equivalence classes. In fact, let I_1 and I_2 be in N_ω, we have that

$$\omega((a + I_1)^*(b + I_2)) = \omega(a^* b) + \omega(I_1^* b) + \omega(a^* I_2) + \omega(I_1^* I_2). \tag{2.36}$$

But all terms containing either I_1 or I_2 can be shown to be zero using Lemma 1, then we have

$$\omega((a + I_1)^*(b + I_2)) = \omega(a^*b), \tag{2.37}$$

thus showing that the product does not depend on the choice of the elements in the equivalence class and giving to A/N_ω the structure of a pre-Hilbert space. We can then complete this space with respect to the scalar product and obtain a Hilbert space, that will be denoted by $\mathcal{H}_\omega$ and called the ***GNS Hilbert space*** of the $C^\star$-algebra A associated to the positive functional ω.

We can then define a *-representation of A on $\mathcal{H}_\omega$, that will be denoted with $\{\pi_\omega, \mathcal{H}_\omega\}$, as

$$\pi_\omega(a)b_\omega = (ab)_\omega, \tag{2.38}$$

where $(ab)_\omega$ is the equivalence class containing ab. First notice that this definition is again independent on the representative chosen for the class b_ω. Notice also that, for all $a, b, c \in A$ and for all $\lambda \in \mathbb{C}$,

$$\begin{aligned} \pi_\omega(a)(\lambda b_\omega + c_\omega) &= \pi_\omega(a)(\lambda b + c)_\omega = (\lambda ab + ac)_\omega \\ &= \lambda(ab)_\omega + (ac)_\omega = \lambda\pi_\omega(a)b_\omega + \pi_\omega(a)c_\omega, \end{aligned} \tag{2.39}$$

hence $\pi_\omega(a)$ is a linear operator. Then we can use again point 4 of Proposition 3 to see that

$$\begin{aligned} \|\pi_\omega(a)b_\omega\|^2 &= ((ab)_\omega, (ab_\omega)) = \omega(b^*a^*ab) \\ &\leq \|a\|^2\omega(b^*b) = \|a\|^2(b_\omega, b_\omega) \\ &= \|a\|^2\|b_\omega\|^2, \end{aligned} \tag{2.40}$$

showing that $\pi_\omega(a)$ is a bounded operator on A/N_ω. This means that $\pi_\omega(a)$ so it can be uniquely extended to the completion $\mathcal{H}_\omega$ of A/N_ω, it can be easily checked that this is indeed a *-representation of A with representation space $\mathcal{H}_\omega$.

Definition 11 (Universal *-representation)**.** *Let $\mathscr{S}(A)$ denote the space of states of the $C^\star$-algebra A, and let $\rho \in \mathscr{S}(A)$. We can then construct the *-representation $\{\pi_\rho, \mathcal{H}_\rho\}$ associated to ρ and then consider the direct sum of the representation spaces for all ρ, i.e.*

$$K = \sum_{\rho\in\mathscr{S}(A)} \mathcal{H}_\rho. \tag{2.41}$$

This is a Hilbert space with respect to the product

$$(\xi, \eta) = \sum_{\rho\in\mathscr{S}(A)} (\xi_\rho, \eta_\rho), \tag{2.42}$$

where the product on the left hand side is the product on K, while the one on the right is performed on each $\mathcal{H}_\rho$. Define then the map

$$U(a) = \sum_{\rho\in\mathscr{S}(A)} \pi_\rho(a). \tag{2.43}$$

*This can be shown to be a *-representation of A on K, $\{U, K\}$ is called **universal *-representation** of A.*

The relevance of the universal *-representation of a $C^\star$-algebra is given by the Gelfand-Naimark theorem [89].

Theorem 5 (Gelfand-Naimark). *The universal *-representation of A is an isometric isomorphism.*

Proof. See [170, Th. 9.18] □

This means that every $C^\star$-algebra is *-isomorphic to a uniformly closed self-adjoint (†-closed) subalgebra of $\mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$. This connects the algebraic approach taken here with the early approach to $C^\star$-algebras as uniformly closed self-adjoint subalgebras of $\mathcal{B}(\mathcal{H})$.

Let us conclude this subsection with a result about the GNS construction for the $C^\star$-algebra $\mathcal{B}(\mathcal{H})$ with $\mathcal{H}$ finite-dimensional.

Proposition 8. *Let ρ be a pure state of the $C^\star$-algebra $\mathcal{B}(\mathcal{H})$ with $\mathcal{H}$ finite-dimensional and let $\{|e_1\rangle, |e_2\rangle, \ldots, |e_n\rangle\}$ be a basis of $\mathcal{H}$. This means, as we saw in Example 3, that exists a vector $|\psi\rangle$ in $\mathcal{H}$ such that*

$$\rho = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}. \tag{2.44}$$

Then:

- *The Gel'fand ideal is given by all linear operator of the form*

$$a = \sum_{k\neq 1} a_{jk} |e_j\rangle\langle e_k| . \tag{2.45}$$

- *The GNS Hilbert space associated with ρ can be identified with $\mathcal{H}$ itself, and the and the universal *-representation of $\mathcal{B}(\mathcal{H})$ on $\mathcal{H}$ sends $\mathcal{B}(\mathcal{H})$ to $\mathcal{B}(\mathcal{H})$ itself.*

- *The universal *-representation is irreducible [173, pp. 32-33] and its commutant, i.e. the commutant of its image (see Equation (2.2)), is given by the operators of the form $\lambda\mathbb{I}$ with $\lambda \in \mathbb{C}$ and $\mathbb{I}$ the identity operator on $\mathcal{H}$.*

Proof. See [53]. □

2.1.3 $W^\star$-algebras

In chapter 3, we will introduce a framework for parametric models in Information Geometry that relies on the use of *normal functionals* in $W^\star$-algebras. For this reason, we will now introduce $W^\star$-algebras in this subsection and focus in particular on properties of normal functionals.

Definition 12 ($W^\star$-algebra). *A $C^\star$-algebra $\mathscr{A}$ is called a* ***W***-algebra*** *if it admits a predual, meaning that there exists a Banach space, that will be denoted by $\mathscr{A}_*$, such that its Banach dual is $\mathscr{A}$.*

Every Banach space can be canonically immersed in its double dual (i.e. the dual of its dual) and in the finite-dimensional case this immersion is also an isomorphism. This means that we can immerse, let us call this immersion map i, the predual $\mathscr{A}_*$ of the $W^\star$-algebra $\mathscr{A}$ in the dual of $\mathscr{A}$. Functionals that are such that they are the image via i of some element in the predual are called *normal*. This means that if η is normal there exists an element $\hat{\eta}$ of the predual $\mathscr{A}_*$ such that $i(\hat{\eta}) = \eta$. Recalling the definition of the canonical immersion map i, this means that

$$\eta(a) = a(\hat{\eta}) \qquad \forall a \in \mathscr{A}, \tag{2.46}$$

i.e. that the action of the functional η can be somehow reproduced by means of some elements of the predual of $\mathscr{A}$.

Remark 6. *In the finite-dimensional case every $C^\star$-algebra is a $W^\star$-algebra, every functional is normal, and the immersion i introduced above is actually a Banach space isomorphism.*

Let us now make two infinite-dimensional examples, so that the reader can appreciate the relevance of the notion of normal functional in this context.

Example 5 (Essentially bounded functions on a σ-finite measure space)**.** *Let (Ω, Σ, μ) be a σ-finite measure space. As we already mentioned in Examples 1 and 4, the space $L^\infty(\mathbb{C}, \Omega, \mu)$ of μ-essentially bounded complex valued functions on Ω is a $C^\star$-algebra whose dual is given by the space $\mathrm{ba}(\Omega, \Sigma, \mu)$ of finitely additive finite signed measures that are absolutely continuous with respect to μ.*

The predual of $L^\infty(\mathbb{C}, \Omega, \mu)$ is the space $L^1(\mathbb{C}, \Omega, \mu)$ of μ-integrable complex valued functions on Ω [28], thus $L^\infty(\mathbb{C}, \Omega, \mu)$ is a $W^\star$-algebra. It is well known that the space $L^1(\mathbb{C}, \Omega, \mu)$ can be immersed in $\mathrm{ba}(\Omega, \Sigma, \mu)$ as

$$i : L^1(\mathbb{C}, \Omega, \mu) \ni f \mapsto f d\mu \in \mathrm{ba}(\Omega, \Sigma, \mu). \tag{2.47}$$

This map represents the immersion of the predual of $L^\infty(\mathbb{C}, \Omega, \mu)$ into its dual $\mathrm{ba}(\Omega, \Sigma, \mu)$. Considering Radon-Nikodym theorem [141, 154], it is easy to see that this immersion fails to recover those finitely additive measures that are not σ-additive. I.e. in this case normal functionals are precisely the ones to which Radon-Nikodym theorem applies.

Example 6 (Bounded operators on a separable Hilbert space)**.** *The image $i(\mathcal{B}(\mathcal{H})_*)$ via the natural immersion map i in $\mathcal{B}(\mathcal{H})^*$ of the predual $\mathcal{B}(\mathcal{H})_*$ of $\mathcal{B}(\mathcal{H})$ can be identified with the Banach space $TC(\mathcal{H})$ of trace-class operators on $\mathcal{H}$. The identification can be achieved by means of the isometric linear map that sends $T \in TC(\mathcal{H})$ into $\eta \in i(\mathcal{B}(\mathcal{H})^*)$ whose action on $a \in \mathcal{B}(\mathcal{H})$ is given by $\eta(a) = Tr(Ta)$ (see [159, Th. 1.15.3])*

$$\eta(a) = Tr(Ta). \tag{2.48}$$

In other words, normal functionals in this case are precisely those functionals whose action can be represented by means of a trace-class operator.

Theorem 6. *A functional is normal iff it is continuous with respect to the weak* topology on* $\mathscr{A}$.

Proof. See [159, Theorem 1.13.2] □

Of course, since $W^\star$-algebras are particular $C^\star$-algebras, the definitions of positive linear functionals and states (See Subsection 2.1.1) apply also to this setting. Since we are interested in positive functionals and states that are also normal, it is useful to define the sets $\mathscr{P}$ of normal positive linear functionals (from now on also *n.p.l.f.s*) and $\mathscr{S}$ of normal states on $\mathscr{A}$.

Proposition 9 (Support projection)**.** *Let ω be a n.p.l.f., there exist a unique non-zero projection $\mathbf{p} \in \mathscr{A}$ such that*

$$\omega(a) = \omega(a\mathbf{p}) = \omega(\mathbf{p}a) = \omega(\mathbf{p}a\mathbf{p}) \tag{2.49}$$

*for all $a \in \mathscr{A}$, and such that ω is faithful when restricted to $\mathscr{A}_{pp} = \mathbf{p}\mathscr{A}\mathbf{p}$. This projection is called the **support projection** of a.*

Proof. See [171, lem. III.3.6, p. 134]. □

Given the support projection $\mathbf{p}$ of ω, we have the topological direct sum decomposition

$$\mathscr{A} = \mathscr{A}_{pp} \oplus \mathscr{A}_{pq} \oplus \mathscr{A}_{qp} \oplus \mathscr{A}_{qq}. \tag{2.50}$$

Here the topology we are referring to is the one induced by the $C^\star$-norm, $\mathbf{q}$ is given by $\mathbf{q} = \mathbb{I} - \mathbf{p}$ and we have set $\mathscr{A}_{pp} = \mathbf{p}\mathscr{A}\mathbf{p}$, $\mathscr{A}_{pq} = \mathbf{p}\mathscr{A}\mathbf{q}$, $\mathscr{A}_{qp} = \mathbf{q}\mathscr{A}\mathbf{p}$, and $\mathscr{A}_{qq} = \mathbf{q}\mathscr{A}\mathbf{q}$. We refer to $\mathscr{A}_{pp}$ as the ***support algebra*** of ω. Note that $\mathscr{A}_{pp}$ and $\mathscr{A}_{qq}$ are $C^\star$-algebras and that the involution $\dagger$ gives a Banach space isomorphism between $\mathscr{A}_{qp}$ and $\mathscr{A}_{pq}$.

Obviously, if ω is faithful, then $\mathbf{p} = \mathbb{I}$ so that $\mathscr{A}_{pp} = \mathscr{A}$ and all other summands in the right hand side of equation (2.50) vanish.

Remark 7. *The decomposition in equation (2.50) allows us to visualize the algebraic operations in $\mathscr{A}$ in terms of matrix operation. Specifically, given an arbitrary $a \in \mathscr{A}$, if we write*

$$a = \begin{pmatrix} a_{pp} = \mathbf{p}a\mathbf{p} & a_{pq} = \mathbf{p}a\mathbf{q} \\ a_{qp} = \mathbf{q}a\mathbf{p} & a_{qq} = \mathbf{q}a\mathbf{q} \end{pmatrix}, \tag{2.51}$$

it is a matter of direct inspection to check that $a + b$ and ab can be computed using matrix-like operations. It then follows that the Gel'fand ideal N_ω of ω (see Equation 2.31) can be written as

$$N_\omega = \mathscr{A}_{pq} \oplus \mathscr{A}_{qq}. \tag{2.52}$$

As for $C^\star$-algebras, also for $W^\star$-algebras is possible to represent them on the space of bounded operator of some Hilbert space.

Definition 13 ($W^\star$-homomorphism and $W^\star$-representation)**.** *Let $\mathscr{A}_1$ and $\mathscr{A}_2$ be two $W^\star$-algebras and ϕ be a *-homomorphism of $\mathscr{A}_1$ into $\mathscr{A}_2$. If ϕ is continuous with respect to the ultraweak topology of $\mathscr{A}_1$ and $\mathscr{A}_2$, it is said to be a* ***W*******-homomorphism****. A $W^\star$-homomorphism π of a $W^\star$-algebra $\mathscr{A}$ into $\mathcal{B}(\mathcal{H})$ is called a* ***W*******-representation****. As for representations of $C^\star$-algebras, we will call the Hilbert space $\mathcal{H}$ the* ***representation space*** *of π.*

We now state a result due to Sakai [157].

Theorem 7. *Every $W^\star$-algebra admits a faithful $W^\star$-representation.*

Proof. See [159, Th. 1.16.7]. □

Therefore, every $W^\star$-algebra is *-isomorphic to a weakly closed, self-adjoint (†-closed) subalgebra of $\mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$, i.e. to a Von Neumann algebra. This connects the definition of $W^\star$-algebras given in this Chapter with the notion of Von Neumann algebra as weakly closed and self-adjoint subalgebras of the space of bounded linear operators on some Hilbert space.

2.2 Banach-Lie groups

In order to understand the relevance of the theory of Banach-Lie groups for our discussion, let us start from a seemingly unrelated instance.

In the context of the geometry of the space of quantum states, group actions can be seen to play a central role. In particular, it can be seen that both the cone of positive operators $\mathcal{P}(\mathcal{H})$ and the space of states $\mathscr{S}(\mathcal{H})$ of a finite-dimensional Hilbert space $\mathcal{H}$ can be seen as the disjoint union of orbits of two different actions of $GL(\mathcal{H})$ on $\mathcal{B}(\mathcal{H})$ [95].

More specifically, let n be the (complex) dimension of $\mathcal{H}$ and let $k \leq n$. We denote by $\mathcal{P}^k(\mathcal{H})$ and $\mathscr{S}^k(\mathcal{H})$ respectively the set of positive operators of rank k and the set of quantum states of rank k. Let us notice here that $\mathscr{S}^1(\mathcal{H})$ represents the space of pure states of the Hilbert space $\mathcal{H}$ and $\mathscr{S}^n(\mathcal{H})$ is the space of faithful states of the Hilbert space $\mathcal{H}$. In [95], the authors show that $\mathcal{P}^k(\mathcal{H})$ is a smooth and connected manifold and that it is also an orbit of the action

$$\beta : GL(\mathcal{H}) \times \mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}(\mathcal{H}) : \quad \beta(g, A) := gAg^\dagger. \tag{2.53}$$

It can be seen that each $\mathscr{S}^k(\mathcal{H})$ is also the orbit of an action of $GL(\mathcal{H})$ that represents some sort of normalized version of the one introduced in (2.53), namely

$$\gamma : GL(\mathcal{H}) \times \mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}(\mathcal{H}) : \gamma(g, A) := \frac{gAg^\dagger}{Tr(gg^\dagger)}. \tag{2.54}$$

This is done for example in [43] and [96].

Another interesting instance about group actions on the space of quantum states is that we can also define an action of the group $\mathcal{U}(\mathcal{H})$ of unitary operators as

$$\delta : \mathcal{U}(\mathcal{H}) \times \mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}(\mathcal{H}) : \delta(U, A) := UAU^{\dagger}, \tag{2.55}$$

and it can be shown that the *manifolds of isospectral states*, i.e. submanifolds of some $\mathscr{S}^k(\mathcal{H})$ such that any two states on the same manifold of isospectral state share the same spectrum, are orbits of this action.

We believe this is enough to bring one to ponder whether something similar happens in our context with the groups $GL(A)$ and $\mathcal{U}(A)$ introduced in Section 2.1. The main difference being of course that here we deal with the technicalities of the infinite-dimensional setting, in order to do so, we need to introduce the notion of Banach-Lie group.

There is also another, more speculative, reason why Banach-Lie groups can be an interesting structure for our discussion and it is related to an unfolding procedure for the space of quantum states that is discussed in [42, 46, 132]. The rough idea behind it is the following: let $\mathcal{H}$ be the n-dimensional Hilbert space associated to an n-levels quantum system (See Example 3). It can be seen that any quantum state of this quantum system can be specified, in a non-unique fashion, by giving a probability vector with n-components and a unitary operator acting on $\mathcal{H}$. Referring now to Example 2, we saw that the space of probability n-vectors can be identified with the $n-1$-dimensional simplex Δ_{n-1}. Thus the spaces of the kind $O \times \mathcal{U}(\mathcal{H})$, where O is an open of the simplex Δ_{n-1}, can always be projected on the space of quantum states $\mathscr{S}_n$.

In Chapter 5 we will take advantage of this instance to define a family of parametric models of quantum states. As we will see, the interesting idea behind this construction is that the model contains Δ_{n-1}, which can be regarded in some sense as the classical skeleton of the considered quantum model. This has been done uniquely for the finite dimensional case, but one is brought to ponder whether it is possible to generalize this construction to the infinite-dimensional setting. A preliminary step for such an analysis is to give an appropriate analogue of the group $\mathcal{U}(\mathcal{H})$ for the infinite-dimensional case, and the notion of Banach-Lie group helps us also in this effort.

For the aforementioned reasons, in this Section a brief introduction to the theory of Banach-Lie groups will be given, then we will introduce some actions of Banach-Lie groups on $C^{\star}$-algebras that can be seen as the generalization to the infinite-dimensional setting of the actions β, γ and δ introduced here, and give some properties of these actions. This can be seen as a preliminary step in giving a full picture of the interplay between group action and the geometry of the space of quantum states in the infinite-dimensional setting.

2.2.1 Some notions on Banach manifolds and Banach-Lie groups

In order to study the orbits of $GL(A)$ and $\mathcal{U}(A)$ on the cone of positive functionals or on the space of states of a $W^\star$-algebra we need to introduce some basic notions on Banach manifolds and Banach-Lie groups. In particular, we want to be able to take quotients of the groups $GL(A)$ and $\mathcal{U}(A)$ with respect to the isotropy group at some point, to establish a bijection with the orbits and give to them the structure of a homogeneous Banach space. All the theoretical machinery necessary to this purpose will be introduced in this Subsection.

A Banach-Lie group is a Lie group that is also a Banach manifold. Let us recall briefly some notions about Banach manifolds, some references on this matter are [2, 39, 125, 178]. Roughly speaking, a Banach manifold is a manifold that is modelled over some Banach space, meaning that the charts are homeomorphisms with some open subset of a Banach space as target space and the transition maps between different charts are bianalytic, see [178, p. 36] for a more precise definition. As an example, every open subset of a Banach space E is a Banach manifold and it can be modelled over the Banach space E itself.

A mapping f between two Banach manifolds M and N is called *analytic* if its representation in terms of local charts is an analytic mapping between the Banach spaces over which M and N are modelled is an analytic mapping between Banach spaces. If f admits an inverse which is also an *analytic mapping*, then f is said to be *bianalytic*.

The concepts of tangent space, differential of a mapping, and tangent bundles for Banach manifolds are introduced in an analogous way to how it is done in standard differential calculus. It can then be seen that the tangent space T_mM of M at one of its points m is also a Banach space, that the tangent bundle TM can be given the structure of a Banach manifold, that the canonical projection $\pi_m : TM \to M$ is an analytic mapping and that the differential Tf of an analytic mapping $f : M \to N$ is an analytic mapping from TM to TN [178, p. 58, 59]. A vector field X over M is said to be *analytic* if it is an analytic mapping from M to TM and the diagram

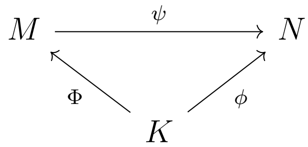
$$\begin{array}{ccc} TM & & \\ X \big\uparrow & \searrow^{\pi_M} & \\ M & \xrightarrow{\mathrm{id}_M} & M \end{array}$$

commutes.

Definition 14. *Let $f : M \to N$ be an analytic mapping between Banach manifolds. Then*

- *f is called an **immersion** if the differential T_mf of f at M is injective and $T_mf(T_mM)$ is a split subspace of $T_{f(m)}N$ for all $m \in M$;*
- *f is called a **submersion** at m if the differential T_mf of f at M is surjective and $\mathrm{Ker}T_mf$ is a split subspace of T_mM for all $m \in M$.*

Proposition 10. *Let M,N and K be Banach manifolds and consider the following diagram:*



If the maps Φ,ψ and ϕ are such that the diagram commmutes, then we have that:

- *If Φ is a analytic submersion and ϕ is analytic, then ψ is analytic;*
- *If Φ is a analytic submersion and ϕ is an analytic submersion, then ψ is an analytic submersion;*
- *If Φ is continuous and ψ is an immersion, and ϕ is analytic, then Φ is analytic;*
- *If Φ is continuous and ψ is an immersion, and ϕ is an analytic immersion, then Φ is an analytic immersion;*
- *The composition of analytic submersions is an analytic submersion;*
- *The composition of analytic immersions is an analytic immersion.*

Proof. See [178, p. 125] □

The usual notions of submanifold and quotient space can be extended to this case.

Definition 15 (Banach submanifolds)**.** *Let $N \subseteq M$, with M a Banach manifold modelled over the Banach space E, let also p be a point of N, U a neighbourhood of p. If for all $p \in N$ there exists a chart (U, ϕ, E) and a split subspace F of E such that*

$$\phi(N \cap U) = F \cap \phi(U), \tag{2.56}$$

*the subset N is said to be a **Banach submanifold** of the Banach manifold M.*

A submanifold N of a Banach manifold M is a Banach manifold in the relative topology induced on N from the same topology that makes M into a Banach manifold.

Proposition 11. *If $g : M \to N$ is an analytic immersion and is a homeomorphism onto $N' := g(M)$, then N' is a submanifold of N and the mapping $g' : M \to N'$ is bianalytic.*

Proof. See [178, Prop. 8.7] □

Let R be an equivalence relation on a Banach manifold M, the fact that M/R is a Banach manifold is not assured, but we have the following theorem that gives a condition for this.

Theorem 8. *Let R be an equivalence relation on a Banach manifold M, $\pi_R : R \to M/R$ be the canonical projection and $\pi_k : R \to M$ for $k = \{1, 2\}$ be the canonical projections on the two factors. let us also endow M/R with the quotient topology.*

The following conditions are equivalent:

- *R is a closed submanifold of $M \times M$ and the projections π_k are analytic submersions;*
- *M/R is a Banach manifold and the canonical projection π_R is an analytic submersion.*

Proof. See [178, Th. 8.14]. □

Having briefly recalled some basic notions on Banach manifolds, we are now ready to give the definition of a Banach-Lie group.

Definition 16 (Banach-Lie group)**.** *A **Banach-Lie group** over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ is a group G such that the underlying set is a Banach manifold over $\mathbb{K}$ and such that the map from $G \times G$ to G given by the group operation is analytic.*

In the theory of Lie groups, a central role is played by the Lie algebra of the Lie group, this idea can also be extended in this framework.

Definition 17 (Banach-Lie algebra)**.** *Let $\mathfrak{g}$ be a Banach space and a Lie algebra with respect to the product $[\cdot, \cdot]$. If the product is continuous with respect to the norm topology, then $\mathfrak{g}$ is called a **Banach-Lie algebra**.*

Remark 8. *The continuity of the Lie product is equivalent to the existence of a positive constant C such that*

$$\|[X, Y]\| \leq C\|X\|\|Y\| \quad \forall X, Y \in \mathfrak{g}. \tag{2.57}$$

The Lie algebra of right-invariant (left-invariant) analytic vector fields of a Banach-Lie group G is a Banach-Lie algebra [39, Theorem 2.1.26], and it is called the Banach-Lie algebra of G. It is then possible to define as customary the *exponential mapping* as the convergent power series $\exp : A \to GL(A)$ defined by

$$\exp(a) := \sum_{n=0}^{\infty} \frac{a^n}{n!}, \tag{2.58}$$

and it is possible to show that it is an analytic mapping [178, Corollary 6.5].

The usual notions of submanifold and quotient space can be extended to this case.

Definition 18 (Banach-Lie subgroup)**.** *A subgroup K of a Banach-Lie group G is called a **Banach-Lie subgroup** if K is also a submanifold of G in the sense specified in Definition 15.*

If a subgroup K of a Banach-Lie group G is closed, then there exists a unique Hausdorff topology that makes it a Banach-Lie group, and the algebra

$$\mathfrak{k} = \{a \in \mathfrak{g} : \quad \exp(ta) \in K \quad \forall t \in \mathbb{R}\} \tag{2.59}$$

is its Lie algebra and it is a closed subalgebra of $\mathfrak{g}$. However, this topology may not coincide with the relative topology on K induced by the topology that makes G a Banach manifold, hence K may not be a Banach-Lie subgroup of K. The following proposition provides a sufficient and necessary condition to show that a closed subgroup of a Banach-Lie group G is indeed a Banach-Lie subgroup of G.

Proposition 12. *Let K be a closed subgroup of a Banach-Lie group G with Banach-Lie algebra $\mathfrak{g}$. Then K is a Banach-Lie subgroup of G if and only if the algebra $\mathfrak{k}$ as in Equation (2.59) is a split subspace of $\mathfrak{g}$ and for every neighbourhood U of $\mathbf{0} \in \mathfrak{k}$ $\exp U \subseteq K$ is a neighbourhood of the identity of K.*

Proof. See [178, Lemma 8.12]. □

The following Theorem uses what said so far to give a useful criterion to say when the quotient of a Banach-Lie group with respect to a subgroup a Banach manifold. As already mentioned this is needed if one wants to study the orbits of such groups on the cone of positive functionals or on the space of states of some $W^\star$-algebra.

Theorem 9. *Let G be a Banach-Lie group, $\mathfrak{g}$ its Banach-Lie algebra, and K a Banach-Lie subgroup of G. Then, the quotient space $M \equiv G/K$ is a Banach manifold, the projection map $\pi : G \to M$ is an analytic submersion (see Theorem 8), and G acts analytically on M by the left translation. Moreover, consider the tangent map $T_{\pi_e} : \mathfrak{g} \to T_{\pi(e)}M$ of the projection π at the identity e, we have that the kernel of this map coincides with the split subalgebra $\mathfrak{k}$ of $\mathfrak{g}$ in equation (2.59) and we have $T_{\pi(e)}M \cong \mathfrak{g}/\mathfrak{k}$.*

Proof. The theorem can be proved using Proposition 10, Proposition 11, Theorem 9 and Proposition 12. See [178, Th. 8.19] for a complete proof. □

2.2.2 Positivity preserving action of $GL(A)$ on A_{sa}^*

The set $GL(A)$ introduced in Section 2.1 is an open subset of A [170, p. 5], hence it is a Banach manifold with respect to the relative topology induced by the strong topology of A. The product on $GL(A)$ is the restriction of a continuous bilinear mapping (see Remark 1) hence $GL(A)$ is a Banach-Lie group. Since it is modelled on A, its tangent space is A itself, moreover it can be shown that A, endowed with the commutator product, can be identified with the Banach-Lie algebra $\mathfrak{gl}(A)$ of $GL(A)$, see [17, Example 2.48] or [178, p. 96] for a more detailed discussion.

Following [50], let us introduce an action α of $GL(A)$ on A_{sa}^* as follows,

$$\alpha : GL(A) \times A_{sa}^* \longrightarrow A_{sa}^* : \alpha(g, \eta) := \eta_g \quad s.t. \quad \eta_g(a) = \eta(g^* a g) \quad \forall a \in A_{sa}. \tag{2.60}$$

We will now prove several properties of the action α.

Proposition 13. *α is a smooth, left action of $GL(A)$ on A^*_{sa}.*

Proof. Let us introduce another map, let us denote it by $\widetilde{\alpha}$, defined as follows,

$$\widetilde{\alpha} : A \times A^*_{sa} \longrightarrow A^*_{sa} : \widetilde{\alpha}(g, \eta) := \eta_g \quad s.t. \quad \eta_g(a) = \eta(g^* a g) \quad \forall a \in A_{sa}. \tag{2.61}$$

It is then clear that the map α can be expressed in terms of $\widetilde{\alpha}$ as

$$\alpha = \widetilde{\alpha} \circ (i_{GL(A)} \times \mathrm{id}_{A^*}), \tag{2.62}$$

where $i_{GL(A)}$ is the canonical immersion of $GL(A)$ in A and id_{A^*} is the identity on A^* (more precisely, the identity on A^* restricted to A^*_{sa}) which are both smooth maps. Thus in order to prove the smoothness of α it suffices to prove that $\widetilde{\alpha}$ is smooth.

In order to prove that $\widetilde{\alpha}$ is smooth, let a and b be two elements of A, η be a self adjoint functional, and let us define a functional η_{ab} whose action on an element c of A is given by

$$\eta_{ab}(c) := \frac{1}{2} \left(\eta(a^* c b) + \eta(b^* c a) \right). \tag{2.63}$$

It is easily checked that η_{ab} is a self-adjoint functional for all $a, b \in A$.

Let us then introduce a map $F : (A \times A^*_{sa}) \times (A \times A^*_{sa}) \times (A \times A^*_{sa}) \to A^*_{sa}$ as

$$F(a, \eta; b, \xi; c, \zeta) = \frac{1}{3}(\eta_{bc} + \xi_{ac} + \zeta_{ab}). \tag{2.64}$$

It can be directly checked that F is a bounded multilinear map and that

$$\widetilde{\alpha}(a, \eta) = F(a, \eta; a, \eta; a, \eta). \tag{2.65}$$

This implies that $\widetilde{\alpha} : A \times A^*_{sa} \to A^*_{sa}$ is a continuous polynomial map, which in turn implies that it is smooth (see [39, p. 63]). ☐

Proposition 14. *The action α is such that:*

- *α preserves the positiveness and faithfulness of linear functionals;*
- *α preserves the set of self-adjoint, normal, linear functionals of a $W^\star$-algebra $\mathscr{A}$.*

Proof. We will only prove that α preserves normality of self-adjoint, linear functionals, the other properties of α in the statement can be checked by direct inspection.

Let $b \in A$ and let us introduce the left and right multiplication maps on A,

$$\begin{aligned} l_b &: A \longrightarrow A, \quad l_b(a) := ba, \\ r_b &: A \longrightarrow A, \quad r_b(a) := ab, \end{aligned} \tag{2.66}$$

it is then clear that

$$\alpha(g, \eta) = \eta \circ l_{g^*} \circ r_g. \tag{2.67}$$

Normal functionals are weakly*-continuous (see Theorem 6), thus the fact that α preserves the weak* continuity (hence the normality) of functionals comes directly from the fact that both l_b and r_b are weak* continuous. ☐

Let now $O \subset A_{sa}^*$ be an orbit of α, then recall that the isotropy subgroup of $\eta \in A_{sa}^*$ with respect to the action α of $GL(A)$ is defined as

$$G_\eta = \{g \in GL(A) : \quad \alpha(g, \eta) = \eta\}. \tag{2.68}$$

We can then define a map $i_\eta : G/G_\eta \to \mathcal{O}$ as

$$i_e([g]) = \alpha(g, e), \tag{2.69}$$

and this defines a bijection of the quotient G/G_η to $\mathcal{O}$ for all $\eta \in \mathcal{O}$. Thus, if G/G_η can be endowed with the structure of a homogeneous Banach manifold, that this structure can be transported to O.

Recall (see Theorem 9) that we saw that the quotient space has a Banach manifold structure if G_η is a Banach-Lie subgroup of $GL(A)$, for this reason, it is reasonable to investigate the conditions under which G_η is actually a Banach-Lie subgroup of $GL(A)$.

An helpful notion in this regard is that of ***algebraic subgroup*** of $GL(A)$.

Definition 19 (Algebraic subgroup of $GL(A)$)**.** *A subgroup K of $GL(A)$ is called algebraic of degree k if there exist a family F of Banach space-valued continuous polynomials on $A \times A$ of degree k such that*

$$K = \{g \in GL(A) : f(g, g^{-1}) = 0 \quad \forall f \in F\}. \tag{2.70}$$

In fact it can be shown [178, p. 117] that if a subgroup K of $GL(A)$ is algebraic, then it is a closed subgroup of $GL(A)$, that it is a real Banach-Lie group with respect to the relative norm topology and that its Lie algebra

$$\mathfrak{k} = \{a \in \mathfrak{k} \cong A : \quad \exp(ta) \in K \quad \forall t \in \mathbb{R}\} \tag{2.71}$$

is a closed subalgebra of $\mathfrak{gl}(A)$. Also, according to [100, p. 667], it is possible to show that whenever K is an algebraic subgroup of $GL(A)$, then $\exp(V)$ is a neighbourhood of the identity in K for every neighbourhood V of $\mathbf{0} \in \mathfrak{k}$. Let us now prove, following [50], that G_η is an algebraic subgroup of $GL(A)$ for all $\eta \in A_{sa}^*$.

Proposition 15. *Let $\eta \in A_{sa}^*$, then the isotropy subgroup G_η is an algebraic subgroup of $GL(A)$ of order 2.*

Proof. Let $F_\eta = \{f_{\eta,a}\}_{a \in A}$ be the family of complex-valued polynomials of order 2 defined as

$$f_{\eta,a}(b, c) := \eta(a) - \eta(b^* a b), \tag{2.72}$$

where the dependence of the polynomial on c is trivial. Since η is a norm-continuous linear functional on A, then also $f_{\eta,a}$ is continuous for all $\eta \in A_{sa}^*$ and for all $a \in A$. It is then immediate to see that $f_{\eta,a}(g, g^{-1}) = 0$ amounts to say $\alpha(g, \eta) = \eta$, meaning that we can write

$$G_\eta = \{g \in GL(A) : f_{\eta,a}(g, g^{-1}) = 0 \quad \forall f_{\eta,a} \in F_\eta\}. \tag{2.73}$$

□

Thus, recalling Proposition 12 and having in mind what said here, in order to prove that G_η is a Banach-Lie subgroup of $GL(A)$, it suffices to prove that $\mathfrak{g}_\eta$ is a split subspace of $\mathfrak{gl}(A) \cong A$.

Let us now summarize the results obtained in [50].

- Whenever $\dim(A) < \infty$, the isotropy subgroup G_η is a Banach-Lie subgroup for all $\eta \in A^*_{sa}$, thus every orbit of α in A^*_{sa} is a homogeneous Banach manifold of $GL(A)$.
- If $A = \mathcal{B}(\mathcal{H})$ for a complex, separable Hilbert space $\mathcal{H}$ and η is a *n.p.l.f.*with arbitrary but finite rank, then G_η is a Banach-Lie subgroup of $GL(A)$ and the orbits are made of *n.p.l.f.s*. Moreover, given two *n.p.l.f.s* η_1 and η_2 with the same finite rank, there always exist an element g of $GL(A)$ that connects them via the action α.
- If A, admitting a faithful normal trace η, the Lie algebra $\mathfrak{g}_\eta$ coincides with the space of skew-adjoint elements in $\mathfrak{gl}(A) \cong A$, and this is always a split subspace of $\mathfrak{gl}(A)$. Thus GL_η is a Banach-Lie subgroup and the orbits have the structure of homogeneous Banach manifold of $GL(A)$. It can then be proved that such orbits can be put in a one-to-one correspondence with the set of positive, invertible elements in A and that if A is finite-dimensional, this coincides with the whole space of faithful, positive linear functional on A. An example of this kind of algebras are the II_1-factors [57]

2.2.3 State preserving action of $GL(A)$ on A^*_{sa}

It can be easily seen that the action α introduced in (2.60) does not preserve the space of states of a $C^\star$-algebra. Meaning that if one acts with α on a state ρ we have

$$\rho_g := (\alpha(g,\rho))(\mathbb{I}) = \rho(g^*g) \tag{2.74}$$

which can be different from 1 for an arbitrary g, hence ρ_g is not a state. This is certainly an undesired property if one wants to deal only with states, as one usually does in Probability Theory and Quantum Mechanics. One possibility is to define a map

$$\Phi : GL(A) \times \mathscr{S}(\mathscr{A}) \ni (g,\rho) \to \Phi(g,\rho) \in \mathscr{S}(\mathscr{A}), \tag{2.75}$$

that represents a kind of "normalized" version of α, meaning that

$$\left(\Phi(g,\rho)\right)(a) := \frac{\rho(g^*ag)}{\rho(g^*g)}, \tag{2.76}$$

in analogy with the definition of the action γ in (2.54). To ensure that the definition in (2.76) is well-posed, we need the following proposition

Proposition 16. *Let A be a unital $C^\star$-algebra. We have*

$$\rho(g^*g) > 0 \tag{2.77}$$

for all $g \in GL(A)$ and $\rho \in \mathscr{S}(\mathscr{A})$.

Proof. See [50, Prop. 9] □

It can then also be seen that the map Φ preserves the space of states of the $C^\star$-algebra A.

This solution, however, presents some technical problems. In fact the action α can be proven to be a smooth, left action of $GL(A)$ on A^*_{sa} (see Proposition 13), while the action Φ can not be meaningfully extended to A^*_{sa}. The reason for that is that the denominator in (2.76) could go to zero if ρ is not a state. On the other hand, even if it can be seen to be a left action on $\mathscr{S}(\mathscr{A})$, there is no meaningful notion of smoothness that can be given to this map, since $\mathscr{S}(\mathscr{A})$ is lacking a differentiable structure.

As an additional note let us notice that, given $\lambda_1, \lambda_2 \in \mathbb{C}$ and $\rho_1, \rho_2 \in \mathscr{S}(\mathscr{A})$, we have

$$\Phi(g, \lambda_1\rho_1 + \lambda_2\rho_2) \neq \lambda_1\Phi(g, \rho_1) + \lambda_2\Phi(g, \rho_2). \tag{2.78}$$

Meaning that the action Φ does not preserve the convex structure of $\mathscr{S}(\mathscr{A})$.

Despite the technical difficulties given by the aforementioned structural properties of Φ, in [50], the authors are able to prove results that are completely analogous to the one proven for the action α in the previous Section.

2.2.4 Action of $\mathcal{U}(A)$ on A^*_{sa}

The group $\mathcal{U}(A)$ introduced in Section 2.1 is actually a Banach-Lie group with respect to the norm topology of the $C^\star$-algebra A, i.e. the same topology that makes $GL(A)$ into a Banach manifold. Its Lie algebra $\mathfrak{u}(A)$ is given by the of skew-adjoint elements of A, i.e.

$$\mathfrak{u}(A) := \{u \in A : h^* = -h^*\}, \tag{2.79}$$

that is a closed real subalgebra of $A \cong \mathfrak{gl}(A)$. It can be also shown that $\mathcal{U}(A)$ is a real Banach-Lie subgroup of $GL(A)$, for a proof of this statements see [178, Corollary 15.22, p. 255].

Let us define an action of $\mathcal{U}(A)$ on A^*_{sa} as

$$\zeta : \mathcal{U}(A) \times A^*_{sa} \longrightarrow \in A^*_{sa} \quad s.t. \quad \zeta(U, \eta)(a) = \eta(U^* a U). \tag{2.80}$$

Proposition 17. *The action ζ is a smooth left action of $\mathcal{U}(A)$ on A^*_{sa}.*

Proof. Let $i_U : \mathcal{U}(A) \longrightarrow GL(A)$ the identification of $\mathcal{U}(A)$ as a subset of $GL(A)$, It is immediate to see that

$$\zeta = \alpha \circ (i_U \times \mathrm{id}^*_A), \tag{2.81}$$

where id^*_A is the identity map on A^*. Thus ζ is a smooth map, being the composition of two smooth maps.

The fact that ζ is a left action can be checked with a direct computation,

$$\zeta(U_1, \zeta(U_2, \rho))(a) = \rho(U_2^* U_1^* a U_1 U_2) = \zeta(U_1 U_2, \rho). \tag{2.82}$$

□

Unlike the action Φ introduced in (2.76), the action ζ preserves the convex structure of $\mathscr{S}(\mathscr{A})$,

$$\begin{aligned}\zeta(U, \lambda_1\rho_1 + \lambda_2\rho_2)(a) &= (\lambda_1\rho_1 + \lambda_2\rho_2)(U^* a U) \\ &= \lambda_1\rho_1(U^* a U) + \lambda_2\rho_2(U^* a U) \\ &= \lambda_1\zeta(U, \rho_1)(a) + \lambda_2\zeta(U, \rho_2)(a).\end{aligned} \tag{2.83}$$

This action is particularly interesting when considering it as acting on the space of states of a particular kind of $W^\star$-algebras, i.e. III_1 factors, see [135] for the characterization of $W^\star$-algebras in terms of factors. In fact, Alain Connes and Erling Størmer have shown in [58] that a factor is of type III_1 if and only if the action defined here is topologically transitive in the norm topology. Let us stress here that factors of type III_1 are used in the Algebraic formulation of Quantum Field Theory [97, 99, 184].

2.3 Riemannian structures from Jordan product on finite-dimensional $C^\star$-algebras

This section is mainly based on [53]. We will see how we can endow the space of states of a finite-dimensional $C^\star$-algebra with a Riemannian structure that is intimately linked with the anti-commutator product on the algebra. This construction can be successfully used to reproduce metric tensors that are well-known in the Information Geometry community: the Fisher-Rao metric tensor [78] for the classical case, the Fubini-Study [84, 167] and the Bures-Helstrom metric tensor [30, 174, 175, 176, 177] for the quantum case. For this section, all the spaces discussed will be finite-dimensional, even when not specified.

Let A be a $C^\star$-algebra, the associative product of the algebra can be exploited to define the non-associative, antisymmetric product

$$[a, b] = \frac{1}{2i}(ab - ba), \tag{2.84}$$

It can be directly checked that this product satisfies the Jacobi identity, giving to A the additional structure of a ***Lie algebra*** [98]. In the case that the algebra A is Abelian, this product is clearly always equal to zero.

On the other hand, we can define also a non-associative, symmetric product as

$$\{a, b\} = \frac{1}{2}(ab + ba). \tag{2.85}$$

It is a matter of direct inspection to see that this is a non-associative product satisfying the Jordan identity, thus giving to the $C^\star$-algebra A the structure of a ***Jordan algebra*** (see Equation (2.1)). Clearly, in the case that A is an Abelian $C^\star$-algebra, this reduces to the usual product on the algebra.

For any element a of A_{sa}, we can canonically immerse it in its double dual A_{sa}^{**}, this means that we associate to a the functional acting on A_{sa}^* defined as

$$l_a(\xi) := \xi(a). \tag{2.86}$$

Since A_{sa} is finite-dimensional, this immersion is actually a bijection, this means that the differentials dl_a of the functions l_a on A_{sa}^* generate the cotangent space $T_\xi^* A_{sa}^*$ at all ξ.

Let us now recall (see Subsection 2.2.2) that we can define an action α on the space A_{sa}^* as

$$\alpha : GL(A) \times A_{sa}^* \ni (g, \eta) \mapsto \eta_g \in A_{sa}^*, \tag{2.87}$$

where the action of η_g is given by

$$\eta_g(a) = \eta(g^* a g), \tag{2.88}$$

and such that it is a smooth, left action of $GL(A)$ that preserves positivity and the faithfulness of linear functionals. Thus the action preserves the space $\mathcal{P}$ of positive functionals on A, and despite $\mathcal{P}$ does not have a smooth structure, it can be seen that it is the disjoint union of orbits of the action β, which are clearly homogeneous spaces and smooth submanifolds of A_{sa}^*. These orbits coincide with positive functionals of a given rank and the orbit of higher dimension is given by maximal rank, i.e. faithful, functionals.

The Lie algebra of the Lie group $GL(A)$ can be identified with A itself, which in turn satisfies $A = A_{sa} \oplus iA_{sa}$. This allows us to use the exponential map [98] to define the smooth curve $g(t)$ on $GL(A)$ as

$$g(t) = e^{\frac{t}{2}(a+ib)}, \tag{2.89}$$

where $a, b \in A_{sa}$. This curve clearly satisfies $g(0) = \mathbb{I}$ and with a straightforward computation allows to get the following expression for fundamental vector fields of the action α,

$$V_{ab} = Y_a + X_b, \tag{2.90}$$

where the vector fields Y_a and X_b are defined as

$$\begin{aligned} Y_a(l_c) &:= l_{\{a,c\}}, \\ X_b(l_c) &:= l_{[b,c]}. \end{aligned} \tag{2.91}$$

From this expression it can be seen that the fundamental vector fields of the action α are given by two terms that are related to the two products defined in Equations (2.85) and (2.84). Moreover, it can be easily seen that the fundamental vector fields X_b associated to the antisymmetric product are precisely the fundamental vector fields of the action ζ given by the restriction of α to the subgroup $\mathcal{U}(A)$ of $GL(A)$. In fact, the Lie algebra of $\mathcal{U}(A)$ can be identified with A_{sa}, and the exponential map for this Lie group is given by

$$U(t) = e^{\frac{it}{2} b} \tag{2.92}$$

with $b \in A_{sa}$. Then a straightforward computation brings that the fundamental vector fields W_a of the action ζ are given by

$$W_a = X_a. \tag{2.93}$$

Let now $\eta \in A_{sa}^*$ and let us exploit once again the antisymmetric product defined in Equation (2.84) to introduce the map

$$\Lambda : T_\eta^* A_{sa}^* \times T_\eta^* A_{sa}^* \longrightarrow \mathbb{C} \tag{2.94}$$

whose action is given by

$$(\Lambda(dl_a, dl_b))\,(\eta) := l_{[a,b]}(\eta) = \eta([a,b]). \tag{2.95}$$

Since the differential of the linear functions l_a with $a \in A_{sa}$ generate the whole cotangent space $T_\eta^* A_{sa}^*$, the map Λ can be extended to the bilinear and antisymmetric tensor

$$(\Lambda(df_1, df_2))\,(\eta) := \eta([df_1(\eta), df_2(\eta)]). \tag{2.96}$$

Notice also that this tensor has non-constant rank and it coincides with the null tensor whenever A is Abelian. If the $C^\star$-algebra considered is $\mathcal{B}(\mathcal{H})$, the tensor Λ coincide with the one defined in [31, 32, 37, 41, 43, 44, 45, 51].

It can be shown that the tensor Λ is a Poisson tensor and it can be inverted on the orbits of the action ζ, thus giving to these orbits the structure of homogeneous symplectic manifolds [53]. This can be seen as a particular instance of the so-called Kirillov-Kostant-Souriau construction [119, 120, 123, 165], where one construct a Poisson tensor starting from the coadjoint action of a Lie group and then shows that this can be inverted on the orbits of such action, thus giving to these orbits the structure of a homogeneous symplectic manifold.

Let $f : A_{sa}^* \longrightarrow \mathbb{R}$ be a smooth function, one can use the Poisson tensor Λ to associate to f the ***Hamiltonian vector field*** H_f as

$$H_f = \Lambda(df, \cdot), \tag{2.97}$$

it can then immediately be noticed that

$$H_{l_b} = X_b = V_{0b}, \tag{2.98}$$

i.e. that the Hamiltonian vector fields associated to the Poisson tensor Λ, when applied to the linear functions l_a, give back a term of the fundamental vector fields of α, which is precisely the term that is related to the action ζ of $\mathcal{U}(A)$.

So far, we have the following picture: the fundamental vector fields of the action α contain a term X_a, which is related to the anti-symmetric product of the algebra A, we then saw that the vector fields X_b can be seen as the Hamiltonian vector fields of linear functions with respect to a Poisson tensor defined using the same anti-symmetric product used in defining X_b. However, the fundamental vector fields of the action α also contain a term Y_a that is related instead to the symmetric product, and one can wonder whether this can be related with some symmetric tensor defined in terms of the Jordan product used to define Y_a. This instance has been investigated, in the case where $A = \mathcal{B}(\mathcal{H})$, also in [47, 48].

Following this idea, we define, at each $\eta \in A^*_{sa}$, the map

$$R_\eta : T^*_\eta A^*_{sa} \times T^*_\eta A^*_{sa} \to \mathbb{R} \tag{2.99}$$

whose action is given by

$$R_\eta(dl_a, dl_b) := l_{\{a,b\}}(\eta) = \eta(\{a, b\}). \tag{2.100}$$

We can then extend by linearity this to $T^*_\eta A^*_{sa}$ as done for Λ, obtaining

$$R_\eta(df, dg) := \eta\left(\{df(\eta), dg(\eta)\}\right). \tag{2.101}$$

We will refer to this tensor as the ***Jordan tensor***.

Remark 9. *The Jordan tensor $\mathcal{R}$ defines a generalized distribution $\mathcal{D} = \{\mathcal{D}_\omega\}_{\omega \in \mathscr{A}^*_{sa}}$ on $\mathscr{A}^*_{sa}$ as*

$$\mathcal{D}_\omega := \{\eta \in T_\omega \mathscr{A}^*_{sa} \cong \mathscr{A}^*_{sa} \mid \ \exists\, \mathbf{a} \in \mathscr{A}_{sa} \,:\, \eta(\mathbf{b}) = \mathcal{R}_\omega(\mathbf{a}, \mathbf{b})\ \forall\, \mathbf{b} \in \mathscr{A}_{sa}\}. \tag{2.102}$$

We will say more about the role of this distribution for our discussion in Section 5.2.

We can then consider the field R given by the collection of these forms, this is a symmetric smooth $(0, 2)$-type tensor field. Its symmetry descends trivially from the symmetry of the product $[\cdot, \cdot]$. If we consider the case $A = \mathcal{B}(\mathcal{H})$, it can be seen that the tensor field introduced here coincides with a contravariant field relevant in Geometric Quantum Mechanics and related to the GKLS evolution [94, 129] that is considered in [31, 32, 37, 41, 43, 44, 45, 51].

We will now see that this tensor is actually invertible on particular subsets of A^*_{sa}, namely, the orbits $\mathcal{O}$ of the action α introduced in (2.60). As we saw in Section 2.2, this action preserves the positivity of functionals, let us then consider an orbit $\mathcal{O}^+$ such that $\mathcal{O}^+ \subset \mathcal{P}$.

Let us denote with i the immersion map of the orbit $\mathcal{O}^+$ in A^*_{sa} and let us introduce the functions

$$l^+_a := i^* l_a. \tag{2.103}$$

Such functions are such that the set $\{dl^+_a(\omega)\}$ is an overcomplete basis for the cotangent space $T_\omega \mathcal{O}^+$ for all $\omega \in \mathcal{O}^+$. Thus allowing us to define a tensor $\mathcal{R}$ on $\mathcal{O}$ by setting

$$\mathcal{R}_\omega(dl^+_a, dl^+_b) := \omega(\{a, b\}). \tag{2.104}$$

and extending it by linearity, as done for R. We clearly have

$$\mathcal{R}(i^*\theta_1, i^*\theta_2) = i^*(R(\theta_1.\theta_2)) \tag{2.105}$$

for all 1-forms θ_1, θ_2 on A^*_{sa}. This tensor may be regarded as the restriction of R to $\mathcal{O}^+$.

Proposition 18. *The* $(0, 2)$*-type tensor* $\mathcal{R}$ *is symmetric, invertible, and positive.*

Proof. For a proof of this statement see [53, p. 10] □

Since $\mathcal{R}$ is invertible, we can then consider

$$G := \mathcal{R}^{-1}, \tag{2.106}$$

and it will be a Riemannian metric tensor on the orbit $\mathcal{O}^+$.

The definition of a metric allows us to define the ***gradient vector field*** D_f of a smooth function f as

$$G(D_f, X) = df(X), \tag{2.107}$$

for all vector fields X on $\mathcal{O}^+$, this expression can easily be inverted to get

$$D_f = \mathcal{R}(df, \cdot), \tag{2.108}$$

and it can then immediately be noticed that

$$D_{l_a} = Y_a = V_{0a}, \tag{2.109}$$

i.e. that the gradient vector fields associated to the metric tensor G, when applied to the linear functions l_a, give back a term of the fundamental vector fields of α. Considering the analogous derivation for the Hamiltonian vector fields defined via Λ, one can say that the action α, when considered at an infinitesimal level, acts in a way that is related to the action of the tensors Λ and G defined by means respectively of the Lie product and the Jordan product.

Remark 10. *In some sense, introducing the tensor* $\mathcal{R}$ *can be considered as some sort of Jordan-analogue of the Kirillov-Kostant-Souriau construction. Notice that this construction has no application when the* $C^\star$*-algebra* A *is Abelian, since the Poisson tensor vanishes identically, while its Jordan-analogue can be carried out also in the Abelian case. In fact, in the next subsection this will be done for an Abelian* $C^\star$*-algebra and we will see that this gives rise to the Fisher-Rao metric tensor.*

Proposition 19. *Let* X_a *be as in Equation* (2.91)*,* G *as in* (2.106) *and let us denote with* L_X *the Lie derivative along* X*. We have*

$$L_{X_a} G = 0. \tag{2.110}$$

I.e. that the metric G *is invariant under the action of the unitary group.*

Proof. For a proof of this statement see [53, p. 11] □

Remark 11. *In Section 3.3 we will prove Proposition 31, which can be considered as a generalization of the Proposition. In fact, we will see that the metrics descending from the Jordan product are invariant under the action of ultra-weakly continuous automorphisms of the algebra considered.*

Until now we worked with orbits contained in the cone of positive functionals, disregarding the normalization constraint needed to deal with states, this has been done to avoid cumbersome computation and show the results in a more clear way. However, since the goal of this section is to obtain a Riemannian metric tensor on the space of states of a finite-dimensional $C^\star$-algebra, we need to replicate this construction for states, i.e. norm-one positive linear functionals. This will be done in complete analogy with what done before and referring the reader to [53] for a more detailed discussion.

Let us now recall (see Subsection 2.2.3) that we can define an action Φ on A^*_{sa} as

$$\Phi : GL(A) \times A^*_{sa} \ni (g, \eta) \to \Phi(g, \eta) \in \mathscr{S}(\mathscr{A}), \tag{2.111}$$

that represents a kind of "normalized" version of α, meaning that

$$\left(\Phi(g, \eta)\right)(a) := \frac{\eta(g^* a g)}{\eta(g^* g)}. \tag{2.112}$$

This is a smooth action of $GL(A)$, it preserves positivity, the faithfulness and the norm of the linear functionals it acts on, thus this action preserves the space $\mathscr{S}(\mathscr{A})$ of states of A. Moreover, $\mathscr{S}(\mathscr{A})$ can be seen as the disjoint union of orbits of the action Φ, these orbits coincide the spaces of states of a given rank and the orbit of higher dimension is given faithful states. Let now ρ be a state and let us denote with $\mathcal{O}_1^+$ the orbit of Φ containing ρ, clearly this is a submanifold of the orbit $\mathcal{O}^+$ of α containing ρ. We will denote by i_1 the immersion of $\mathcal{O}_1^+$ in $\mathcal{O}^+$.

We can then restrict the linear functions defined in (2.86) to $\mathcal{O}_1^+$, i.e. define

$$e_a := i_1^* l_a, \tag{2.113}$$

and the vector fields

$$\mathbb{V}_{ab} = \mathbb{Y}_a + \mathbb{X}_b, \tag{2.114}$$

with

$$\begin{aligned} \mathbb{X}_a(e_c) &= e_{[a,c]}, \\ \mathbb{Y}_b(e_c) &= e_{\{b,c\}} - e_b e_c. \end{aligned} \tag{2.115}$$

can be seen to be the fundamental vector fields for the action Φ.

The vector fields $\mathbb{X}_a$ can clearly still be considered the Hamiltonian vector field associated to e_a by the tensor

$$\Lambda_1 := i_1^* \Lambda. \tag{2.116}$$

We can define the tensor

$$G_1 := i_1^* G, \tag{2.117}$$

which can be shown to be a metric on the space of quantum states. Let f be a smooth function defined on $\mathcal{O}_1^+$ and let us denote with $\mathbb{D}_f$ the gradient vector field associated to f via G_1, we have that

$$\mathbb{D}_{e_a} = \mathbb{Y}_a. \tag{2.118}$$

Finally, we have that a result analogous to Proposition 19 holds, i.e.

$$L_{\mathbb{X}_a} G_1 = 0. \tag{2.119}$$

In the next subsections we will see that the tensor G_1 gives exactly the Fisher-Rao metric tensor, the Fubini-Study metric tensor and the Bures-Helstrom metric tensor, when the appropriate choice of A and $\mathcal{O}$ is made.

2.3.1 The Fisher-Rao metric

Let now A be the finite-dimensional, Abelian $C^{\star}$-algebra $A = \mathbb{C}^n$. Consider the canonical basis $\{e^j\}$, with $j = 1, \dots, n$, we have

$$a = a_j e^j \tag{2.120}$$

with $a_j \in \mathbb{C}$ for all $a \in A$. The components a_j will all be real if and only if a is a self-adjoint element of A. We can then consider the dual basis of $\{e^j\}$ and give usual Cartesian coordinates on $A_{sa}^* \cong (\mathbb{R}^+)^n$, $\mathcal{P}$ is the positive orthant and $\mathscr{S}$ may be identified with the standard simplex Δ^n in $\mathbb{R}^n$ (see Example 2).

The linear functional l_a associated to $a \in A_{sa}$ is given by

$$l_a = a_j p^j, \tag{2.121}$$

and we have

$$l_{\{a,b\}} = \sum_{j=1}^{n} a_j b_j p^j. \tag{2.122}$$

It is then easy to see that the tensor R may be written as

$$R = \sum_{j=1}^{n} p^j \frac{\partial}{\partial p^j} \otimes \frac{\partial}{\partial p^j}, \tag{2.123}$$

so that we can consider the covariant tensor

$$G = \sum_{j=1}^{n} \frac{1}{p^j} dp^j \otimes dp^j, \tag{2.124}$$

whose restriction to states (i.e. to the standard n-simplex) gives exactly the Fisher-Rao metric tensor.

2.3.2 The Fubini-Study metric

We now consider the case where $A = \mathcal{B}(\mathcal{H})$ with $\mathcal{H}$ finite-dimensional and the orbit considered is given by all pure states. In fact, recall (see the introduction to Section 2.2) that the orbits of the action α of $GL(A)$ on quantum states gives the spaces of states of a given rank, thus pure (i.e. rank one) states are clearly an orbit of α.

Taking into account what said in Example 3, for any pure state ρ, there exist a vector $|\psi\rangle \in \mathcal{H}$ such that

$$\rho = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}, \tag{2.125}$$

realizing a diffeomorphism between the orbit of α given by pure states and the complex projective space of $\mathcal{H}$ [77].

Taking into account Proposition 31, then our metric tensor G_1 has to be the only metric tensor on the complex projective space which is invariant under the action of the unitary group, i.e. Fubini-Study metric [84, 167].

2.3.3 The Bures-Helstrom metric

Let us now consider the case $A = \mathcal{B}(\mathcal{H})$, with $\mathcal{H}$ a finite-dimensional Hilbert space, and the orbit of α considered is the one given by faithful states.

The trace defined on $\mathcal{B}(\mathcal{H})$ allows to establishes an isomorphism between A and its dual A^* that identifies the functional η with the linear operator $\hat{\eta}$ such that

$$\eta(a) = Tr(\hat{\eta}a) \tag{2.126}$$

for all $a \in A$. The space of faithful states $\mathscr{S}_+$ can be thus identified with the manifold of invertible positive operators with trace one and its tangent space $T_\rho\mathscr{S}_+$ at $\rho \in \mathscr{S}_+$ can be identified with the affine hyperplane A^0_{sa} of self-adjoint linear operators with trace zero.

Considering the aforementioned isomorphism between A and A^*, the definition of the gradient vector fields Y_a in Equation (2.91) can be rewritten as

$$Y_a(\eta) = \{\hat{\eta}, a\}, \tag{2.127}$$

Remark 12. *Let $\rho \in \mathscr{S}$, notice that the vector $Y_a(\rho)$ provides the geometric analogue of the so-called Symmetric Logarithmic Derivative, which is widely used in Quantum Estimation Theory [103, 143, 169].*

Let us define the constant vector field on A^*_{sa}

$$Z_a(\eta) = a \tag{2.128}$$

associated to the element $a \in A_{sa}$, let ω be in the space $\mathcal{P}_+$ of positive invertible operators and let us define the operator

$$A_\omega(a) := \{\omega, a\} = \frac{1}{2}(L_\omega + R_\omega)(a), \tag{2.129}$$

where L_ω and R_ω are the invertible linear operators on A_{sa} defined as

$$\begin{aligned} L_\omega : A_{sa} \longrightarrow A_{sa} \quad \text{s.t.} \quad L_\omega(a) := \omega a, \\ R_\omega : A_{sa} \longrightarrow A_{sa} \quad \text{s.t.} \quad R_\omega(a) := a\omega. \end{aligned} \tag{2.130}$$

Since ω is a positive, invertible operator on $\mathcal{H}$, we have that also A_ω is an invertible linear operator.

Then we clearly have

$$a = A_\omega^{-1}(A_\omega(a)) = A_\omega^{-1}(\{\omega, a\}), \tag{2.131}$$

meaning that

$$\begin{aligned} Z_a(\omega) &= Y_{A_\omega^{-1}(a)}(\omega), \\ Y_a(\omega) &= Z_{A_\omega(\omega)}. \end{aligned} \tag{2.132}$$

Finally, let us define the vector field $\mathbb{Z}_a$ as the vector field that is i_1-related to Z_a.

We can now easily compute, for any state ρ,

$$\begin{aligned} G_1(\mathbb{Z}_a, \mathbb{Z}_b)(\rho) &= (i_1^* G)_\rho(\mathbb{Z}_a(\rho), \mathbb{Z}_b(\rho)) \\ &= G_\rho(Z_a(\rho), Z_b(\rho)) \\ &= G_\rho(Y_{A_\rho^{-1}(a)}(\rho), Y_{A_\rho^{-1}(b)}(\rho)), \end{aligned} \tag{2.133}$$

where we used the definition of G_1, Equation (2.132), and the definition of $\mathbb{Z}_a$. Considering now the fact that our metric is given in terms of the Jordan product, we have

$$\begin{aligned} G_1(\mathbb{Z}_a, \mathbb{Z}_b)(\rho) &= G_\rho(Y_{A_\rho^{-1}(a)}(\rho), Y_{A_\rho^{-1}(b)}(\rho)) \\ &= \rho\left(\{A_\rho^{-1}(a), A_\rho^{-1}(b)\}\right), \end{aligned} \tag{2.134}$$

from which immediately follows

$$\begin{aligned} G_1(\mathbb{Z}_a, \mathbb{Z}_b)(\rho) &= Tr\left(\{\hat{\rho}, A_\rho^{-1}(a)\} A_\rho^{-1}(b)\right) \\ &= Tr\left(a A_\rho^{-1}(b)\right). \end{aligned} \tag{2.135}$$

Which is precisely the expression of the Bures-Helstrom metric as given in [72, 73, 176].

Chapter 3

Information Geometry on smooth parametric models of normal positive linear functionals on $W^\star$-algebras

In this chapter, the central idea of this work, the use of parametric models of normal positive linear functionals on a $W^\star$-algebra in Information Geometry, will be discussed. This idea has been introduced in [49], which is the main reference for this Chapter. Let us discuss in detail the content of each section.

- In Section 3.1, smooth parametric models of normal positive linear functionals will be introduced. This will immediately give rise to the need to discuss in detail the concept of ***tangent double cone***, a surrogate of the concept of tangent space that is relevant in this framework. We will also see how this structure is related to a concept of absolute continuity of linear functionals that is a generalization of the concept of absolute continuity at the heart of Radon-Nikodym theorem. This concept of absolute continuity for linear functionals is introduced, building on the previous works [76, 158, 162, 181], by Niestegge in [140] and developed further in [131], where an analogue of Radon-Nikodym theorem in this framework is proven.

- In Section 3.2 the Riemannian structure on smooth parametric models of normal positive linear functionals stemming from the product of the algebra, that we will call ***J-metric*** of the parametric model, will be introduced. As we will see, this will need to take care of some technical difficulties given by the infinite-dimensional setting. In particular, this will bring us to the introduction of the definitions of ***J-regular*** and ***J-smooth*** parametric models.

- In Section 3.3 a result about monotonicity of the ***J-metrics*** for ***J-smooth*** parametric models will be proved. This result can be linked to the equivariance of Fisher-Rao metric tensor in Classical Information Geometry, i.e. to Čencov theorem [34], and to the monotonicity of the metric tensors characterized by Petz in [149], that are the metric tensors used in Quantum Information Geometry.

3.1 Smooth parametric models of normal positive linear functionals

The use of parametric models is widespread in statistical inference [21, 88, 134], and the use of statistical inference is ubiquitous in physics and in science in general.

One reason for the use of parametric models is given by experimental constraints. It is a typical situation, when performing an experiment, to prepare a system in a state that is not completely determined, while still not being completely arbitrary, there are some assumption that are reasonable to make on the state of the experimental apparatus, which can be of different nature.

This can be conceptualized by imagining that the possible configuration of our experimental setting are specified by giving some set of parameters, that can assume a certain range of values. Each element of our model (i.e. any possible set of values of the parameters) then represents one of the possible configurations the apparatus can be in. This is an extremely powerful tool that allows to disregard the (typically unapproachable) complexity of *all possible configurations* while preserving the ability to consider different configurations for the system we are working on.

The same idea is found to be useful also from a purely theoretical point of view, many problems that are completely unapproachable in their most general setting become more tame when using parametric models. Again the simplification comes from disregarding the complexity given by considering all possible configurations and focusing on some family of possible solutions, parametrized by some set of parameters.

More specifically, in this chapter we will introduce smooth parametric models of *n.p.l.f.s*, the idea is that by doing so we can exploit the fact that the parameters live on a manifold to use the tools of differential geometry, even if the space of states of a $W^\star$-algebra is not a manifold.

Definition 20 (Smooth parametric model of normal positive linear functionals)**.** *Let $\mathscr{A}$ be a $W^\star$-algebra, $\mathcal{M}$ be a Banach manifold and $\mathrm{j}:\mathcal{M}\longrightarrow(\mathscr{A}_{sa})^*$ be a smooth map such that $\mathrm{j}(\mathcal{M})\subset\mathscr{P}$. Then the triple $(\mathcal{M},\mathrm{j},\mathscr{A})$ is called a **smooth parametric model of n.p.l.f.s**. If $\mathrm{j}(\mathcal{M})\subset\mathscr{S}\subset\mathscr{P}$, we refer to $(\mathcal{M},\mathrm{j},\mathscr{A})$ as a smooth parametric model of normal states.*

Remark 13. *In order to avoid repetitions of words, we write "parametric model" instead of "smooth parametric model of n.p.l.f.s" or "smooth parametric model of normal states".*

A relevant property of parametric models is whether they represent faithfully the parameter space on $\mathscr{P}$.

Definition 21 (Identifiable and locally identifiable models)**.** *A smooth parametric model of n.p.l.f.s (normal states)* $(\mathcal{M}, j, \mathscr{A})$ *is said to be* ***identifiable*** *if* j *is injective, and it is said to be* ***locally identifiable*** *if* $T_m\mathrm{j}$ *is injective for all* $m \in \mathcal{M}$*.*

Every identifiable model is locally identifiable, while the converse is not true. Indeed, as the name suggests, for every locally identifiable model $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ and for every point $m \in \mathcal{M}$, we can find an open neighbourhood U of m such that the model $(U, \mathrm{j}|_U, \mathscr{A})$ is identifiable.

The ultimate goal of this chapter is to introduce a Riemannian structure on our models, in order to do this we need to introduce some sort of surrogate of the tangent space that can be defined in our framework. In fact, a Banach manifold clearly admits a tangent space at each point. It would then be possible to consider the tangent map of the immersion map j introduced in Definition 20, and such a map would send a vector in $T_m\mathcal{M}$ to some vector in the tangent space of $\mathscr{P}$ at the point $\mathrm{j}(m)$. This idea can not be applied here, since $\mathscr{P}$ is not a Banach manifold, not even in the finite-dimensional case.

For this reason we will now devote some space to the study of a geometric object that can be used in our context to mimic the role of the tangent space, i.e. the *tangent double cone.*

3.1.1 The tangent double cone

Let $\mathcal{B}$ be a real Banach space, X be a subset (not necessarily a submanifold) of $\mathcal{B}$ and x be a point of X. We can exploit the fact that X is a subset of a Banach space to consider C^1 curves on X and consider the derivatives of such curves. The construction is quite similar to the construction of the tangent space of a manifold, but since X is not necessarily a submanifold of $\mathcal{B}$, the result of this construction will not, in general, be a vector space.

Definition 22 (Tangent Double Cone)**.** *Let* $\mathcal{B}$ *be a real Banach space,* X *be a subset of* $\mathcal{B}$ *and* $x \in X$*.*

If there exist $\epsilon > 0$ *and a* C^1*-curve* $\gamma : (-\epsilon, \epsilon) \longrightarrow \mathcal{B}$ *s.t.*

$$\begin{aligned} \gamma(t) \in X \quad \forall t \in (-\epsilon, \epsilon); \\ \gamma(0) = \xi; \\ \dot{\gamma}(0) = \eta. \end{aligned} \tag{3.1}$$

The ***tangent double cone*** T_xX *of* X *at* x *is the collection of all* $\eta \in \mathcal{B}$ *satisfying such properties. I.e.* $\eta \in T_xX$ *if there exist a* C^1*-curve passing through* x *that has* η *as derivative at the* x*.*

The tangent double cone is, from the geometrical point of view, a double cone, this can be easily checked by seeing that if $\eta \in T_xX$, any multiple of η can be obtained by reparameterization of the curve γ in Definition 22. As anticipated, the tangent double cone represents a surrogate of the concept of tangent space, in fact whenever the subset X in Definition 22 is also a submanifold of $\mathcal{B}$, the tangent double cone reduces to the ordinary tangent space.

Example 7. *Let us consider a simple but clarifying example of tangent double cone. Let γ_1 and γ_2 be two open, simple, smooth curves in $\mathbb{R}^3$, intersecting at a point P, see Figure 3.1. The set $X = \gamma_1 \cup \gamma_2$, endowed with the relative topology induced on*

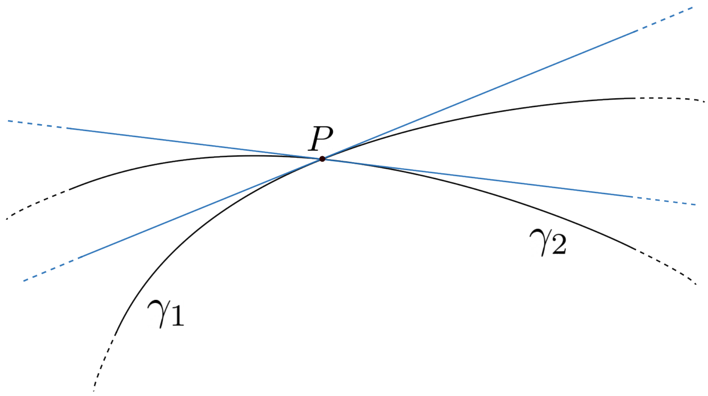


Figure 3.1: A simple example of tangent double cone; the union of the two blue straight lines represents the tangent double cone $T_P X$ of the set $X = \gamma_1 \cup \gamma_2$ at the point P.

X by the standard topology in $\mathbb{R}^3$, is not a manifold, to see this, consider an open neighbourhood of P and notice that this cannot be mapped continuously on an open of either $\mathbb{R}$ or $\mathbb{R}^2$. Nonetheless, we may consider the the tangent double cone of the set X.

The only C^1 curves passing through P that we can construct are exactly the ones following either γ_1 or γ_2. The tangent double cone $T_P X$ will then be given by the union of the tangent spaces $T_P\gamma_1$ and $T_P\gamma_2$. This is no longer a vector space, since the sum of an element of $T_P\gamma_1$ with an element of $T_P\gamma_2$ will not lay in $T_P X$, but it is still a double cone. Notice however then on a point $Q \neq P$ of γ_1 (or γ_2) the concept of tangent double cone and tangent space coincide completely.

The relevance of the tangent double cone in our discussion is even more evident from the following proposition regarding the tangent double cone of the set $\mathscr{P}$ of *n.p.l.f.s*.

Proposition 20. *Let $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ be a smooth parametric model of n.p.l.f.s according to definition 20. Then, $T_m\mathrm{j}(T_m M)$ lies in the tangent double cone $T_{\mathrm{j}(m)}\mathscr{P}$ for all $m \in \mathcal{M}$.*

Proof. The proof follows directly from the definition of tangent map. □

This circumstance motivates us to study in detail structure and properties of the tangent double cone $T_\xi\mathscr{P}$ of $\mathscr{P}$ at one of its points ξ. A first relevant property is that $T_\omega\mathscr{P}$ is actually a vector space containing only normal linear functionals.

Proposition 21. *Let $\mathcal{B}$ be a real Banach space, X be a subset of $\mathcal{B}$ and let also $\xi \in X$. Then, if X is a convex cone, $T_\xi X$ is a vector space. On the other hand, if V is a Banach subspace of $\mathcal{B}$ containing X, then $T_\xi X \subseteq \mathcal{V}$.*

Proof. Let X be a convex cone, η, ζ be in $T_\xi X$ and let also γ and σ be smooth curves in X starting at ξ and having respectively η and ζ as tangent vectors at $t = 0$. Consider then the domain $I_\gamma = (-\epsilon_\gamma, \epsilon_\gamma)$ of γ and let us define the interval $I_{a\gamma}$ as follows,

$$\begin{aligned} I_{a\gamma} &= (-\epsilon_\gamma/a, \epsilon_\gamma/a) \quad \text{if} \quad a > 0; \\ I_{a\gamma} &= (\epsilon_\gamma/a, -\epsilon_\gamma/a) \quad \text{if} \quad a < 0; \\ I_{a\gamma} &= I_\gamma \quad \text{if} \quad a = 0; \end{aligned} \tag{3.2}$$

The interval $I_{a\gamma}$ is the domain of the curve $\gamma_a(t) = \gamma(at)$, which has $a\eta$ as tangent vector at its starting point ξ. The same can be clearly done for I_σ.

Consider then two positive (real) numbers a and b, and let also $0 < \alpha < 1$, $\beta = (1 - \alpha)$, and $I_\mu = (-\epsilon_\mu, \epsilon_\mu)$, where

$$\epsilon_\mu = \min\left(\frac{\alpha}{a}, \frac{\beta}{b}\right). \tag{3.3}$$

Let us also define the curve $\mu\colon I_\mu \to \mathcal{B}$ as

$$\mu(t) := \alpha\gamma(at/\alpha) + \beta\sigma(bt/\beta). \tag{3.4}$$

Since X is a convex cone, $\mu(t)$ lies in X for every $t \in I_\mu$. It can be checked that μ is a smooth curve starting at ξ having $a\eta + b\zeta$ as tangent vector at $t = 0$. The cases in which a and b are either positive, negative or 0 can be handled similarly. It then follows that $T_\xi X$ is a vector space as claimed.

Now, assume there is a Banach subspace $\mathcal{V}$ of $\mathcal{B}$ such that $X \subseteq \mathcal{V}$ and let η be in $T_\xi X$. According to the definition of tangent double cone, exists a curve $\gamma(t)$ such that $\gamma(0) = \xi$ and $\dot\gamma(0) = \eta$, meaning that

$$\eta = \lim_{\epsilon \to 0} \frac{\gamma(\epsilon) - \gamma(0)}{\epsilon}. \tag{3.5}$$

Since $X \subseteq \mathcal{V}$ and $\mathcal{V}$ is a vector space, it is clear that $\gamma(\epsilon)$, $\gamma(0)$, $\gamma(\epsilon) - \gamma(0)$ are in $\mathcal{V}$. Recall that the limit in equation (3.5) is taken in the norm topology of $\mathcal{B}$. Being $\mathcal{V}$ a Banach subspace of $\mathcal{B}$, this topology coincides with the norm topology on $\mathcal{V}$. Thus we have that $\eta \in \mathcal{V}$ as claimed. □

Corollary 1. *The tangent double cone $T_\omega\mathscr{P}$ of $\mathscr{P}$ at ω is a vector space for every $\omega \in \mathscr{P}$ sitting inside $i\left(\mathscr{A}_{sa}\right)_*)$ for every $\omega \in \mathscr{P}$. I.e. it is a vector space containing only normal linear functionals.*

Proof. Recall that $\mathscr{P}$ is a convex cone and it is contained in $i\left((\mathscr{A}_{sa})_*\right)$ (recall that we denoted with i the canonical immersion of $\mathscr{A}_*$ in its double dual), and consider that $i\left((\mathscr{A}_{sa})_*\right)$ is a Banach subspace of $\mathscr{A}_{sa}^*$ [159, p. 29]. Then the claim follows from Proposition 21. □

Let us now turn briefly our attention to the classical case, in this case we have the following theorem.

Theorem 10. *The tangent double cone of the convex cone of positive-definite measures at any of its points ω coincides with the Banach space of signed measures that are absolutely continuous with respect to μ_ω. I.e., it coincides with $L^1(\mathcal{X}, \mu_\omega)$.*

The tangent double cone of the set of probability measures at any of its points ρ coincides with the Banach space of signed measures that are absolutely continuous with respect to μ_ρ and such that

$$\int_{\mathcal{X}} d\mu_\rho = 0. \tag{3.6}$$

Proof. See [13, thm. 3.1, p. 142]. □

This instance can bring one to wonder if this identification of absolute continuous functionals and elements in the tangent double cone exists at a more general level.

3.1.2 Absolute continuity of functionals and a Radon-Nikodym type theorem for functionals

A notion of absolute continuity in the context of bounded linear functionals on $C^\star$-algebras has been introduce by Niestegge in [140]. This represents a generalization of the measure-theoretic notion of absolute continuity to the case of $C^\star$-algebras.

Definition 23 (Absolute continuity)**.** *Let ω be a positive functional on a $C^\star$-algebra $\mathscr{A}$, $\eta \in (\mathscr{A}_{sa})^*$ and let also $\mathbf{B}_1$ be the unit ball in $\mathscr{A}$. Then η is said to be absolutely continuous with respect to ω if one of the following equivalent statements is true:*

- $\forall \epsilon > 0, \exists\, \delta > 0 \quad s.t. \quad \omega(a) < \delta \implies |\eta(a)| < \epsilon \quad \forall\, a \in \mathscr{A}_+ \cap \mathbf{B}_1.$
- *For every sequence $\{a_n\}_{n\in\mathbb{N}}$ with $a_n \in \mathscr{A}_+ \cap \mathbf{B}_1$, then $\lim \omega(a_n) = 0 \implies \lim \eta(a_n) = 0$.*
- *For every net $\{a_\lambda\}_{\lambda\in\Lambda}$ with $a_\lambda \in \mathscr{A}_+ \cap \mathbf{B}_1$, then $\lim \omega(a_\lambda) = 0 \implies \lim \eta(a_\lambda) = 0$.*

The set of all self-adjoint linear functionals absolutely continuous with respect to a given positive linear functional ω will be denoted by AC_ω.

Remark 14. *In the Classical case $\mathscr{A} = C_0(\mathcal{X})$, $\xi \in (\mathscr{A}_{sa})^*$ is in AC_ω iff μ_ξ is absolutely continuous with respect to μ_ω in the measure-theoretic sense [140, Prop. 1.3];*

A crucial property of absolute continuous functional in the classical case is given by Radon-Nikodym theorem, see [141, 154] or [24, Section 3.2]. Now we will state a theorem that somehow states a similar property in the more general setting of funcionals on a $C^\star$-algebra and that refers to the notion of absolute continuity introduced in Definition 23.

Theorem 11. *Let ω be a positive linear functional on $\mathscr{A}$, let us define the set*

$$J_\omega = \left\{ \eta \in (\mathscr{A}_{sa})^* \,|\, \exists\, a \in \mathscr{A}_{sa} \,:\, \eta(b) = \omega(\{\mathbf{a}, \mathbf{b}\}) \;\forall\, b \in \mathscr{A} \right\}. \tag{3.7}$$

The set AC_ω is a Banach subspace of $(\mathscr{A}_{sa})^*$ which is the closure of J_ω in the norm topology.

Proof. See [131, Thm. 1.3]. □

Remark 15. *The space J_ω is the analogous of the distribution defined by the Jordan product, as introduced in Remark 9, for this reason, in this chapter we will follow the idea of defining the action of our metrics.*

In this case, however, we have that this space is closed only in the finite-dimensional case. So we will have to investigate who is the closure of this space, and we will find out that this is precisely the tangent double cone, see Definition 22.

The following proposition gives two properties of AC_ω in the case that ω is a *n.p.l.f.* on a $W^\star$-algebra $\mathscr{A}$ that will be useful in the following.

Proposition 22. *If ω in Theorem 11 is also normal, then every $\xi \in J_\omega$ is normal [159, Sec. 1.8]. Then, since $i(\mathscr{A}_*)$ is norm-closed in $\mathscr{A}^*$, Theorem 11 implies that $AC_\omega \subseteq i\left((\mathscr{A}_{sa})_*\right)$. Meaning that every functional in AC_ω is normal.*

Moreover, $\xi \in i\left((\mathscr{A}_{sa})_*\right)$ is in AC_ω iff

$$\omega(p) = 0 \quad \Longrightarrow \quad \xi(p) = 0 \tag{3.8}$$

for every projection $p \in \mathscr{A}$.

Proof. See [140, Prop. 1.5]. □

Remark 16. *In the case $\mathscr{A} = L^\infty(\mathcal{X}, \mu)$, the result in Propostion 22 amounts to say that μ_ξ is absolutely continuous with respect to μ_ω in the measure-theoretic sense. To see this, consider that in this case projections in $\mathscr{A}$ are given by the characteristic functions on some measurable subsets of $\mathcal{X}$ with finite μ-measure.*

The set J_ω in equation (3.7) will turn out to be particularly useful for our investigation. We will refer to an element in J_ω, as η_a with $a \in \mathscr{A}_{sa}$, where the two are related via

$$\eta_a(b) = \omega(\{a, b\}) \tag{3.9}$$

for all $b \in \mathscr{A}_{sa}$.

The following proposition states the aforementioned link between the tangent double cone and the set AC_ω of bounded linear functional that are absolutely continuous with respect to a *n.p.l.f.* ω and with respect to the notion of absolute continuity given in Definition 23. As we will see, this is in complete analogy with what stated in Theorem 10 about the classical case.

Proposition 23. *Let ω be a n.p.l.f. on a $W^\star$-algebra $\mathscr{A}$, the norm-closure $\overline{T_\omega \mathscr{P}}$ of the tangent double cone $T_\omega \mathscr{P}$ at $\omega \in \mathscr{P}$ coincides with AC_ω.*

Proof. We will first prove that $AC_\omega \subseteq \overline{T_\omega \mathscr{P}}$.

Consider the smooth curve ω_t in $(\mathscr{A}_{sa})_*$ given by

$$\omega_t(b) = \omega(e^{\frac{ta}{2}} b \, e^{\frac{ta}{2}}), \tag{3.10}$$

we clearly have that $\omega_0 = \omega$ and $\omega_t \in \mathscr{P}$ for all $t \in \mathbb{R}$. By definition of tangent double cone, we have that the tangent vector $\dot{\omega}_t(b)\Big|_{t=0}$ of the smooth curve ω_t at $t = 0$ is in $T_\omega \mathscr{P}$.

A quick computation shows that

$$\dot{\omega}_t(b)\Big|_{t=0} = \frac{1}{2}\, \omega(ab + ba) = \eta_a(b), \tag{3.11}$$

with $\eta_a \in J_\omega$ as in (3.9), we thus have that ù$J_\omega \subseteq T_\omega \mathscr{P}$. Recalling that the set AC_ω is a Banach subspace of $(\mathscr{A}_{sa})^*$ which is the closure of J_ω in the norm topology (see Theorem 11), we have that

$$AC_\omega = \overline{J_\omega} \subseteq \overline{T_\omega \mathscr{P}}. \tag{3.12}$$

Let us now turn to proving that $\overline{T_\omega \mathscr{P}} \subseteq AC_\omega$.

Suppose ξ be a functional in the double tangent cone $T_\omega \mathscr{P}$ of $\mathscr{P}$ at ω, let $\gamma \colon (-\epsilon, \epsilon) \to (\mathscr{A}_{sa})_*$ be a smooth curve such that its image is contained in $\mathscr{P}$, we have $\gamma(0) = \omega$, and its tangent vector at $t = 0$ is precisely ξ. Consider then a projection $p \in \mathscr{A}$ such that $\omega(p) = 0$ and consider the function

$$f(t) = \omega_t(p) = (\gamma(t))(p). \tag{3.13}$$

This is a real-valued smooth function which is also non-negative and vanishes when $t = 0$, so that its derivative $f'(0)$ at $t = 0$ vanishes. It can be then directly checked that

$$f'(0) = \xi(p) = 0. \tag{3.14}$$

Thus we have that if $\xi \in T_\omega \mathscr{P}$, then $\omega(p) = 0$ implies $\xi(p) = 0$ for every projection $p \in \mathscr{A}$. According to Proposition 22 this amounts to say that $\overline{T_\omega \mathscr{P}} \subseteq AC_\omega$. □

Proposition 24. *Given a n.p.l.f. ω, the Banach space dual $(AC_\omega)^*$ of $AC_\omega = \overline{J_\omega}$ can be identified with*

$$\mathscr{A}_{sa}^\omega := (\mathscr{A}_{pp} \oplus \mathscr{A}_{pq} \oplus \mathscr{A}_{qp})_{sa}. \tag{3.15}$$

Proof. Let us denote with $\mathrm{Ann}(AC_\omega)$ the annihilator of AC_ω, which is defined by

$$\mathrm{Ann}(AC_\omega) = \{a \in \mathscr{A}_{sa} \colon \eta(a) = 0 \quad \forall \eta \in AC_\omega\}, \tag{3.16}$$

Since $AC_\omega \subseteq (\mathscr{A}_{sa})_*$ and $\mathscr{A}_{sa} \cong ((\mathscr{A}_{sa})_*)^*$, we have that

$$(AC_\omega)^* \cong \mathscr{A}_{sa}/\mathrm{Ann}(AC_\omega). \tag{3.17}$$

Let now $a \in \mathrm{Ann}(AC_\omega)$ then we have that

$$0 = \eta_a(a) = \omega(a^2), \tag{3.18}$$

which in turn implies that a lies in the Gel'fand ideal of ω (*cf.* equation (2.31)).

Since ω is a *n.p.l.f.*, we have that equation (2.52) holds, considering then the fact that a is self-adjoint, we have that $\mathrm{Ann}(AC_\omega) = (\mathscr{A}_{qq})_{sa}$. Plugging this in (3.17), we have

$$(AC_\omega)^* \cong \mathscr{A}_{sa}/(\mathscr{A}_{qq})_{sa} \cong (\mathscr{A}_{pp} \oplus \mathscr{A}_{pq} \oplus \mathscr{A}_{qp})_{sa} =: \mathscr{A}^\omega_{sa} \tag{3.19}$$

as claimed. □

3.2 Riemannian structure on parametric models of *n.p.l.f.s*

In this Section, we will introduce a Riemannian structure on the manifold $\mathcal{M}$ of a parametric model of *n.p.l.f.s* $(\mathcal{M}, \mathrm{j}, \mathscr{A})$.

The idea is then to introduce a product on the tangent double cone $T_\omega\mathscr{P}$, that is the analogous in our framework of the tangent space. Mimicking what done in Section 2.3 in the finite dimensional case, we want also this product to be related to the Jordan product of the algebra. Then, given a smooth parametric model $(\mathcal{M}, \mathrm{j}, \mathscr{A})$, one could define a map from $T_m\mathcal{M} \times T_m\mathcal{M}$ to $\mathbb{R}$ mapping tangent vectors at $m \in \mathcal{M}$ through the map $T_m\mathrm{j}$ and then exploit the product on the tangent double cone to perform the product between the images of vectors. In fact, according to 20, we have that the image of a vector in $T_m\mathcal{M}$ via the map $T_m\mathrm{j}$ is in the tangent double cone $T_{\mathrm{j}(m)}\mathscr{P}$ of $\mathscr{P}$ at $\mathrm{j}(m)$. As we will see, this idea cannot be really followed without taking care of some technical details given by the infinite-dimensional setting.

Consider a *n.p.l.f.* ω, let us define a map $G_\omega\colon J_\omega \times J_\omega \to \mathbb{R}$ as

$$G_\omega(\eta_a, \eta_b) := \omega(\{a, b\}), \tag{3.20}$$

where η_a and η_b are related respectively to a and b as in equation (3.9), notice that the same equation also implies

$$G_\omega(\eta_a, \eta_b) = \omega(\{a, b\}) = \eta_a(b) = \eta_b(a). \tag{3.21}$$

Proposition 25. *The map G_ω defines a bilinear, symmetric, positive definite product on each J_ω for every n.p.l.f. ω. Thus J_ω endowed with G_ω becomes a real pre-Hilbert space.*

Proof. First notice that the map sending a in η_a and b in η_b, which are related as in Equation (3.9), is a linear map. This, together with the linearity of ω, gives the bilinearity of G_ω. The symmetry of G_ω comes from the fact that the Jordan product is symmetric.

Let us move now to prove that G_ω is positive definite. Recalling equation (3.9), we have that

$$G_\omega(\eta_a, \eta_a) = \omega(a^2) = 0. \tag{3.22}$$

This implies (see equation (2.31)) that a is in the Gel'fand ideal of ω. Recall equation (2.52), and the fact that the involution gives a Banach space isomorphism between $\mathscr{A}_{pq}$ and $\mathscr{A}_{qp}$ (see equation (2.50)). Then the self-adjointness of a implies that $a \in \mathscr{A}_{qq}$, i.e.

$$a = q\, a\, q. \tag{3.23}$$

This in turn implies

$$\eta_a(b) = \omega(\{a, b\}) = \omega(p\,\{q\,a\,q, b\}\,p) = 0 \quad \forall b \in \mathscr{A}, \tag{3.24}$$

where we used Proposition 9.

We thus conclude that η_a is the null functional and thus G_ω is positive definite on each J_ω for every *n.p.l.f.* ω. □

Next, let us complete the space J_ω with respect to the inner product defined by G_ω.

Definition 24. *The closure of J_ω with respect to the topology determined by the inner product G_ω (cf. equation (3.20)) is a real Hilbert space, it is denoted by $\mathcal{J}_\omega$ and it is referred to as the **J-Hilbert space** of ω.*

Recall that J_ω is in general also not closed with respect to the norm topology inherited from $(\mathscr{A}_{sa})_*$, instead, according to Theorem 11, its norm closure is exactly the space AC_ω of self-adjoint normal linear functionals which are absolutely continuous with respect to ω (in the sense of Definition 23).

Proposition 26. *Let ω be a n.p.l.f. and $\mathcal{J}_\omega$ as in definition 24. We have that $\mathcal{J}_\omega$ is a subset of AC_ω which is dense in AC_ω with respect to the norm-topology.*

Moreover, the dual space $(AC_\omega)^* \cong \mathscr{A}_{sa}^\omega$ (see proposition 24) can be identified with a subset of $\mathcal{J}_\omega^*$ which is dense in $\mathcal{J}_\omega^*$ with respect to the norm-topology.

Proof. Let $\eta \in \mathcal{J}_\omega$, and let us define a linear function $l_\eta\colon \mathscr{A}_{sa} \to \mathbb{R}$ as

$$l_\eta(a) := G_\omega(\eta, \eta_a) \tag{3.25}$$

for every $a \in \mathscr{A}_{sa}$. Because of Cauchy-Schwarz inequality we have that

$$|l_\eta(a)|^2 = |G_\omega(\eta, \eta_a)|^2 \leq G_\omega(\eta, \eta)\, G_\omega(\eta_a, \eta_a) = ||\eta||^2_\omega\, \omega(a^2). \tag{3.26}$$

Then, using (2.12), we have

$$||\eta||^2_\omega\, \omega(a^2) \leq ||\eta||^2_\omega\, ||\omega||\, ||a||^2, \tag{3.27}$$

i.e., l_η is a bounded linear map, thus it can be looked at as an element of $\mathscr{A}^*_{sa}$. With an abuse of notation, we denote by η the element in $\mathscr{A}^*_{sa}$ determined by l_η.

Next, using the general Cauchy-Schwarz inequality for positive functionals (See Lemma 1), the definition of G_ω and again (2.12), we have

$$|\eta_a(b)|^2 \leq \omega(a^2)\,\omega(b^2) \leq G_\omega(\eta_a,\eta_a)\;||\omega||\;||b||^2. \tag{3.28}$$

Thus, the norm determined by the inner product on J_ω is stronger than the norm on J_ω inherited from $(\mathscr{A}_{sa})_*$, implying that $\mathcal{J}_\omega$ is a subset of $(\mathscr{A}_{sa})_*$.

Then, following the reasoning below equation (3.22), if $\mathbf{P}$ is a projection such that $\omega(\mathbf{P}) = 0$, then $\eta_{\mathbf{P}} = \mathbf{0}$. In turn implying

$$\eta(\mathbf{P}) = l_\eta(\mathbf{P}) = 0, \tag{3.29}$$

which, because of Proposition 22, means that $\eta \in AC_\omega$.

We showed that the norm determined by the inner product on J_ω is stronger than the norm on J_ω inherited from $(\mathscr{A}_{sa})_*$, thus the identification map $\mathrm{i}\colon J_\omega \to AC_\omega$ can be extended to a bounded linear map $\mathrm{i}\colon \mathcal{J}_\omega \to AC_\omega$ whose image is a dense subset of AC_ω, since it contains J_ω. We can then consider the dual map $\mathrm{i}^*\colon (AC_\omega)^* \cong \mathscr{A}_{sa}^\omega \to (\mathcal{J}_\omega)^*$ (see Proposition 24).

We will now prove that the map i^* is injective. First notice that

$$\left(\mathrm{i}^*(a-b)\right)(\eta) = \eta(a-b) = 0 \;\;\forall \eta \in \mathcal{J}_\omega \tag{3.30}$$

implies that

$$\left(\mathrm{i}^*(a-b)\right)(\eta_{a-b}) = \omega((a-b)^2) = 0, \tag{3.31}$$

implying that $(a-b)$ lies in the Gel'fand ideal of ω (see Equation (2.31)). Recall that ω is a *n.p.l.f.*, that $(a-b)$ is self-adjoint and equation (2.52), then we have that $(a-b) \in (\mathscr{A}_{qq})_{sa}$. But if we take into account the fact that

$$a, b \in \mathscr{A}_{sa}^\omega = (\mathscr{A}_{pp} \oplus \mathscr{A}_{pq} \oplus \mathscr{A}_{qp})_{sa} \tag{3.32}$$

then it must be $(a-b) = \mathbf{0}$, and thus i^* is injective.

Since it is a Hilbert space, $\mathcal{J}_\omega$ is isomorphic to its dual. Thus equation (3.30) implies that i^* has an image that is dense in $(\mathcal{J}_\omega)^*$. □

With the previous results, we provided $\mathcal{J}_\omega \subseteq AC_\omega$ with an inner product G_ω for every *n.p.l.f.* ω. However, as already noticed in the beginning of this section, we want to be able to transport this inner product to the tangent space $T_m\mathcal{M}$ of the manifold $\mathcal{M}$ of our parametric model using the tangent $T\mathrm{j}$ map of j, but the set $T\mathrm{j}(T_m\mathcal{M})$, while being for sure a subset of $T_{\mathrm{j}(m)}\mathscr{P}$, it is not necessarily contained in $\mathcal{J}_{\mathrm{j}(m)}$.

This motivates the following definitions.

Definition 25 (J-regular parametric model)**.** *A parametric model $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ is said to be **J-regular** if the image $T_m\mathrm{j}(T_m M)$ of the tangent space at $m \in \mathcal{M}$ through the map $T_m\mathrm{j}$ is a subset of the real Hilbert space $\mathcal{J}_{\mathrm{j}(m)}$ (see definition 24) for all $m \in \mathcal{M}$.*

For a J-regular model, one can define a map $\mathcal{G}_m$ from $T_m\mathcal{M} \times T_m\mathcal{M}$ to $\mathbb{R}$ as

$$\mathcal{G}_m(X, Y) := G_{\mathrm{j}(m)}\left(T_m\mathrm{j}(X(m)), T_m\mathrm{j}(Y(m))\right). \tag{3.33}$$

This map is bilinear and symmetric, since $G_{\mathrm{j}(m)}$ is bilinear and symmetric.

Definition 26 (J-smooth parametric models and J-metrics)**.** *A J-regular parametric model is said to be **J-smooth** if Equation* (3.33) *defines a smooth tensor field on* $\mathcal{M}$*.*

Let $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ *be a J-smooth, identifiable parametric model. The smooth tensor field defined by equation* (3.33) *is called the **J-metric** associated to the model.*

The prefix J in Definitions (25) and (26) stands for Jordan and it is used here to keep in mind that the properties and the structures defined here are all reminiscent of the Jordan product on the $W^\star$-algebra $\mathscr{A}$.

Remark 17. *Recall that an identifiable model is also locally identifiable (see Definition 21), this property is required here to ensure the non-degeneracy of the J-metric. The same construction can be followed also for a non-identifiable model, and this will still give rise to a non-negative, symmetric tensor. As we will see in Subsection 5.2.1, the tensor obtained, despite not being a metric, can still be interesting from an information theoretical perspective.*

Remark 18. *Consider the orbits of* $GL(A)$ *considered in Section 2.3 can be seen as parametric models, also, comparing Equation* (3.33) *with the expression of the tensor* $\mathcal{R}$ *in Equation* (2.104)*, it is easy to see that the construction proposed here reduces to the one presented in Section 2.3 in the finite-dimensional case. Moreover, the fact that we are able to define a metric as in Equation* (2.106) *implies that these models are actually J-smooth.*

We saw in this section how to define a metric on $\mathcal{M}$ that comes from the Jordan product of the algebra. One could ask why is this Riemannian structure relevant for Information Geometry. A first consideration that can be done is given just by the analogy to the finite-dimensional case treated in Section 2.3. There we saw how an analogous structure succeeds in reproducing metric tensors that are relevant in both Classical and Quantum Information Geometry, namely the Fisher-Rao metric, the Fubini-Study metric and the Bures-Helstrom metric.

In the next section, we will add another block to this picture, showing that the Riemannian structure introduced here has the property of being monotone with respect to maps that are dual to completely positive, unital maps. As we already stressed in Chapter 2, this kind of maps are the general $W^\star$-algebraic notion that allows to recover both Markov maps and CPTP maps, which are the relevant classes of maps considered respectively in Classical and Quantum Information Geometry. To some extent, the monotonicity property is reminiscent of the ***Data processing inequality*** [139, 160, 163].

In Chapter 4 we will see that this same construction can be used, to some extent, to obtain a result that is preliminary to obtaining two inequalities that are relevant in

Estimation Theory, namely the Cramér-Rao bound and the Helstrom bound. Finally, in Chapter 5 we will explicitly use the tools given in this Section for specific models and see that this construction will succeed in reconstructing some well-known metrics in Information Geometry.

3.3 J-smooth parametric models and maps; monotonicity and invariance

The aim of this Section is to prove two results. One is the monotonicity of J-metrics under norm-continuous, ultra-weakly continuous maps, that are the dual of a completely positive, unital map. The other is the invariance of these same metrics with respect to ultra-weakly continuous automorphisms of the algebra. This will be done in Subsection 3.3.1.

In order to prove these results, we need some additional properties about the behaviour of $W^\star$-algebras and of parametric models under mappings.

Proposition 27. *Let $\mathscr{A}$ and $\mathscr{B}$ be two $W^\star$-algebras, let $\Phi^*: \mathscr{B}^* \to \mathscr{A}^*$ be a bounded linear map, that is the dual of a completely positive map and let also $(\mathcal{M}, \mathrm{j}, \mathscr{B})$ be a smooth parametric model of n.p.l.f.s. Then $(\mathcal{M}, \mathrm{i}, \mathscr{A})$, with $\mathrm{i} = \Phi^* \circ \mathrm{j}$, is also a smooth parametric model of n.p.l.f.s.*

If Φ is also a unital map and $(\mathcal{M}, \mathrm{j}, \mathscr{B})$ is a smooth parametric model of states. Then $(\mathcal{M}, \mathrm{i}, \mathscr{A})$, with $\mathrm{i} = \Phi^* \circ \mathrm{j}$, is also a smooth parametric model of states.

Proof. The first result comes by considering the fact that completely positive maps send positive functionals in positive functionals and the fact that a bounded linear map is smooth and that the composition of smooth maps is a smooth map.

The second comes from the same considerations, taking into account the fact that the dual of a completely positive, unital map send states in states. □

Let $\mathscr{A}$ and $\mathscr{B}$ be $W^\star$-algebras, let also f be a map from $\mathscr{A}$ to $\mathscr{B}$, the ***predual map*** f_* of f is that map such that

$$f = (f_*)^*. \tag{3.34}$$

If such a map exists, we say that the map f admits a predual. It is easily seen that a map admits a predual if and only if it is ultra-weakly continuous.

Proposition 28. *Let $\mathscr{A}$ and $\mathscr{B}$ be $W^\star$-algebras and let also $f: \mathscr{A} \longrightarrow \mathscr{B}$ admit a predual as in Equation (3.34). The dual f^* of f sends normal functionals in $\mathscr{B}$ in normal functionals in $\mathscr{A}$.*

Proof. Let ρ be a normal functional in $\mathscr{B}^*$ and $a \in \mathscr{A}$ and consider the quantity

$$(f^*\rho)(a) = \rho(f(a)) = (f(a))(\hat{\rho}), \tag{3.35}$$

where we used the definition of dual map, the fact that ρ is a normal functional and denoted with $\widehat{\rho}$ the element in $\mathscr{B}_*$ s.t. $\mathrm{i}_{\mathscr{B}}(\widehat{\rho}) = \rho$, with $\mathrm{i}_{\mathscr{B}}$ being the natural immersion of $\mathscr{B}_*$ in $\mathscr{B}^*$. Then, taking into account that f admits a predual and using again the definition of dual map, we get

$$(f(a))(\widehat{\rho}) = ((f_*)^* a)(\widehat{\rho}) = a(f_*(\widehat{\rho})). \tag{3.36}$$

Thus we have that

$$(f^*\rho)(a) = a(f_*(\widehat{\rho})), \tag{3.37}$$

which implies immediately that $f^*\rho$ is a normal functional . □

Proposition 29. *Let $\Phi\colon \mathscr{A} \to \mathscr{B}$ be a norm-continuous, ultra-weakly continuous, completely positive map. Let $\Phi^*\colon \mathscr{B}^* \to \mathscr{A}^*$ the dual map of Φ, and let $\rho \in \mathscr{P}(\mathscr{A})$ be the image of $\omega \in \mathscr{P}(\mathscr{B})$via Φ^*. Φ^* defines a bounded linear map from $\mathcal{J}_\rho^*$ to $\mathcal{J}_\omega^*$.*

Proof. First of all, we note that dual map Φ^* sends positive linear functionals into positive linear functionals because Φ is positive. Moreover, since Φ is also ultra-weakly continuous, there is a bounded linear map $\varphi\colon \mathcal{B}(\mathscr{B}) \to \mathcal{B}(\mathscr{A})$ between the predual of $\mathscr{B}$ and the predual of $\mathscr{A}$ of which Φ is the dual map. Therefore, Φ^* sends normal linear functionals into normal linear functionals.

In general, if $\rho = \Phi^*(\omega)$, the map Φ does not send $\mathscr{A}_{sa}^\rho$ into $\mathscr{B}_{sa}^\omega$ (*cf.* proposition 24) . However, denoting with $\mathbf{P}_\omega$ the projection onto $\mathscr{B}_{sa}^\omega$ (which exists because $\mathscr{B}_{sa}^\omega$ is complemented in $\mathscr{B}_{sa}$ since ω is a *n.p.l.f.*), it is clear that the map

$$a \mapsto \Phi_\omega(a) := \mathbf{P}_\omega(\Phi(a)) \tag{3.38}$$

is a linear map between $\mathscr{A}_{sa}^\rho$ and $\mathscr{B}_{sa}^\omega$.

What is interesting is that, when looking at $\mathscr{A}_{sa}^\rho$ as a dense subspace of $\mathcal{J}_\rho^*$ and at $\mathscr{B}_{sa}^\omega$ as a dense subspace of $\mathcal{J}_\omega^*$ according to proposition 26, it defines a bounded linear map between $\mathcal{J}_\rho^*$ and $\mathcal{J}_\omega^*$. Indeed, for every $a \in \mathscr{A}_{sa}^\rho$, it holds

$$||\Phi_\omega(a)||_\omega^2 = G_\omega(\mathbf{P}_\omega(\Phi(a)), \mathbf{P}_\omega(\Phi(a))) = \omega(\mathbf{P}_\omega(\Phi(a))\,\mathbf{P}_\omega(\Phi(a))). \tag{3.39}$$

Let us denote by $\mathbf{Q}_\omega$ the complement projection of $\mathbf{P}_\omega$. Obviously, we have that

$$\mathbf{P}_\omega \mathbf{Q}_\omega = \mathbf{Q}_\omega \mathbf{P}_\omega = \mathbf{0} \tag{3.40}$$

and thus

$$\omega(ab) = \omega(\mathbf{P}_\omega(a)\mathbf{P}_\omega(b)) + \omega(\mathbf{Q}_\omega(a)\mathbf{Q}_\omega(b)). \tag{3.41}$$

Moreover, since $\mathbf{Q}_\omega(\mathscr{B}) = \mathscr{B}_{qq}$ is a $C^\star$-algebra of $\mathscr{B}$ and ω vanishes on $\mathscr{B}_{qq}$ (*cf.* discussion below equation (2.49)), we conclude that

$$\omega(ab) = \omega(\mathbf{P}_\omega(a)\mathbf{P}_\omega(b)), \tag{3.42}$$

so that

$$||\Phi_\omega(a)||_\omega^2 = \omega(\Phi(a)\Phi(a)). \tag{3.43}$$

Since Φ is completely positive, it follows that $\Phi(a)\Phi(a) \leq \Phi(a^2)$ (see Proposition 6) and thus

$$||\Phi_\omega(a)||_\omega^2 \leq \omega(\Phi(a^2)) = \rho(a^2) = G_\rho(a,a) = ||a||_\rho^2. \tag{3.44}$$

Equation (3.44) implies that Φ_ω can be extended to a bounded linear map between $\mathscr{J}_\rho^*$ and $\mathscr{J}_\omega^*$ as claimed. □

Corollary 2. *The linear map $\Phi_\omega : \mathscr{A}_{sa}^\rho \to \mathscr{B}_{sa}^\omega$ defined in (3.38) is a contraction.*

Proof. It is enough to inspect Equation (3.44), which also implies that $||\Phi_\omega|| \leq 1$. □

Corollary 3. *In the finite dimensional case, a map Φ^* as in Proposition 29 sends $J_{\Phi^*(\omega)}$ to J_ω.*

Proof. It is enough to consider that every finite-dimensional linear space is closed. □

3.3.1 Monotonicity and invariance properties of metrics on parametric models of *n.p.l.f.s*

Let us now prove a monotonicity property for G_ω under norm-continuous, ultra-weakly continuous, completely-positive maps that are the dual of some completely positive, unital map.

Proposition 30 (Monotonicity property of J-metrics)**.** *Let $\rho \in \mathscr{P}(\mathscr{A})$ be such that $\rho = \Phi^*(\omega)$ with $\omega \in \mathscr{P}(\mathscr{B})$ and let $\Phi^* \colon \mathscr{B}^* \to \mathscr{A}^*$ be the dual of a norm-continuous, ultra-weakly continuous map, completely positive, unital map Φ.*

For every $\eta \in \mathscr{J}_\omega$ it holds $\Phi^*(\eta) \in \mathscr{J}_\rho$ and

$$G_\rho(\Phi^*\eta, \Phi^*\eta) \leq G_\omega(\eta,\eta). \tag{3.45}$$

*Equation (3.45) will be referred to as the **monotonicity property** of G_ρ.*

Proof. Consider the map $\Phi_\omega^* \colon \mathscr{J}_\omega \to \mathscr{J}_\rho$ defined as in (3.38). Denoting with $(\,|\,)_\bullet$ the pairing between $\mathscr{J}_\bullet$ and its dual, it holds

$$(a|\Phi_\omega^*(\eta))_\rho = (\Phi_\omega(a)|\eta)_\rho = \eta(\Phi_\omega(a)) = \eta(\mathbf{P}_\omega(\Phi(a))) \tag{3.46}$$

for all $a \in \mathscr{A}_{sa}^\rho \subseteq \mathscr{J}_\rho^*$. According to proposition 26, every $\eta \in \mathscr{J}_\omega$ is also an element of AC_ω, and proposition 24 implies that

$$\eta(\Phi(a)) = \eta(\mathbf{P}_\omega(\Phi(a))) + \eta(\mathbf{Q}_\omega(\Phi(a))) = \eta(\mathbf{P}_\omega(\Phi(a))). \tag{3.47}$$

Accordingly, it follows that

$$(a|\Phi_\omega^*(\eta))_\rho = \eta(\Phi(a)) = (\Phi^*(\eta))(a) \tag{3.48}$$

which means that Φ_ω^* coincides with Φ^* on $\mathscr{J}_\omega \subseteq AC_\omega$. Eventually, recalling that Φ_ω is a contraction (see Corollary 2) and that the norm of Φ_ω^* coincides with that of Φ_ω, it immediately follows that

$$G_\rho(\Phi^*(\eta), \Phi^*(\eta)) = G_\rho(\Phi_\omega^*(\eta), \Phi_\omega^*(\eta)) \leq ||\Phi_\omega|| \; G_\omega(\eta, \eta) \leq G_\omega(\eta, \eta) \tag{3.49}$$

as claimed. □

As already stated in the Introduction of this work, this result is crucial if one wants to claim that the geometric picture developed in this work is somehow related to Information Theory. In fact, both in Classical and Quantum Information Geometry there exists a result that describes or characterizes the relevant metric tensors by looking at their behaviour under some class of maps. In Classical Information Geometry this is the invariance under congruent embeddings, that selects Fisher-Rao metric among all possible metrics on probability distributions, that is the content of Čencov's theorem [34]. In Quantum Information Geometry this is the characterization, due to Petz [149, 151], of the metrics that are monotone with respect to quantum channels.

While connecting this result with Petz characterization is immediate, it is not really clear to us at the moment if (or in which way) this is connected with Čencov theorem.

Proposition 31 (Invariance of G_ω under ultra-weakly continuous automorphisms)**.** *Let Φ be an ultra-weakly continuous automorphism of $\mathscr{A}$ and let also $\omega \in \mathscr{P}$, $a \in \mathscr{A}$, $\rho = \Phi^*(\omega)$ and η be a functional on $\mathscr{A}$ related to a as in Equation (3.9).*

Then we have

$$G_\rho(\Phi^*(\eta_a), \Phi^*(\eta_a)) = G_\omega(\eta_a, \eta_a), \tag{3.50}$$

meaning that the inner product G_ω with $\omega \in \mathscr{P}$ with respect to ultra-weakly continuous automorphisms.

Proof. Since Φ is an automorphism of $\mathscr{A}$, it is also an automorphism of the Jordan product.

Recall then that Φ_ω^* coincides with Φ^* on $\mathcal{J}_\omega \subseteq AC_\omega$ (*cf.* proposition 30), a direct computation shows that

$$\begin{aligned} \left(\Phi^*(\eta_a)\right)(b) &= \eta_a\left(\Phi(b)\right) = \omega\left(\{a, \Phi(b)\}\right) \\ &= \omega\left(\Phi\left(\Phi^{-1}\left(\{a, \Phi(b)\}\right)\right)\right) = \omega\left(\Phi\left(\{\Phi^{-1}\left(a\right), b\}\right)\right) \\ &= \rho\left(\{\Phi^{-1}(a), b\}\right) \end{aligned} \tag{3.51}$$

for all $a, b \in \mathscr{A}_{sa}$. It then follows that

$$G_\rho(\Phi^*(\eta_a), \Phi^*(\eta_a)) = \rho\left(\Phi^{-1}\left(a^2\right)\right) = \omega(a^2) = G_\omega(\eta_a, \eta_a), \tag{3.52}$$

as claimed. □

Remark 19. *When considering parametric models there is an additional aspect to be considered. Let $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ be a J-smooth, identifiable parametric models, $\omega = \mathrm{j}(m) \in \mathscr{P}$ and let Φ be an automorphism of $\mathscr{A}$. If $\Phi(\omega)$ is in the image of $\mathcal{M}$ via j, the* *result stated in Proposition 31 holds and also the property is translated in a property of the J-metric* $\mathcal{G}$*, meaning that*

$$\mathcal{G}_m(X(m), Y(m)) = \mathcal{G}_{m'}(X(m'), Y(m')). \tag{3.53}$$

Where $m' \in \mathcal{M}$ *is such that* $\mathrm{j}(m') = \Phi(\omega)$ *and* X *and* Y *are two vector fields on* $\mathcal{M}$*.*

What can also happen, however, is that $\Phi(\omega)$ *is not in the image of our model. In this case of course the result given in Proposition 31 can not really be applied.*

Remark 20. *In the case that the* $W^\star$*-algebra considered is* $\mathcal{B}(\mathcal{H})$*, this implies that the tensors we obtain are always invariant under the action* δ *of* $\mathcal{U}(\mathcal{H})$ *defined in Equation 2.55, since these are all automorphisms of the space of quantum states.*

Referring now to Remark 19, in order to be able to use Proposition 31 in a model on $\mathcal{B}(\mathcal{H})$ *what one can do is to consider models that have the property that if they contain a quantum state* ρ*, then they contain the whole orbit of the action* δ *passing through* ρ*.*

Chapter 4

Estimation theory on $W^\star$-algebras

In this Chapter, we will carry out a preliminary analysis regarding how the $W^\star$-algebraic framework introduced in the previous chapters can be applied to Estimation Theory. Before describing in detail the content of the Chapter, some basic concepts of Estimation Theory will be recalled, and in particular we will introduced the so-called Cramér-Rao bound.

A statistical model is a pair $(\Omega, \mathcal{P})$, where Ω is the sample space, and $\mathcal{P}$ is a set of probability distributions on Ω. A typical situation is the following, the given data, i.e. the product of some experiment, is a certain sample from the event space Ω. Then what we want to do is to choose, in the set $\mathcal{P}$, a probability distribution that will be our approximation of the sample given. In general, there is no single criterion or unique method to select a distribution that can be regarded as the *best approximation* of the sample given [62, Chapter 32].

Let us now assume that our space $\mathcal{P}$ is an open of $\mathbb{R}^n$, and let us consider a parametrization of $\mathcal{P}$ given by the vector of parameters $\Theta = \{\theta_1, \dots, \theta_N\}$. The idea is then to set a rule, which goes under the name of ***estimator***, that allows to make a prediction $\hat{\Theta}$, which will be called the ***estimate***, on the set of parameters Θ, which is the ***estimand*** of the estimation problem. In particular we will focus on ***point estimators***, meaning that the estimate $\hat{\Theta}$ is just one possible combination of values of Θ, i.e. the estimator S is a map from the sample space Ω to the set $\mathcal{P}$. A relevant property of estimators is if they converge in probability to its true value θ_0 [60] as the sample size grows. An estimator that has this property is called ***consistent***.

One usually also wants some notion of distance on the space $\mathcal{P}$, an important thing to bear in mind is that there is no unique choice for such a notion of distance and this choice depends heavily on the context in which the specific model is used. This notion of distance is usually given in terms of some ***cost function***, i.e. a two-point function $C(p_1, p_2)$ of the space $\mathcal{P}$ that is non-negative such that it is zero if and only if $p_1 = p_2$.

It is useful to see point estimators as ***random variables*** with values in $\mathbb{R}^n$. This will allow us to compute quantities such as the ***expected value***, the ***covariance matrix*** or the ***bias*** of an estimator.

Definition 27. *Let (Ω, Σ, p) be a probability space and $(E, \mathscr{E})$ a measurable space, a* *random variable X is a measurable function from Ω to E [80].*

The space E represents the set of possible values that the random variable can attain. The random variable X can be used to push-forward the probability measure p to E, thus giving a measure p_X on E, that is called the ***probability distribution*** of X. Let us focus on the case where $E = \mathbb{R}^n$ and let us equip it with the usual structure of measure space given by the Lebesgue measure μ_L, then, if the probability distribution p_X of X is absolutely continuous with respect to μ_L, we can define the ***density*** of the random variable X as the Radon-Nikodym derivative f_X of p_X with respect to μ_L [22].

The ***expected value*** $\mathbb{E}[X]$ of a random variable X is defined by

$$\mathbb{E}[X] := \int_{\Omega} X dp. \tag{4.1}$$

When we are dealing with more than one random variable, another relevant quantity, related with the concept of correlation, can be defined.

Definition 28. *Let $X = (X_1, \ldots, X_n)$ be a vector of random variables, the **covariance matrix** (or **dispersion matrix**, or **variance–covariance matrix**) K_X associated to X is the matrix whose entries are given by*

$$(cov(X))_{ij} = cov(X_i, X_j), \tag{4.2}$$

where with $cov(X_i, X_j)$ we denote the usual covariance of the two random variables X_i and X_j, i.e.

$$cov(X_i, X_j) = \mathbb{E}\left[X_i - \mathbb{E}[X_i]\right] \mathbb{E}\left[X_j - \mathbb{E}[X_j]\right]. \tag{4.3}$$

The inverse matrix of the covariance matrix is called the ***precision matrix*** or ***concetration matrix*** [70, 182].

If the space $\mathcal{P}$ is an open of $\mathbb{R}^n$, we can apply the definitions given here to point estimators, allowing us to compute their expectation values and variances and to define the ***bias*** of a point estimator as

$$B(S) = \mathbb{E}[S] - \theta_0. \tag{4.4}$$

Definition 29 (Unbiased estimator)**.** *An estimator is said to be **unbiased** if $B(S) = 0$, i.e. if its expectation value coincides with its true value.*

Wanting now to introduce the so-called Cramér-Rao inequality, we need the following definition.

Definition 30. *Let $\Theta = \{\theta_1, \ldots, \theta_N\}$ be the parameter vector specifying the probability distribution $p(x, \Theta)$ with $x \in \Omega$. The $N \times N$ matrix whose components are given by*

$$(I(\Theta))_{ij} = \mathbb{E}\left[\frac{\partial}{\partial \theta_i} \log p(x, \Theta) \frac{\partial}{\partial \theta_j} \log p(x, \Theta)\right] \tag{4.5}$$

*is called the **Fisher information matrix**.*

The Fisher information matrix plays a crucial role in Estimation Theory. As already stated in the Introduction of this work, it has been introduced by Ronald Fisher in his seminal work [78] and is then been used to define a Riemannian structure, the so called Fisher-Rao metric, on spaces of probability distributions. We will refer to it alternatively as Fisher information matrix and denote it by I or to Fisher-Rao metric and denote it by G_{FR}. As we will see in Section 5.1, this will be precisely the J-metric one obtains when the space $\mathcal{P}$ is the image of a J-smooth model on an Abelian $W^\star$-algebra.

As already stated in the Introduction of this work, the Cramér-Rao bound is a lower bound in the attainable precision of an unbiased estimator S [62, 156]. The estimator S can be regarded as a vector of random variables, so it makes sense to consider its covariance matrix $\mathrm{cov}\,(S)$, let the Fisher information matrix $I(\Theta)$ be defined and be invertible for all Θ and let also the following regularity condition hold:

$$\frac{\partial}{\partial\theta_j}\left[\int S(x)p(x,\Theta)d\mu\right] = \int S(x)\frac{\partial}{\partial\theta_j}[p(x,\Theta)]d\mu. \tag{4.6}$$

Given the stated conditions, we have that the inequality

$$\mathrm{cov}\,(S) \geq I(\Theta)^{-1} \tag{4.7}$$

holds. This inequality in is in the sense of matrix inequalities, meaning that the matrix $\mathrm{cov}\,(S) - I(\Theta)^{-1}$ is a semi-positive definite matrix.

Definition 31 (Efficient estimator)**.** *An estimator is said to be* ***efficient*** *if it achieves minimum variance.*

This means that for an unbiased and efficient estimator we have that (4.7) holds with the equality sign.

One estimation method whose use is widespread is the ***maximum likelihood method***, and it consists on maximizing the so called ***likelihood function***. The likelihood function is just the density of the probability distribution p_Θ, seen as a function of Θ instead of a function on the sample space. The value of the parameter Θ that attains the maximum of the likelihood function serves as a point estimate $\hat{\Theta}$ for Θ. The estimate $\hat{\Theta}$ obtained with this method of estimation can be proven to have the properties of being consistent, efficient, and its precision is given by the Fisher information matrix.

We are now ready to introduce with some more detail the content of the present Chapter.

- In Section 4.1, a notion of *parametric statistical models* suitable for our framework is introduce. In particular, we will see how the definition of parametric model of *n.p.l.f.s* introduced in Section 3.1 can be enriched with more structure in order to deal also with *parametric statistical models*. Let $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ be a smooth parametric model of *n.p.l.f.s*, the main idea here is that the role of the set $\mathcal{P}$ in the previous discussion is played by the manifold $\mathcal{M}$. In this same Section, the concept of *measurement procedure* will be discussed and some examples will be presented.

- We discussed briefly some properties of the so-called Cramér-Rao bound. In Section 4.2 the possibility to formulate this same bound in the $W^\star$-algebraic framework is discussed. The proof of the Cramér-Rao inequality in this context is carried out up to a certain point, obtaining what we call a *Pre-Cramér-Rao bound*. Then the technical difficulties on completing the proof to obtain the inequality are discussed. Finally, another statistical bound is introduced, namely the Helstrom bound [15, 103, 104, 105, 109, 110, 143]. This bound is related to Quantum Information Theory and represents some sort of universal bound for quantum system, independent of the classical measurement is performed on it. Also for the Helstrom bound, as for the Cramér-Rao bound, we will obtain a weaker result, that we will refer to as a Pre-Helstrom bound. As we will see, an essential ingredient for proving the Pre-Helstrom bound will be the monotonicity property (Proposition 30) of J-metrics proved in Section 3.3.

4.1 Parametric statistical models on $W^\star$-algebras

Before going to the definition of *parametric statistical model* in our framework, we will make two examples that show one of the perks of choosing to work in the general framework of $W^\star$-algebras, in fact we will see how this approach allows to compare structures arising from the classical world and the quantum world, this due to the fact that both are cast in the same language and can in this way live on the same spaces and be compared.

Example 8 (Stern-Gerlach experiment)**.** *The Stern-Gerlach experiment is an experiment, devised by Otto Stern in 1921 and conducted by him and Walther Gerlach in 1922 [90], whose relevance for the understanding of Quantum Mechanics is difficult to overestimate. In its simplest form, it shows the existence of an intrinsic angular momentum for electrons, and also clearly shows that such an intrinsic angular momentum is quantized, this is maybe the main evidence that brought physicists to introduce the concept of* ***spin****. Schematically, the apparatus consists of a source producing a beam of silver atoms, which then get deviated by passing through a non-homogeneous magnetic field. A screen collects the particles after the deviation and allows to measure which fraction of the beam is deflected upwards and which fraction is reflected downwards, see Figure 4.1.*

When several apparatuses of this kind are arranged sequentially and oriented in different directions, it captures the concept of incompatibility of measures on non-commuting observables. These two aforementioned features, namely quantization of outcomes of measurements and incompatibility of measures on non-commuting observables, are among the most striking in Quantum Mechanics, and maybe also the ones that challenge the most our intuition about the nature of these phenomena. Given its relevance for the development of basic concepts in Quantum Mechanics, the experiment is discussed, used, or at least mentioned in most textbooks in Quantum Mechanics and Quantum Information Theory. We refer the reader to [56, 139, 146],

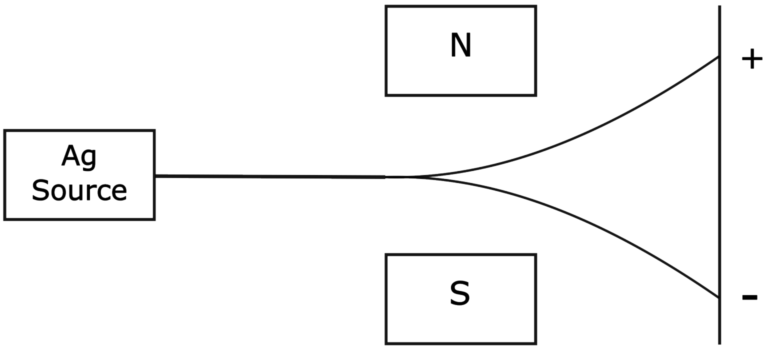


Figure 4.1: A pictorial representation of the Stern-Gerlach apparatus.

that introduce the experiment in three quite different ways, focusing on different aspects of it. Let us now turn to see how this is relevant for our discussion.

The beam of silver atoms is basically a realization of a ***qubit****, thus the appropriate way of describing the quantum state of the beam is as a state on the $W^\star$-algebra $\mathcal{B}(\mathbb{C}^2)$ of bounded linear operators on the Hilbert space $\mathbb{C}^2$ (see Example 3). However, the quantity that is actually measured in the experiment is just the fraction N_+ of the atoms in the beam that are deflected upwards and the fraction N_- of those that are deflected downwards, i.e. a probability distribution on the space of events $\chi = \{+, -\}$. If we imagine the direction of the magnetic field in the apparatus to be fixed, we can conceptualize the measurement as a map that to each possible state ρ of the atoms in the beam associates a probability vector $p = (p_+, p_-)$ with*

$$p_+ = \frac{N_+}{N_+ + N_-}; \quad p_- = \frac{N_-}{N_+ + N_-}. \tag{4.8}$$

Clearly, changing the orientation of the apparatus gives in general a different way of associating states of the qubit to probability vectors, meaning that selecting a way of doing measurements amounts to select a map from the space of states $\mathscr{S}_2$ of $\mathcal{B}(\mathbb{C}^2)$ to the space $\mathcal{P}(\chi)$ of probability vectors on χ. As pointed out in Example 2, $\mathcal{P}(\chi)$ is itself the space of states of the $W^\star$-algebra $L^\infty(\chi, \#)$, where $\#$ is the counting measure, thus we implement the measurement procedure as a map $\mathfrak{m}$ such that its dual $\mathfrak{m}^*$ is a completely positive and trace-preserving map from $\mathcal{B}(\mathbb{C}^2)^*$ to $L^\infty(\chi, \#)^*$. The notation choice for the map may seem odd at the moment, but it will be clarified in what follows.

Let now $(\mathcal{M}, \mathrm{j}, \mathcal{B}(\mathbb{C}^2))$ be a parametric model of positive linear functionals, we have that the commutative diagram

$$\begin{array}{ccc} \mathscr{S}_2 & \xrightarrow{\mathfrak{m}^*} & \mathcal{P}(\chi) \\ \uparrow{\scriptstyle \mathrm{j}} & \nearrow{\scriptstyle \mathrm{j}^c} & \\ \mathcal{M} & & \end{array}$$

implicitly defines j^c *and gives rise to another parametric model* $(\mathcal{M}, \mathrm{j}^c, L^\infty(\chi, \#))$. *This is what in our framework will be called the parametric statistical model associated to the parametric model* $(\mathcal{M}, \mathrm{j}, \mathcal{B}(\mathbb{C}^2))$ *via the measurement procedure* $\mathfrak{m}$*. In this case, because of Proposition 27, if* $(\mathcal{M}, \mathrm{j}, \mathcal{B}(\mathbb{C}^2))$ *is J-smooth then also* $(\mathcal{M}, \mathrm{j}^c, L^\infty(\chi, \#))$ *will be.*

This allows us to define two Riemannian tensors on $\mathcal{M}$*, one being the J-metric* G *on* $\mathcal{M}$ *associated to the model* $(\mathcal{M}, \mathrm{j}, \mathcal{B}(\mathbb{C}^2))$ *and the other being the J-metric* G^c *associated to the model* $(\mathcal{M}, \mathrm{j}^c, L^\infty(\chi, \#))$*. The reason for the subscript c on the second metric stands for classical, in fact this metric is related to the structure of the space* $L^\infty(\chi, \#)$ *and is independent on the fact that* $\mathcal{B}(\mathbb{C}^2)$ *is non-Abelian.*

Example 9 (Double-slit experiment)**.** *The double-slit experiment was crucial in the process of corroborating fundamental ideas of Quantum Mechanics, it was devised and performed by Clinton Davisson and Lester Germer in the years 1923-27 [66, 67, 68], here we will present a simplified version of the experiment, to avoid difficulties tied to the experimental setting that would not influence the conceptual idea of our description.*

The setting is the following: a source emits electrons towards a barrier presenting two slits, the pattern produced by the interaction of the electrons with the barrier is then observed on a detector screen. The idea behind it is to investigate the wave-particle duality conjectured by Louis de Broglie in his Ph.D. thesis [69]; a particle-like object would present a pattern that has only two maxima, given by the objects that go through one of the slit or the other, while a wave-like object would give the diffraction-interference pattern given by the solution of some kind of wave equation.

According to standard quantum mechanics, the beam can be described by elements of $\mathcal{B}(\mathcal{H})$ *for some Hilbert space* $\mathcal{H}$*, we will not focus on the description of the quantum system, since it is not the aim of this example. What we want to stress is that the outcome of the measurement is a probability distribution on the detector screen, i.e. a state of the algebra* $L^\infty(\mathbb{R}^2, \mu_L)$*, where* μ_L *is the Lebesgue measure. Thus we define the measurement procedure* $\mathfrak{m}$ *as a ultra-weakly continuous, completely positive and unital map from* $L^\infty(\mathbb{R}^2, \mu_L)$ *to* $\mathcal{B}(\mathcal{H})$*, so that its dual* $\mathfrak{m}^*$ *will send n.p.l.f.s on* $\mathcal{B}(\mathcal{H})$ *to n.p.l.f.s of* $L^\infty(R^2, \mu_L)$*.*

Let now $\mathcal{M}$ *be a Banach manifold and* $(\mathcal{M}, \mathrm{j}, \mathcal{B}(\mathcal{H}))$ *be a parametric model of n.p.l.f.s, also let us define* j^c *denote with* $\mathrm{ba}(\mathbb{R}^2, \mu_L)$ *the space of finitely additive, finite signed measures on* $\mathbb{R}^2$ *that are absolutely continuous with respect to* μ_L*, which is the dual of* $L^\infty(\mathbb{R}^2, \mu_L)$*. We then have that the diagram*

$$\begin{array}{ccc} \mathscr{S}(\mathcal{H}) & \xrightarrow{\mathfrak{m}^*} & \mathrm{ba}(R^2, \mu_L) \\ {\scriptstyle \mathrm{j}}\uparrow & \nearrow_{\mathrm{j}^c} & \\ \mathcal{M} & & \end{array}$$

commutes and gives rise to the parametric model $(\mathcal{M}, \mathrm{j}^c, L^\infty(\mathbb{R}^2, \mu_L))$*. This is what in our framework will be called the parametric statistical model associated to the parametric model* $(\mathcal{M}, \mathrm{j}, \mathcal{B}(\mathcal{H}))$ *via the measurement procedure* $\mathfrak{m}$*.*

Another possible way of treating this same experiment is by taking into account the fact that the number of electrons in the beam is finite, the detector screen is a finite region, and the precision of the position measurements on the detector screen can not be infinite. This would demote $L^\infty(\mathbb{R}^2, \mu_L)$ *to a finite-dimensional* $W^\star$*-algebra, since the outcome of the measurement can be given in terms of an histogram. However, treating the infinite-dimensional case allows to have a description that on one hand does not depend on the specific accuracy of the apparatus considered. Moreover, the transition to continuous distribution to histograms can always be performed in a way that is analogous to what will be presented in Remark 22.*

In our framework, we have a J-regular parametric model $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ of normal states on the $W^\star$-algebra $\mathscr{A}$. And the object on which the measurement is performed is given by an element of $\mathscr{S}(\mathscr{A})$, the outcome of the measurements is still a probability distribution, but we have no reason at this level to restrict ourselves in the choice of the measure space (Ω, Σ, μ) that gives the possible outcomes of the measurement. For simplicity we will refer to the measure space (Ω, Σ, μ) only as Ω. Notice that the space $\mathcal{P}(\Omega, \mu)$ of probability distributions on Ω that are absolutely continuous with respect to μ is itself the space of states of the $W^\star$-algebra $L^\infty(\Omega, \mu)$ of μ-essentially bounded functions on the space Ω, whose dual is given by the space $S(\Omega, \mu)$ of signed measures on Ω that are absolutely continuous with respect to μ. Again, for simplicity we refer to the $W^\star$-algebra $L^\infty(\Omega, \mu)$ only as $L^\infty(\Omega)$.

Definition 32 (Measurement procedure)**.** *A ultraweakly continuous, positive, unital map* $\mathfrak{m}$ *from* $L^\infty(\Omega)$ *to* $\mathscr{A}$ *is called a* ***measurement procedure*** *for the parametric model* $(\mathcal{M}, \mathrm{j}, \mathscr{A})$.

Remark 21. *Conceptually, one can say that this map creates a correspondence between classical observables defined on* Ω *and the (possibly quantum) observables in the* $W^\star$*-algebra* $\mathscr{A}$*. From a more technical point of view, with this definition the dual map* $\mathfrak{m}^*$ *of the measurement procedure* $\mathfrak{m}$ *associates probability distributions on* Ω *to states of the* $W^\star$*-algebra* $\mathscr{A}$*.*

In order to see this, first recall here that the concept of complete positivity and positivity of a map coincide whenever either the source or the target of the map is an Abelian algebra (see Proposition 5). Then, since the dual of a positive map is positive (see Proposition 4), we have that $\mathfrak{m}^*$ *is positive, and for the same reason as before, also completely positive. Again because of Proposition 4, we can then also conclude that* $\mathfrak{m}^*$ *also preserves the norm of positive functionals.*

As stated in Section 3.3, requiring that $\mathfrak{m}$ is weak*-continuous is equivalent to require that it admits a predual map, meaning that

$$\exists\, \mathfrak{m}_* : \mathscr{A}_* \longrightarrow L^1(\Omega) \quad \text{s.t.} \quad \mathfrak{m} = (\mathfrak{m}_*)^*. \tag{4.9}$$

Taking into account Proposition 28, this implies that the dual $\mathfrak{m}^*$ of the measurement procedure $\mathfrak{m}$ sends normal functionals in normal functionals, i.e. its image is made of probability distributions that can be represented in terms of a function in $L^1(\Omega)$.

The dual $\mathfrak{m}^*$ of measurement procedure $\mathfrak{m}^*$, together with the parametric model $(\mathcal{M}, \mathrm{j}, \mathscr{A})$, give rise to the commutative diagram

$$\begin{array}{ccc} \mathscr{S}(\mathscr{A}) & \xrightarrow{\mathfrak{m}^*} & \mathcal{P}(\Omega) \\ \uparrow{\scriptstyle \mathrm{j}} & \nearrow_{\mathrm{j}^c} & \\ \mathcal{M} & & \end{array}$$

where clearly $\mathrm{j}^c := \mathfrak{m}^* \circ \mathrm{j}$. This map associates to each point on our parameter manifold $\mathcal{M}$ a probability distribution, morally an integrable function, on Ω. The superscript c stands for classical, since the $W^\star$-algebra $\mathscr{A}$ is possibly non-commutative, hence associated to a quantum system, while the image of j^c will always be a probability distribution. In the following, we will use the shortcut

$$p_m = \mathrm{j}^c(m). \tag{4.10}$$

What we will do now is to enrich the notion of statistical model given in the introduction to this section. There, the main role was played by the set $\mathcal{P}$ of probability distribution chosen to model the sampling. Here we have a richer structure given by the manifold of parameters, the $W^\star$-algebra chosen to describe the setting, allowing us to apply the ideas also to Quantum Mechanics, and finally the probability distributions given by sampling. Even if it may look quite different from the definition given in the introduction to this Section, one can claim that the spirit is quite the same.

Definition 33 (Statistical parametric model)**.** *Let $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ be a parametric model of normal states and $\mathfrak{m}^*$ be a measurement procedure for $(\mathcal{M}, \mathrm{j}, \mathscr{A})$. The* parametric statistical model *associated to the parametric model of normal states $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ via the measurement procedure $\mathfrak{m}^*$ is given by the triple $(\mathcal{M}, \mathrm{j}^c, L^\infty(\Omega))$, where $\mathrm{j}^c = \mathfrak{m}^* \circ \mathrm{j}$.*

Notice that $(\mathcal{M}, \mathrm{j}^c, L^\infty(\Omega))$ defines a smooth parametric model, because of Proposition 27.

Remark 22. *In the particular case in which the algebra $\mathscr{A}$ is an algebra of functions on some measure space, to each parametric model $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ there is a parametric statistical model that can canonically be constructed using the identity as measurement procedure. Let us make the example $\mathscr{A} = L^\infty(\Omega)$, then we have that the diagram*

$$\begin{array}{ccc} \mathscr{S}(\mathscr{A}) \equiv \mathcal{P}(\Omega) & \xrightarrow{\mathbb{I}} & \mathcal{P}(\Omega) \\ \uparrow{\scriptstyle \mathrm{j}} & \nearrow_{\mathrm{j}} & \\ \mathcal{M} & & \end{array}$$

commutes trivially.

Let us stress, however, that this is not the only parametric statistical model that can be constructed, and that one can choose probability distributions on a different space of events as target space for the measurement procedure. Let us make a small example to make this clear; let us imagine that we have an object that can lie somewhere on a line, the appropriate framework for this situation would be to take $\mathscr{A} = L^\infty(\mathbb{R}, \mu_L)$ *with* μ_L *being the Lebesgue measure on* $\mathbb{R}$*. The space of states for this situation is given by probability distributions on* $\mathbb{R}$*. Imagine then that we only want to know whether the object lies on a region* I *of* $\mathbb{R}$ *or not. In this case the measurement procedure* $\mathfrak{m}^*$ *will be a map that takes a probability distribution* p *on* $\mathbb{R}$ *and gives the two component probability vector*

$$\vec{p} := \mathfrak{m}^*(p) = \left(\int_I p d\mu_L, \int_{I/\mathbb{R}} p d\mu_L \right), \tag{4.11}$$

in particular, in this case, the target space of the statistical parametric model brings us from a infinite-dimensional setting to a finite-dimensional one.

In Chapter 3, we saw that we can endow $\mathcal{M}$ with a Riemannian structure that comes from the Jordan product on the $W^\star$-algebra $\mathscr{A}$ as

$$\mathcal{G}_m(X(m), Y(m)) := G_{\mathrm{j}(m)} \left(T_m \mathrm{j}(X(m)), T_m \mathrm{j}(Y(m)) \right), \tag{4.12}$$

where $m \in \mathcal{M}$, X and Y are smooth vector fields on $\mathcal{M}$, and provided that the model $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ is J-regular.

Here, with the addition of the space $\mathcal{P}(\Omega)$ in the picture, we have another possibility to endow $\mathcal{M}$ with a Riemannian structure $\mathcal{G}^c_m$ coming from the $W^\star$-algebra $L^\infty(\Omega)$. Again, the superscript c stands for classical, the tensor G^c can not be influenced by the possible non-commutativity of $\mathscr{A}$ and depends only on the classical structure of $\mathcal{P}(\Omega)$. In Section 5.1, we will see that every time a model has as target space a space of probability distributions, what we refer to as a ***classical model***, it is possible to write the J-metric of the model (see Equation 3.33) in the form

$$\mathcal{G}^c_m(X(m), Y(m)) = \int_\Omega \left[T_m \mathrm{j}^c(X(m)) \log p_m \right] \left[T_m \mathrm{j}^c(Y(m)) \log p_m \right] dp_m, \tag{4.13}$$

which is precisely the expression of the Fisher-Rao metric given in [13]. In our case the measurement procedure allows us to obtain a classical model from the (possibly non-classical) model $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ so that the object defined in (4.13) is always defined.

4.1.1 Some notions of Estimation Theory on $W^\star$-algebras

In our framework, we will say that an estimator is a measurable map $S : \Omega \to \mathcal{M}$, this is clearly inspired by the definition given in the introduction to this Section, there we had a map from Ω to $\mathcal{P}$, thus selecting a probability distribution for every event on the sample space.

In our framework, the probability distributions are introduced to make a connection with the sampling related to the experiment, but we also have a possibly non-commutative structure lying behind it and. Thus the definition of estimator gives rise to the following picture.

$$\begin{array}{ccc} \mathscr{S}(\mathscr{A}) & \xrightarrow{\mathfrak{m}^*} & \mathcal{P}(\Omega) \\ {\scriptstyle j}\big\uparrow & & \\ \mathcal{M} & \xleftarrow{\;S\;} & \Omega \end{array}$$

With the same spirit, we define the cost function on the parameter manifold $\mathcal{M}$.

Definition 34. *A* cost function C *is a function C from $\mathcal{M} \times \mathcal{M}$ that is smooth, non-negative, and such that*

$$C(m_1, m_2) = 0 \iff m_1 = m_2. \tag{4.14}$$

Let us define the function $C_m : \mathcal{M} \to \mathbb{R}$ such that $C_m(\overline{m}) = C(m, \overline{m})$. So that $C_m \circ S$ is a function from Ω to $\mathbb{R}$, let us also denote it as f_m. We are interested in this quantity because it represents a distance between the estimate associated to the point ω and a point m on the parameter manifold.

Then the likelihood function is just a function that, given two points m_1 and m_2 of $\mathcal{M}$, gives the expected value of the function f_{m_1} with respect to the probability distribution p_{m_2} on Ω associated to $j(m_2)$ as in (4.10). In order for the definition of the likelihood function to be well-posed, the function f_m should be in $L^\infty(\Omega)$. Clearly, this will depend on the choice of the cost function C and the estimator S, and it can not be assured in general. From now on, let us assume that C and S are such that $f_m \in L^\infty(\Omega)$. We are now ready to define the likelihood function as

$$L : \mathcal{M} \times \mathcal{M} \longrightarrow \mathbb{R}, \quad L(m_1, m_2) = p_{m_2}(f_{m_1}). \tag{4.15}$$

Notice that $p_{\overline{m}}$ is absolutely continuous with respect to μ, this comes from the fact that $\mathfrak{m}^*$ admits a predual (see Proposition 28 and Example 5). Thus integration with respect to the measure $p_{\overline{m}}$ can be written as

$$L(m_1, m_2) = \int_\Omega f_{m_1} p_{m_2} d\mu, \tag{4.16}$$

where, with an abuse of notation, we denoted with $p_{\overline{m}}$ also the Radon-Nikodym derivative of the probability distribution $p_{\overline{m}}$ with respect to μ.

Let us also define the function $L_m : \mathcal{M} \to \mathbb{R}$, that is just the likelihood function, but with the point on m that defines the probability distribution fixed, i.e.

$$L_m(\overline{m}) = L(\overline{m}, m). \tag{4.17}$$

Definition 35 (Stationary and unbiased estimator)**.** *An estimator is said to be* stationary *if $L_{\overline{m}}$ is stationary at $\overline{m}$ for all $\overline{m} \in \mathcal{M}$. An estimator is said to be* unbiased *if $L_{\overline{m}}$ has a minimum at $\overline{m}$ for all $\overline{m} \in \mathcal{M}$.*

Let now $V \in T_{\overline{m}}\mathcal{M}$, the stationary condition given in the last definition can be written as

$$(VL_{\overline{m}})(\overline{m}) = 0 \tag{4.18}$$

for all $\overline{m} \in \mathcal{M}$. This definition gives back Definition 29 whenever $\mathcal{M}$ is chosen to be an open of $\mathbb{R}^n$ and the cost function is chosen to be the Euclidean distance, this computation is explicitly performed in [106].

4.2 Statistical bounds

In the introduction of this Chapter we glanced at the role of the Cramér-Rao bound (see Equation 4.7) in the contest of Estimation Theory and maximum likelihood estimation in particular. We will now tackle the problem of obtaining the Cramér-Rao bound in the $W^\star$-algebraic framework. After that we will move to discussing another statistical bound, the Helstrom bound.

4.2.1 Cramér-Rao bound

Let us now turn to our discussion of the Cramér-Rao bound in the $W^\star$-algebraic framework. An essential ingredient for the derivation of this bound is the Hessian form H of the likelihood function, defined as

$$H_{\overline{m}}(X(\overline{m}), Y(\overline{m})) := (YXL_{\overline{m}})(\overline{m}), \tag{4.19}$$

with X and Y vector fields on $\mathcal{M}$. In order to manipulate this definition and obtain an expression of this related with the Fisher-Rao metric on $\mathcal{P}(\Omega)$, we need to introduce some mathematical machinery.

Let us define the following maps, called respectively the diagonal immersion, the left projection and the right projection,

$$\begin{aligned} \pi_l &: \mathcal{M} \times \mathcal{M} \longrightarrow \mathcal{M}, \quad \pi_l(m_1, m_2) := m_1, \\ \pi_r &: \mathcal{M} \times \mathcal{M} \longrightarrow \mathcal{M}, \quad \pi_r(m_1, m_2) := m_2, \\ i_D &: \mathcal{M} \longrightarrow \mathcal{M} \times \mathcal{M}, \quad i_D(m) := (m, m). \end{aligned} \tag{4.20}$$

This allows us to define, for every vector field X on $\mathcal{M}$, two vector fields on $\mathcal{M} \times \mathcal{M}$.

Definition 36 (Left and right lift of a vector field)**.** *Let X be a vector field on M, its left and right lift are two vector fields X_l and X_r on $M \times M$ such that*

$$\begin{aligned} X_l(\pi_l^* f) &= \pi_l^*(Xf), \\ X_r(\pi_r^* f) &= \pi_r^*(Xf), \end{aligned} \tag{4.21}$$

for all smooth functions f on M.

From a more concrete point of view, the vector fields X_l and X_r have the purpose of taking derivative of a function of $\mathcal{M} \times \mathcal{M}$ only with respect to one of its argument. To see this, let $\mathcal{M}$ be a finite-dimensional manifold, and let $\{x^j\}$ and $\{y^j\}$ be two (possibly identical) coordinate systems on $\mathcal{M}$, then $\{x^j, y^j\}$ is a coordinate system on $\mathcal{M} \times \mathcal{M}$ and we have

$$\begin{aligned} X_l(\pi_l^* f) &= \pi_l^*(Xf) = \pi_l^* \left(X^j(x) \frac{\partial}{\partial x^j} f(x) \right), \\ X_r(\pi_r^* f) &= \pi_r^*(Xf) = \pi_r^* \left(X^j(y) \frac{\partial}{\partial y^j} f(y) \right). \end{aligned} \tag{4.22}$$

This, together with the definition on $L_{\overline{m}}$, implies that it is possible to rewrite the stationary condition for the estimator S in (4.18) as

$$i_D^*(X_l L) = 0, \tag{4.23}$$

and the expression of the Hessian form in (4.19) as

$$H_{\overline{m}}(X(\overline{m}), Y(\overline{m})) = \left(i_D^*(Y_l X_l L)\right)(\overline{m}). \tag{4.24}$$

The vector fields X, X_l and X_r and the map i_D are related by the following proposition.

Proposition 32. *The vector field X is i_D-related to $X_l + X_r$, i.e.*

$$X(i_D^* f) = i_D^*[(X_l + X_r)f] \tag{4.25}$$

for all smooth functions f on $\mathcal{M} \times \mathcal{M}$.

Proof. In [46], the authors prove this statement in the finite-dimensional setting. However, an analogous proof can be carried out in the infinite dimensional case. □

Finally, for technical reasons, let us define the functions from F_{X_m} and G_{Y_m} from Ω to $\mathbb{R}$ and the bilinear form C as

$$\begin{aligned} F_{X_m} &:= X f_m, \\ G_{Y_m} &:= Y \log p_m, \\ C(X_m, Y_m) &:= \mathbb{E}_m \left[F_X F_X \right]. \end{aligned} \tag{4.26}$$

Where X and Y are vector fields on $\mathcal{M}$, X_m and Y_m are their values at m and $\mathbb{E}_m$ denotes the expectation value with respect to the probability distribution p_m.

We are now ready to prove an inequality, that we will refer to as a ***pre-Cramér-Rao bound***, since it can be seen as a preliminary step towards proving the actual bound in this framework, a more detailed discussion will follow the proof of the inequality.

Proposition 33 (Pre-Cramér-Rao bound). *Let $\overline{m} \in \mathcal{M}$, X and Y be vector fields on $\mathcal{M}$, S be a stationary estimator, the Hessian form $H_{\overline{m}}$ defined as in Equation (4.19), the form C defined as in Equation (4.26) and the metric $\mathcal{G}^c_m$ defined as in Equation (4.13). Let also the functions $F_{X_{\overline{m}}}$ and $G_{Y_{\overline{m}}}$ defined as in Equation (4.26) be in $L^\infty(\Omega)$.*

The inequality

$$(H_{\overline{m}}(X(\overline{m}), Y(\overline{m})))^2 \leq \mathcal{C}(X(\overline{m}), X(\overline{m}))\mathcal{G}^c_m(Y(\overline{m}), Y(\overline{m})) \tag{4.27}$$

holds.

Proof. Since S is a stationary estimator, $i^*_D(X_l L)$ is the null function on $\mathcal{M}$, this implies that also its derivative with respect to any vector field will be zero, i.e.

$$Y\left(i^*_D(X_l L)\right) = 0. \tag{4.28}$$

Then we can use Proposition 32 to obtain

$$0 = Y\left(i^*_D(X_l L)\right) = i^*_D\left(Y_l X_l L + Y_r X_l L\right), \tag{4.29}$$

that in turns implies

$$i^*_D\left(Y_l X_l L\right) = -i^*_D\left(Y_r X_l L\right), \tag{4.30}$$

thus getting to the expression for the Hessian form

$$H_{\overline{m}}(X(\overline{m}), Y(\overline{m})) = -\left(i^*_D(Y_r X_l L)\right)(\overline{m}) = -(Y p_{\overline{m}})(X f_{\overline{m}}). \tag{4.31}$$

Recalling (4.16), we have

$$H_{\overline{m}}(X(\overline{m}), Y(\overline{m})) = -\int_\Omega (X f_{\overline{m}})(Y p_{\overline{m}}) d\mu = -\int_\Omega (X f_{\overline{m}})(Y \log p_{\overline{m}}) p_{\overline{m}} d\mu. \tag{4.32}$$

This last expression can be seen as the expectation value of the product of the functions $F_{X_{\overline{m}}}$ and $G_{Y_{\overline{m}}}$ defined as in Equation (4.26).

We then have

$$H_{\overline{m}}(X(\overline{m}), Y(\overline{m})) = -\mathbb{E}_{\overline{m}}[F_X G_Y], \tag{4.33}$$

where $\mathbb{E}_{\overline{m}}[f]$ denotes the expectation of the function f with respect to the probability distribution $p_{\overline{m}}$. Notice that the expression $\mathbb{E}_{\overline{m}}[F_X G_Y]$ can be seen as an inner product on the probability space $(\Omega, p_{\overline{m}})$, so that we can use the Cauchy-Schwarz inequality on (4.33) to obtain

$$(H_{\overline{m}}(X(\overline{m}), Y(\overline{m})))^2 \leq \mathbb{E}_{\overline{m}}[F_X F_X]\mathbb{E}_{\overline{m}}[G_Y G_Y]. \tag{4.34}$$

Taking a closer look at the quantity $\mathbb{E}_{\overline{m}}[G_Y G_Y]$, we can see that

$$\mathbb{E}_{\overline{m}}[G_Y G_Y] = p_{\overline{m}}(Y \log p_{\overline{m}} Y \log p_{\overline{m}}) = p_{\overline{m}}(\{Y \log p_{\overline{m}}, Y \log p_{\overline{m}}\}), \tag{4.35}$$

where the curly brackets here denote the Jordan product on $L^\infty(\Omega)$, and the last equality comes trivially from the fact that $L^\infty(\Omega)$ is Abelian. Comparing the last expression with Equation (4.13) we can write

$$\mathbb{E}_{\overline{m}}[G_Y G_Y] = \mathcal{G}^c_m(\widetilde{Y}, \widetilde{Y}). \tag{4.36}$$

The vectors $\widetilde{Y}$ in the last expression are those vectors in $T_{\overline{m}}\mathcal{M}$ such that $T\mathrm{j}(\widetilde{Y}) \in T_{\mathrm{j}(\overline{m})}\mathcal{P}(\Omega)$ is that element of the tangent double cone at $\mathrm{j}(\overline{m})$ that is associated to the element $Y \log p_{\overline{m}}$ of L^∞ as in (3.9).

Recalling now the definition of the bilinear form C in Equation (4.26), we finally have that Equation (4.34) can be rewritten as

$$(H_{\overline{m}}(X(\overline{m}), Y(\overline{m})))^2 \leq \mathcal{C}(X(\overline{m}), X(\overline{m}))\mathcal{G}^c_m(Y(\overline{m}), Y(\overline{m})). \tag{4.37}$$

□

Let us now clarify in which sense we call this inequality Pre-Cramér-Rao bound, in fact, one could argue that this inequality has little in common with the Cramér-Rao bound in Equation (4.7).

Let us assume that the Hessian form is invertible, then we can define a map $\widetilde{H} : T_{\overline{m}}\mathcal{M} \longrightarrow T^*_{\overline{m}}\mathcal{M}$ as

$$\left(\widetilde{H}_{\overline{m}}(X(\overline{m}))\right)(Y(\overline{m})) = H_{\overline{m}}(X(\overline{m}), Y(\overline{m})). \tag{4.38}$$

The key passage to go from the Pre-Cramer-Rao bound we stated and the Cramér-Rao bound is given by a process of maximizing the value of the factor

$$(H_{\overline{m}}(X(\overline{m}), Y(\overline{m})))^2 \tag{4.39}$$

over the assumption that $Y(m)$ has Fisher-Rao norm one, and showing that

$$\max_{\|Y(\overline{m})\|_{FR}=1} (H_{\overline{m}}(X(\overline{m}), Y(\overline{m})))^2 = I^{-1}(\widetilde{H}_{\overline{m}}(X(\overline{m})), \widetilde{H}_{\overline{m}}(X(\overline{m}))), \tag{4.40}$$

i.e. that the maximum attained by this function is proportional to the inverse of the Fisher Information matrix.

We can see that this maximization procedure allows on one hand to introduce the inverse of the Fisher Information matrix and at the same time of getting rid of the factor $\mathcal{G}^c_m(Y(\overline{m}), Y(\overline{m}))$, since clearly now it has value 1. However, the proof in the finite-dimensional case (see for example [106] or [52]) relies on the fact that the maximization is done on the Fisher-Rao unit sphere, hence a compact set. As it is well-known, this is not true in the infinite-dimensional case, representing a big obstacle in completing the proof and getting to the Cramér-Rao inequality in our framework. What looks like is that, in this framework, a case-by-case analysis is needed, and whether the maximization procedure can be successfully accomplished may depend on the specific model, or the specific estimator one is considering.

Then, following [106], one can define the covariance bivector of the estimator S as

$$\mathrm{cov}\,(S)(\xi_1,\xi_2) := C\left(\widetilde{H}_{\overline{m}}^{-1}(\xi_1), \widetilde{H}_{\overline{m}}^{-1}(\xi_2)\right), \tag{4.41}$$

for every ξ_1 and ξ_2 in $T^*_{\overline{m}}\mathcal{M}$.

Combinining Equations (4.40) and (4.41) with the Pre-Cramér-Rao bound in Proposition 33 we get that

$$\mathrm{cov}\,(S)(\xi,\xi) \geq I^{-1}(\xi,\xi) \tag{4.42}$$

for all $\xi \in T^*_m\mathcal{M}$, i.e. the covariance bivector of the estimator S is bounded from below by the inverse of the Fisher Information matrix.

It can then be easily shown that if the estimator takes value on an open set of $\mathbb{R}^n$ and the cost function is chosen to be the Euclidean distance, this expression coincides precisely with Equation (4.7) [106].

4.2.2 Helstrom bound

In Quantum Detection Theory and Quantum Estimation Theory, one usually tries to replicate results of Classical Statistical Decision Theory or Estimation Theory with density operators replacing probability distributions. This brings to formulate a bound on accuracy of estimators that goes under the name of ***Helstrom bound*** and that can be seen as the quantum counterpart of the Cramér-Rao bound [15, 103, 104, 105, 109, 110, 143]

Let us recall Equation (4.13), right after this equation, we mentioned that the structure introduced there can not feel the possible non-commutativity of the $W^\star$-algebra $\mathscr{A}$. Now we will formulate a bound that is actually responsive towards the non-commutative structure of $\mathscr{A}$, and in order to do so, we will have to take in the picture the tensor $\mathcal{G}$ given in Equation (4.13).

Proposition 34. *Let $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ be a parametric model of states on $\mathscr{A}$, $\mathfrak{m}$ a measurement procedure, $\mathrm{j}^c = \mathfrak{m}^* \circ \mathrm{j}$ and $(\mathcal{M}, \mathrm{j}^c, \mathcal{P}(\Omega))$ the statistical parametric model associated to $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ via the measurement procedure $\mathfrak{m}$. Let also $\mathcal{G}$ be the J-metric (see Definition 26) of the parametric model $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ and $\mathcal{G}^c$ be as in (4.13). Then we have*

$$\mathcal{G}_m(X,X) \geq \mathcal{G}^c_m(X,X) \tag{4.43}$$

for all $m \in \mathcal{M}$ and for all $X \in T_m\mathcal{M}$.

Proof. Let X and Y be two vector fields on $\mathcal{M}$ and let us recall that

$$\mathcal{G}_m(X(m),Y(m)) := G_{\mathrm{j}(m)}\left(T_m\mathrm{j}(X(m)), T_m\mathrm{j}(Y(m))\right), \tag{4.44}$$

while

$$\mathcal{G}^c_m(X(m),Y(m)) := G^{FR}_{\mathrm{j}^c(m)}\left(T_m\mathrm{j}(X(m)), T_m\mathrm{j}(Y(m))\right). \tag{4.45}$$

The map $\mathfrak{m}^*$ is a ultra-weakly continuous, CPTP map such that

$$\begin{aligned} \mathrm{j}^c(m) &= \mathfrak{m}^*(\mathrm{j}(m)), \\ T_m\mathrm{j}^c(X(m)) &= \mathfrak{m}^*(T_m\mathrm{j}(X(m))), \end{aligned} \tag{4.46}$$

so we can use Proposition 30 to get

$$G^{FR}_{\mathrm{j}^c(m)}\left(T_m\mathrm{j}(X(m)), T_m\mathrm{j}(X(m))\right) \leq G_{\mathrm{j}(m)}\left(T_m\mathrm{j}(X(m)), T_m\mathrm{j}(X(m))\right). \tag{4.47}$$

Which immediately implies

$$\mathcal{G}_m(X,X) \geq \mathcal{G}^c_m(X,X) \tag{4.48}$$

□

What we just proved can be seen as the formulation in the $W^\star$-algebraic framework of Equation (24) in [27].

This inequality allows us to formulate what we call a ***pre-Helstrom bound***.

Proposition 35 (Pre-Helstrom bound)**.** *Let $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ be a parametric model of states on $\mathscr{A}$, $\mathfrak{m}$ a measurement procedure, $\mathrm{j}^c = \mathfrak{m}^* \circ \mathrm{j}$ and $(\mathcal{M}, \mathrm{j}^c, \mathcal{P}(\Omega))$ the statistical parametric model associated to $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ via the measurement procedure $\mathfrak{m}$. Let also $\mathcal{G}$ be the J-metric (see Definition 26) of the parametric model $(\mathcal{M}, \mathrm{j}, \mathscr{A})$ and $\mathcal{G}^c$ be as in (4.13). Let also all the hypotheses of Proposition 33 hold, together with the additional requirement that*

$$C(X(\overline{m}), X(\overline{m})) \neq 0. \tag{4.49}$$

Then we have

$$\frac{\left(H_{\overline{m}}(X(\overline{m}), X(\overline{m}))\right)^2}{\mathcal{C}(X(\overline{m}), X(\overline{m}))} \leq \mathcal{G}^c_{\overline{m}}(X(\overline{m}), X(\overline{m})) \leq \mathcal{G}_{\overline{m}}(X(\overline{m}), X(\overline{m})) \tag{4.50}$$

Proof. The result comes trivially from combining Propositions 34 with Proposition 33. □

Remark 23. *There are two ways one can look at the result we just stated:*

If $\mathscr{A}$ is Abelian the kind of measurement procedures one can define are of the type described in Remark 22, meaning that is either some kind of ***coarse graining*** *or just a symmetry of the space of states of the algebra. So this result becomes a rephrasing of the principle that you can not gain information via coarse graining [33, 114, 155].*

If, on the other hand, $\mathscr{A}$ is non-Abelian, we look at this as performing some measurements on a quantum system, as in Examples 9 and 8. In the following we want to focus on this case, to investigate what this proposition says about the relation between Classical and Quantum Information Theory. For this reason for the rest of this section we will consider the case that $\mathscr{A}$ is non-Abelian and call $\mathcal{G}^c$ the classical metric and $\mathcal{G}$ the quantum metric.

This is called a Pre-Helstrom bound because, even if it is a weaker result than the Helstrom bound, it is on one hand preliminary to obtain the Helstrom bound and on the other hand shares with it some features. In particular, even if the specific form of the classical metric tensor $\mathcal{G}^c$ depends on the choice of the measurement procedure $\mathfrak{m}$, the bound holds in any case, and it holds for the same quantum metric $\mathcal{G}$ coming from the Jordan product on $\mathscr{A}$. This means that this inequality is universal in the sense that it is related to the quantum system associated to $\mathscr{A}$ and does not depend on which measurement one wants to perform on such system.

Now, if the model in consideration satisfies a Cramér-Rao bound, one can actually use the Pre-Helstrom in Proposition 35 to obtain

$$\operatorname{cov}(S) \geq I_F^{-1} \geq I_Q^{-1}, \tag{4.51}$$

which is precisely the Helstrom bound used in Quantum Detection and Estimation Theory.

Chapter 5

Examples

The aim of this work is to convince the reader that the formalism developed in Chapters 2 and 3 is suitable to describe both Classical and Quantum Information Geometry. Clearly, this can not be fully claimed if one is not able to reproduce some examples in both fields in this formalism. For this reason, we decided to devote a whole chapter, the present one, to carefully present several examples.

- In Section 5.1, the formalization of infinite-dimensional Classical Information Geometry given in [12, 13, 14] will be obtained by using the approach to Information Geometry introduced in this work. This will immediately give back also the finite-dimensional case as a particular case.
- In Section 5.2 quantum models will be discussed. First some general remarks on the Information Geometry for finite-dimensional quantum system will be given. Then a particular class of models, making use of an unfolded perspective of the space of quantum states [42, 46, 132], will be introduced. Finally, the infinite-dimensional quantum model of rank-one, strongly-continuous, unitary models will be discussed.
- In Section 5.3 we will introduce the concept of ***singular model*** and propose a list of examples that we plan to investigate in the future.

5.1 The classical case

The aim of this Section is to check whether the formalism developed in this work is a genuine generalization of Classical Information Geometry.

In Example 2, we already saw how the simplices, which are the appropriate setting for discussing finite-dimensional Classical Information Geometry, can be seen as the state of spaces of a finite-dimensional Abelian $C^{\star}$-algebra. What one could do is to take a parametric model (see Definition 20) that has as image the interior of the simplex, construct the J-metric associated to the parametric model as in 26 and obtain Fisher-Rao (see Equation (1.1)). We will follow a different approach and tackle directly a more interesting goal, i.e. to see if this framework can actually reproduce the elegant formalization of Classical Information Geometry in the infinite dimensional setting given in [12, 13, 14]. It will then be immediate to see how this can reproduce also the finite-dimensional case.

Let (Ω, Σ, μ) be a σ-finite measure space and consider $\mathscr{A} = \mathcal{L}^\infty(\Omega, \Sigma, \mu)$, as seen in Example 5, this is an Abelian $W^\star$-algebra. Again referring to the same example, recall that linear functionals are identified with complex measures and the space $\mathscr{P}$ of *n.p.l.f.s* can be identified with the space of non-negative functions in $L^1(\Omega, \Sigma, \mu)$. Let ω be a *n.p.l.f.* on $\mathscr{A}$, according to Theorem 10, the tangent double cone $T_\omega\mathscr{P}$ of $\mathscr{P}$ at ω is closed and it can be identified with $L^1(\Omega, \Sigma, \mu_\omega)$, that is, every $\xi \in T_\omega\mathscr{P}$ acts as

$$\xi(f) = \int_{\mathcal{X}} f \,\mathrm{d}\mu_\xi = \int_{\mathcal{X}} f\, F_\xi \,\mathrm{d}\mu_\omega, \tag{5.1}$$

where $F_\xi \in L^1(\Omega, \Sigma, \mu_\omega)$ the Radon-Nikodym derivative of μ_ξ with respect to μ_ω. Notice the last equation is just a realization of Equation (3.9).

Let now $\mathcal{M}$ be a Banach manifold and $(\mathcal{M}, \mathbf{p}, \mathscr{A})$ be a parametric model of *n.p.l.f.s* as in Definition 20, then we have that

$$\mathbf{p}(m) = p_m \mu, \tag{5.2}$$

where p_m is a positive function in $L^1(\Omega, \Sigma, \mu)$. This in turn implies that $(\mathcal{M}, \Omega, \mu, \mathbf{p})$ is a *parametrized measure model dominated by* μ in the languange of Classical Information Geometry [13, defn. 3.4, p. 150]. In particular, let $v_m \in T_m\mathcal{M}$, we have that $T_m\mathbf{p}(v_m) \in T_{\mathbf{p}(m)}\mathscr{P}$ can be associated to a Radon-Nikodym derivative

$$F_{\mathbf{p}(m)} = \frac{\mathrm{d}T_m\mathbf{p}(v_m)}{\mathrm{d}\mathbf{p}(m)}, \tag{5.3}$$

which is precisely what in [13, defn. 3.6, p. 152] is referred to as the logarithmic derivative of $\mathbf{p}$ at m in the direction v_m.

Consider now the characteristic function 1_A associated with the measurable subset $A \in \Sigma$ with $\mu_\omega(A) < \infty$. It is clearly an element of our $W^\star$-algebra$\mathscr{A} = \mathcal{L}^\infty(\Omega, \Sigma, \mu)$ and, again according to Equation (3.9), determines an element $\xi_A \in J_\omega$ as

$$\xi_A(f) = \omega(\{1_A, f\}) = \int_A f \,\mathrm{d}\mu_\omega. \tag{5.4}$$

This means that the J-Hilbert space $\mathcal{J}_\omega$ of Definition 24 can be identified with the real Hilbert space $L^2_{\mathbb{R}}(\mathcal{X}, \mu_\omega)$, this is consistent with the well-known fact that $L^q(\mathcal{X}, \mu_\omega) \subseteq L^p(\mathcal{X}, \mu_\omega)$ for all $1 \leq p \leq q \leq \infty$.

Considering what said until now, we can clearly write the tensor $\mathcal{G}$ defined in equation 3.33 as

$$\left(\mathcal{G}(X, Y)\right)(m) = \int_{\mathcal{X}} \frac{\mathrm{d}T_m\mathbf{p}(X_m)}{\mathrm{d}\mathbf{p}(m)} \frac{\mathrm{d}T_m\mathbf{p}(Y_m)}{\mathrm{d}\mathbf{p}(m)} \,\mathrm{d}\mathbf{p}(m). \tag{5.5}$$

Equation (5.5) coincides with the Fisher-Rao metric tensor as given in [13, eqn. 3.41, p. 136]. There it is also found that this defines a Riemannian structure only if the model $(\mathcal{M}, \Omega, \mathbf{p})$ is 2-integrable parametrized measure model, we can then claim that the notion of 2-integrable measure model is the realization in Classical Information Geometry of the notion of J-regular parametric model given in definition 25.

It is then immediate that if the space Ω is finite and we use as reference measure μ the counting measure, we fall back in the finite-dimensional case and the metric we obtain is precisely Fisher-Rao metric.

5.2 Quantum models

In this section, we will deal with quantum models, i.e. models where the $W^\star$-algebra considered is the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on some Hilbert space, we will introduce a family of models that are related to an unfolding procedure used in [42, 46] and then move to an infinite-dimensional example in Subsection 5.2.2.

Before doing so, let us briefly look again at what done in Section 2.3, i.e. to the finite-dimensional case, bearing in mind the notions introduced in Chapter 3 about parametric models of *n.p.l.f.s* on $W^\star$-algebras.

Let $\mathcal{H}$, and thus $\mathscr{A} = \mathcal{B}(\mathcal{H})$, be finite-dimensional, in this case, as seen in Section 2.3, there is a class of Riemannian manifolds that are given by the orbits of the action α defined in 2.60 and whose metric structure emerge from some sort of Jordan-analogue of Konstant-Kirillov-Souriau theory of coadjoint orbits of a Lie group (see [119, 120, 123, 165] and Remark 10). We can now see these as a *preferred* family of J-regular parametric models. In particular, since $\mathscr{A}$ is finite-dimensional, all linear functional are normal, and we have that $\mathscr{A}^* = \mathscr{A}_*$ and $\mathscr{A}_{sa}^* = (\mathscr{A}_{sa})_*$, and this allows us to define, for any $\omega \in A_{sa}^*$, the 2-times contravariant tensor

$$\mathcal{R}_\omega : T_\omega \mathscr{A}_{sa}^* \times T_\omega \mathscr{A}_{sa}^* \cong \mathscr{A}_{sa} \times \mathscr{A}_{sa} \longrightarrow \mathbb{R}, \tag{5.6}$$

whose action on $a, b \in A_{sa}$ is given by

$$\mathcal{R}_\omega(a, b) := \omega\left(\{a, b\}\right), \tag{5.7}$$

as in Equation 2.104. This can be seen as a geometric structure that encapsulates the algebraic structure of the Jordan product of the $W^\star$-algebra $\mathscr{A}$. The contravariant tensor $\mathcal{R}$ is smooth since it is linear in ω and it is symmetric since it is defined using a symmetric product, we will refer to this tensor in the following as the ***Jordan tensor*** on $\mathscr{A}_{sa}^*$.

The Jordan tensor $\mathcal{R}$ defines a generalized distribution $\mathcal{D} = \{\mathcal{D}_\omega\}_{\omega \in \mathscr{A}_{sa}^*}$ on $\mathscr{A}_{sa}^*$ as

$$\mathcal{D}_\omega := \{\eta \in T_\omega \mathscr{A}_{sa}^* \cong \mathscr{A}_{sa}^* \mid\ \exists\, \mathbf{a} \in \mathscr{A}_{sa} : \eta(\mathbf{b}) = \mathcal{R}_\omega(\mathbf{a}, \mathbf{b})\ \forall\, \mathbf{b} \in \mathscr{A}_{sa}\}. \tag{5.8}$$

Notice that, if ω is a *n.p.l.f.*, then we can use Equations (5.8) and (3.7), together with the fact that J_ω is closed when $\mathscr{A}$ is finite-dimensional, to conclude that

$$\mathcal{D}_\omega = J_\omega = AC_\omega. \tag{5.9}$$

The distribution $\mathcal{D} = \{\mathcal{D}_\omega\}_{\omega \in \mathscr{A}_{sa}^*}$ is referred to as the canonical distribution generated by the tensor $\mathcal{R}$.

Then again in analogy with the Kirillov-Kostant-Souriau theory of coadjoint orbits of a Lie group, it is possible to invert the tensor $\mathcal{R}$ on maximal leaves of the canonical distribution generated by the tensor $\mathcal{R}$. Its inverse is a symmetric tensor, that can be seen as a symmetric analogue of the Konstant-Kirillov-Souriau symplectic form on coadjoint orbits [54]. It can be, however, noticed a difference between the Konstant-Kirillov-Souriau construction and its symmetric analogue. In fact, the Poisson tensor on the dual of a Lie algebra in the Konstant-Kirillov-Souriau theory generates a symplectic foliation of $\mathscr{A}_{sa}^*$, while the canonical distributions of $\mathcal{R}$ does not generate any foliation of $\mathscr{A}_{sa}^*$. The tangent vectors η_a as defined in equation (3.7) span the tangent space at each point of any of the maximal leaves of $\mathcal{D}$ and the inverse of $\mathcal{R}$ is given pointwise by equation (3.20). Then it can also be seen that the space $\mathscr{P}$ of *n.p.l.f.s* of $\mathscr{A}$ is decomposed into the disjoint union of the maximal leaves of $\mathcal{D}$ and that the inverse $\mathcal{G}$ of $\mathcal{R}$ on any such leave is a Riemannian metric tensor.

This family of maximal leaves coincides with the *preferred* family of Jordan regular parametric models mentioned in the beginning of this section, i.e. it coincides with the orbits of the action α defined in Equation (2.60). It is worth mentioning that in [63] the authors obtain a Withney stratification of $\mathscr{P}$. We believe that it is possible that the analysis carried out in this work can be adapted to show that this stratification is given precisely by the maximal leaves of $\mathcal{D}$.

In the infinite dimensional case this picture breaks down, this is, in some way, the reason for the change of perspective to tackle the infinite-dimensional case that is the content of Chapter 3. In particular, instead of trying to generalize the foliation picture described here, what one does is to look at the maximal leaves of *n.p.l.f.s* of the canonical distribution of $\mathcal{R}$ in finite dimensions as a particular case of a parametric model in the sense of Definition 20. This models are then endowed with a Riemannian metric tensor that comes from the Jordan product in $\mathscr{A}_{sa}$ through the inverse of the tensor $\mathscr{R}$, a reformulation of this instance leads quite naturally to the notion of J-regular parametric model given in definition 25, which has the advantage of being applicable also to the infinite-dimensional case.

Remark 24. *Let us briefly discuss the infinite-dimensional generalization of the picture discussed here, in few words, the idea here is that equation (5.7) allows to translate the algebraic structure of Jordan algebra of $\mathscr{A}_{sa}$ into a $(2,0)$-contravariant tensor $\mathcal{R}$ on $(\mathscr{A}_{sa})_*$, and this is possible because $T_\xi^*(\mathscr{A}_{sa})_* \cong \mathscr{A}_{sa}$. Let us define the map*

$$\sharp\colon \mathscr{A}_{sa} \cong T_\xi^*(\mathscr{A}_{sa})_* \to T_\xi^{**}(\mathscr{A}_{sa})_* \cong \mathscr{A}_{sa}^* \tag{5.10}$$

whose action is given by

$$(\sharp(\mathbf{a}))\,(\mathbf{b}) := \mathcal{R}_\xi(\mathbf{a},\mathbf{b}) = \xi(\{\mathbf{a},\mathbf{b}\}). \tag{5.11}$$

The image of $T_\xi^*(\mathscr{A}_{sa})_*$ via this map is contained in $(\mathscr{A}_{sa})_* \subseteq \mathscr{A}_{sa}^* \cong T_\xi^{**}(\mathscr{A}_{sa})_*$, and this means that $\mathcal{R}$ satisfies a condition that is analogous to the one that is used in

[18] to define a Banach Lie-Poisson structure on $\mathscr{A}_{sa}$ stemming from the Lie algebra structure of $\mathscr{A}_{sa}$ given by the (scaled) commutator product in $\mathscr{A}$.

However, studying in a rigorous way the geometry of the distribution generated by $\mathcal{R}$ requires care for technical details given by the infinite-dimensional setting. As an example, notice that the vector space $\mathcal{D}_\xi$ is in general not closed. It would be of course interesting to try to pursue such an intent and see to what extent, or in which way, one can use the finite-dimensional theory developed in [54] in the infinite-dimensional setting and we hope that we will be able to investigate this issue in the future.

5.2.1 Unfolding of the space of quantum states as parametric model

Here we can exploit the unfolding procedure used in [42, 46, 71, 132] to produce a class of example, these models are intrinsically non-identifiable, but they have the interesting property of containing, in a sense that will be clear later, the Riemannian structures of a simplex equipped with the Fisher-Rao metric. Also they are a large class of examples, they contain, but are not limited to, all spaces of faithful and pure states of a finite-dimensional quantum system.

Let $\mathcal{H}$ be a finite-dimensional Hilbert space with $\dim_{\mathbb{C}} \mathcal{H} = n$, and $\{|e_1\rangle, \ldots, |e_n\rangle\}$ be a basis of $\mathcal{H}$. Consider the $C^\star$-algebra $\mathcal{B}(\mathcal{H})$, the basis $\{|e_1\rangle, \ldots, |e_n\rangle\}$ of $\mathcal{H}$ induces basis on $\mathcal{B}(\mathcal{H})$ given by elements of the form

$$e_{jk} = |e_j\rangle\langle e_k| \,. \tag{5.12}$$

Recalling Example 3, we can regard states as density matrices, and recalling the introduction to Section 2.2, we know that the unitary group $\mathcal{U}(\mathcal{H})$ can act on the space of quantum states $\mathscr{S}(\mathcal{H})$ with an action defined as

$$\delta : \mathcal{U}(\mathcal{H}) \times \mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}(\mathcal{H}) : \delta(U, A) := UAU^\dagger. \tag{5.13}$$

Let then ρ be a quantum state and let us denote its eigenvalues with p_j, it is well-known that every state ρ can be written as

$$\rho = U\rho_0 U^\dagger, \tag{5.14}$$

where $U \in \mathcal{U}(\mathcal{H})$ and ρ_0 is defined as

$$\rho_0 = \mathrm{diag}(p_1, p_2, \ldots, p_n). \tag{5.15}$$

Remark 25. *The element U of the unitary group $\mathcal{U}(\mathcal{H})$ in Equation (5.14) is not uniquely defined, in fact, the action defined in Equation (5.13) has always a non-trivial isotropy group. To see this, consider the action of an element $U' = e^{-i\phi}U$ with $0 \leq \phi < 2\mathrm{j}$, we clearly have*

$$U\rho\, U^\dagger = U'\rho\,(U')^\dagger, \tag{5.16}$$

for every state ρ, so that the isotropy group of the action defined in 5.13 at any point always contains $\mathcal{U}(1)$. The isotropy group of this action then also depends on the *point, in particular, it depends on the degeneracy of the eigenvalues of the considered state.*

*As an extreme example, consider the **maximally mixed state**, i.e. the state*

$$\rho_{mm} := diag(1/n, 1/n, \dots, 1/n) = \frac{1}{n}\mathbb{I}, \tag{5.17}$$

where we clearly have that the unitary group acts trivially, hence the isotropy group at this point coincides with the whole $\mathcal{U}(\mathcal{H})$.

Since ρ is a state, i.e. is semi-positive and has trace 1, its eigenvalues have to satisfy

$$\begin{aligned} p_j &\geq 0 \quad \forall j, \\ \sum_{j=1}^{n} p_j &= 1. \end{aligned} \tag{5.18}$$

Meaning that the n-tuple $\mathbf{p} = (p_1, p_2, \dots, p_n)$ can be seen as a probability vector over a space of events of cardinality n, hence as a point on the $(n-1)$-simplex (see Example 2). Moreover, if the state ρ is faithful, then none of its eigenvalues can be zero. This means that we can define an immersion map i from the interior Δ^o_{n-1} of the $(n-1)$-simplex into the space of faithful states $\mathscr{S}_n(\mathcal{H})$ as

$$\mathrm{i} : \Delta^o_{n-1} \ni \mathbf{p} = (p_1, p_2, \dots, p_n) \mapsto \begin{pmatrix} p_1 & 0 & \dots & 0 \\ 0 & p_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p_n \end{pmatrix} \in \mathscr{S}_n(\mathcal{H}). \tag{5.19}$$

Let now $\mathcal{M}(\mathcal{H}) = \mathcal{U}(\mathcal{H}) \times \Delta^o_{n-1}$, we can exploit Equation (5.14) to define a map j from $\mathcal{M}(\mathcal{H})$ to $\mathscr{S}_n(\mathcal{H})$ as

$$\mathrm{j}(U, \mathbf{p}) = U \mathrm{i}(\mathbf{p}) U^\dagger. \tag{5.20}$$

Recalling Remark 25, it is clear that the map j is non-injective. The map j is what in [42] is referred to as an unfolding map for the space of quantum states.

Here we would like to stress a different aspect, i.e. that the triple $(\mathcal{M}(\mathcal{H}), \mathrm{j}, \mathcal{B}(\mathcal{H}))$ is a parametric model of *n.p.l.f.s* (see Definition 20). It is immediately seen that this model is smooth and non-identifiable, and since the model is non-identifiable, it is clear that it is not possible to define a J-metric (see Definition 26), since the tensor obtained will always be degenerate. Nonetheless, let us proceed in constructing the tensor $\mathcal{G}$ on $\mathcal{M}(\mathcal{H})$.

According to Equation 3.33, in order to construct the bilinear form $\mathcal{G}_m$ on $\mathcal{M}(\mathcal{H})$ at m we need to compute the image via the tangent map of j of vectors in $T_m\mathcal{M}(\mathcal{H})$, let us denote by ρ the image of m via the map j of our model. Let then $X, Y \in T_m\mathcal{M}(\mathcal{H})$ and let $m := (U, p)$, we have

$$\mathrm{j}(m) = \mathrm{j}(U, p) = U \mathrm{i}(p) U^\dagger = \rho. \tag{5.21}$$

Given the product structure of $\mathcal{M}(\mathcal{H})$, we have that

$$T_m\mathcal{M}(\mathcal{H}) = T_U\mathcal{U}(\mathcal{H}) \oplus T_p\Delta^o_{n-1}, \tag{5.22}$$

and we also have that

$$T_U\mathcal{U}(\mathcal{H}) \cong \mathfrak{u}(\mathcal{H}) \cong \mathcal{B}(\mathcal{H})_{sa}, \tag{5.23}$$

while $T_p\Delta^o_{n-1} \cong \mathbb{R}^{n-1}$ and can be identified with vectors of $\mathbb{R}^n$ satisfying the constraint that their components have to sum to zero. It is then clear that we can write the vectors X and Y as

$$\begin{aligned} X &= (v, A), \\ Y &= (w, B), \end{aligned} \tag{5.24}$$

with $v, w \in \mathbb{R}^{n-1}$ and $A, B \in \mathcal{B}(\mathcal{H})_{sa}$.

Let then γ be a curve in $\mathcal{M}(\mathcal{H})$ such that $\gamma(0) = m$ and $\dot{\gamma}(t)|_{t=0} = X$, and let also $p(t)$ and $U(t)$ be curves respectively on Δ^o_{n-1} and $\mathcal{U}(\mathcal{H})$ such that $\gamma(t) = (U(t), p(t))$, then clearly we have that

$$\begin{aligned} \dot{p}(t)|_{t=0} &= v, \\ \dot{U}(t)|_{t=0} &= A. \end{aligned} \tag{5.25}$$

Taking into account what just said and by definition of tangent map, we have that

$$\begin{aligned} T_m\mathrm{j}X &= \frac{d}{dt}\left(\mathrm{j} \circ \gamma(t)\right)_{t=0} = \left(U(t)\mathrm{i}(p(t))(U(t))^\dagger\right)_{t=0} \\ &= U\mathrm{i}(v)U^\dagger + iA\mathrm{i}(p)U^\dagger - iU\mathrm{i}(p)A, \end{aligned} \tag{5.26}$$

where $\mathrm{i}(v)$ is the image via the immersion map i of v seen as a vector in $\mathbb{R}^n$ whose component sum to zero. In an analogous way we clearly get

$$T_m\mathrm{j}Y = U\mathrm{i}(w)U^\dagger + iB\mathrm{i}(p)U^\dagger - iU\mathrm{i}(p)B. \tag{5.27}$$

This would easily allow to compute the bilinear form at m as

$$\mathcal{G}_m(X, Y) = \mathrm{j}(m)(\{T_m\mathrm{j}X, T_m\mathrm{j}Y\}) = Tr\left(U\mathrm{i}(p)U^\dagger\{T_m\mathrm{j}X, T_m\mathrm{j}Y\}\right). \tag{5.28}$$

We will not give the explicit expression of the tensor $\mathcal{G}_m$ here, let us instead compute explicitly this expression for $X = (v, 0)$ and $Y = (w, 0)$,

$$\begin{aligned} \mathcal{G}_m((v,0),(w,0)) &= Tr\left(U\mathrm{i}(p)U^\dagger\{U\mathrm{i}(v)U^\dagger, U\mathrm{i}(w)U^\dagger\}\right) \\ &= Tr\left(U\mathrm{i}(p)\{\mathrm{i}(v),\mathrm{i}(w)\}U^\dagger\right) = Tr\left(\mathrm{i}(p)\mathrm{i}(v)\mathrm{i}(w)\right) \\ &= \sum_{j=1}^n p^j v^j w^j. \end{aligned} \tag{5.29}$$

Where we used the fact that U is a unitary operator, the cyclic property of the trace, and the fact that $\mathrm{i}(v)$ and $\mathrm{i}(w)$ commute. It is then immediate to see that

$$\mathcal{G}_m((v,0),(w,0)) = \sum_{j=1}^n p^j dp^j \otimes dp^j(v,w) \tag{5.30}$$

This shows that, even if this tensor does not define a metric, it has the remarkable properties that it contains a part that depends only on the classical probabilities given by the eigenvalues of the state, and its expression is the same as the Fisher-Rao metric (see Equation (1.1)). More specifically, one can define the map π as

$$\pi_{\Delta^o_{n-1}} : \mathcal{M}(\mathcal{H}) \ni (U, \mathbf{p}) \mapsto \mathbf{p} \in \Delta^o_{n-1}, \tag{5.31}$$

giving rise to the following picture,

$$\begin{array}{ccc} \mathcal{M}(\mathcal{H}) & \xrightarrow{\mathrm{j}} & \mathcal{B}(\mathcal{H}) \\ \downarrow{\scriptstyle \pi} & & \\ \Delta^o_{n-1} & & \end{array} \tag{5.32}$$

and, denoting by G^{n-1}_{FR} the Fisher-Rao metric defined on Δ^o_{n-1}, we have

$$\mathcal{G}_m((v,0),(w,0)) = (\pi^* G^{n-1}_{FR})((v,0),(w,0)). \tag{5.33}$$

Looking at Equation (5.33) is understood in which sense, as stated in the beginning of this subsection, this models contains in some sense the Riemannian structure of a simplex equipped with the Fisher-Rao metric.

On the other hand, even without making any explicit computation, we can make some remarks on the action of $\mathcal{G}_m$ on vectors of the type $(0,U),(0,V)$. In fact, let us consider the orbit $\mathcal{O}_{\mathbf{p}}$ of the action δ passing through $\mathbf{p}$, vectors of this type have to be tangent to such orbits, and to such orbits one can apply Proposition 31 to say that the metric obtained there must be invariant under the action of the unitary group. This gives rise for example, in the case that $\mathrm{i}(\mathbf{p})$ is a pure state, to the so-called Fubini-Study metric [84, 167].

The same idea used for defining the maps j and i can actually be used to define other models, one possibility is to define these maps in an analogous fashion, but taking probability vectors only from the interior of a k-dimensional face of the simplex. This would give a non-identifiable model where the image of the map j coincides with the stratum $\mathscr{S}_{k-1}(\mathcal{H})$ of quantum states with rank $k-1$. A particular case of this would be to take only one of the extremal points of the simplex. This would give rise to a model where the image of j coincides with the orbit of the unitary group given by pure states.

In general, instead of Δ^o_{n-1} one can choose any open subset of it or any open subset of any of the faces of the simplex and will always get a model that clearly satisfies the following property:

$$\rho \in \mathrm{j}(\mathcal{M}(\mathcal{H})) \implies U\rho U^\dagger \in \mathrm{j}(\mathcal{M}(\mathcal{H})) \quad \forall U \in \mathcal{U}(\mathcal{H}). \tag{5.34}$$

The relevance of this property can be understood by recalling Proposition 19 and Remark 20.

5.2.2 Rank-one, strongly-continuous unitary models

Let now $\mathscr{A}$ be the $W^\star$-algebra $\mathcal{B}(\mathcal{H})$ of bounded linear operator on the possibly infinite-dimensional, separable complex Hilbert space $\mathcal{H}$. The predual $\mathscr{A}_*$ of $\mathscr{A} = \mathcal{B}(\mathcal{H})$ is given by the space $\mathcal{T}(\mathcal{H})$ of trace-class operators on $\mathcal{H}$ (see Example 6). This means that every non-negative trace-class operator ω defines a *n.p.l.f.* on $\mathscr{A} = \mathcal{B}(\mathcal{H})$ as

$$\omega(a) = \mathrm{Tr}(\omega a) \quad \forall a \in \mathscr{A}, \tag{5.35}$$

where, with an abuse of notation, we denoted with ω the functional defined in terms of the trace-class operator ω.

We will now discuss a type of parametric models, that stems from strongly continuous unitary representations of Banach-Lie groups, and that is general enough to contain most of the parametric models used in the literature when describing with a quantum system in terms of a separable Hilbert space [85, 109, 110].

Let G be a Banach-Lie group, a ***strongly-continuous unitary representation*** of G on $\mathcal{H}$ is a group homomorphism π between G and the unitary group $\mathscr{U}(\mathcal{H})$ such that the map $\pi_\psi \colon \mathrm{G} \to \mathcal{H}$ defined by

$$\pi_\psi(\mathrm{g}) := (\pi(\mathrm{g}))(\psi) \tag{5.36}$$

is continuous with respect to the norm topology in $\mathcal{H}$ for all $\psi \in \mathcal{H}$. A vector $\varphi \in \mathcal{H}$ is then called ***smooth*** for π if the map π_ψ is smooth. If G is a finite-dimensional Lie group, then the existence of smooth vectors is guaranteed and they also form a dense subset of $\mathcal{H}$ [87].

Proposition 36. *Let $\pi \colon \mathrm{G} \to \mathscr{U}(\mathcal{H})$ be a strongly-continuous unitary representation, let φ be a smooth vector for π and let $\mathrm{j} : \mathrm{G} \to \mathscr{A} = \mathcal{B}(\mathcal{H})$ be defined as*

$$\mathrm{j}(\mathrm{g}) := |(\pi(\mathrm{g}))(\varphi)\rangle\langle(\pi(\mathrm{g}))(\varphi)| \equiv |\varphi_{\mathrm{g}}\rangle\langle\varphi_{\mathrm{g}}|. \tag{5.37}$$

Then the triple $(\mathrm{G}, \mathrm{j}, \mathscr{A} = \mathcal{B}(\mathcal{H}))$, is a J-smooth parametric model (see Definition 26).

Moreover, the tensor $\mathcal{G}$ defined as in equation (3.33) is invariant with respect to the canonical left action of G on itself.

Proof. Let us first prove that $(\mathrm{G}, \mathrm{j}, \mathcal{B}(\mathcal{H}))$ is a smooth parametric model as in Definition 20.

To see this, notice that j can be written as the composition of two smooth maps, namely the map π_φ and the map F from $\mathcal{H}$ to the space $\mathcal{T}_{sa}(\mathcal{H})$ of self-adjoint trace-class operators on $\mathcal{H}$ whose action is defined as $F(\psi) = |\psi\rangle\langle\psi|$ the smooth map $\pi_\varphi \colon \mathrm{G} \to \mathcal{H}$.

Then, by definition of tangent map, we have that

$$T_{\mathrm{g}}\pi_\varphi(v_{\mathrm{g}}) = \frac{\mathrm{d}}{\mathrm{d}t}\left((\pi(\mathrm{g}_t))(\varphi)\right)_{t=0} \tag{5.38}$$

for all $v_{\mathrm{g}} \in T_{\mathrm{g}}\mathrm{G}$, and again by definition of tangent map, we have

$$T_\psi F(\phi) = \frac{\mathrm{d}}{\mathrm{d}t}\left(|\psi_t\rangle\langle\psi_t|\right)_{t=0} = |\phi\rangle\langle\psi| + |\psi\rangle\langle\phi| \tag{5.39}$$

for all $\phi \in T_\psi\mathcal{H} \cong \mathcal{H}$. This means that we can write

$$T_{\mathrm{g}}\mathrm{j}(v_{\mathrm{g}}) = T_{\pi_\varphi(\mathrm{g})}F \circ T_{\mathrm{g}}\pi_\varphi(v_{\mathrm{g}}) = |v_{\mathrm{g}}\rangle\langle\varphi_{\mathrm{g}}| + |\varphi_{\mathrm{g}}\rangle\langle v_{\mathrm{g}}| \tag{5.40}$$

where, with an abuse of notation, we denoted on the right-hand side by v_{g} its image $T_{\mathrm{g}}\pi_\varphi(v_{\mathrm{g}})$ via the tangent map $T_{\mathrm{g}}\pi_\varphi$ of ϕ_φ at g. Notice that the quantity $C_\varphi := \langle\varphi_{\mathrm{g}}|\varphi_{\mathrm{g}}\rangle$ is a constant depending only on the reference vector φ and not on g and that it holds

$$\langle v_{\mathrm{g}}|\varphi_{\mathrm{g}}\rangle + \langle\varphi_{\mathrm{g}}|v_{\mathrm{g}}\rangle = 0 \tag{5.41}$$

for all $\mathrm{g} \in \mathrm{G}$ and all $v_{\mathrm{g}} \in T_{\mathrm{g}}\mathrm{G}$.

Since G is a Banach-Lie group, it is possible to naturally define a left action L of G on itself that satisfies

$$\mathrm{j} \circ L_{\mathrm{h}} = Ad_{\pi(\mathrm{h})} \circ \mathrm{j}, \tag{5.42}$$

where

$$Ad_{\pi(h)}(a) = \pi(h)a\pi(\mathrm{h})^\dagger \tag{5.43}$$

for all $a \in \mathscr{A}$ and for all $h \in \mathrm{G}$. Consequently, it we have that

$$T_{\mathrm{g}}\left(\mathrm{j} \circ L_h\right) = T_{\mathrm{g}}\left(Ad_{\pi(h)} \circ \mathrm{j}\right) = T_{\mathrm{j(g)}}Ad_{\pi(h)} \circ T_{\mathrm{g}}\mathrm{j} = Ad_{\pi(h)} \circ T_{\mathrm{g}}\mathrm{j}, \tag{5.44}$$

where the last equality follows from the fact that $Ad_{\pi(h)}$ is linear in a.

Again because of the fact that G is a Banach-Lie group, we can focus only on left-invariant vector fields in order to be able to write down $\mathcal{G}$ following equation (3.33). Let then X be a left-invariant vector field on G, it holds

$$X(\mathrm{g}) = T_{\mathrm{e}}L_{\mathrm{g}}(X(\mathrm{e})) \tag{5.45}$$

where we denoted with e the identity element in G. We can then use Equation (5.44) to obtain

$$\begin{aligned} T_{\mathrm{g}}\mathrm{j}(X(\mathrm{g})) = T_{\mathrm{g}}\mathrm{j} \circ T_{\mathrm{e}}L_{\mathrm{g}}(X(\mathrm{e})) &= \left(T_{\mathrm{g}}(\mathrm{j} \circ L_{\mathrm{g}})\right)(X(\mathrm{e})) \\ &= Ad_{\pi(\mathrm{g})} \circ T_{\mathrm{g}}\mathrm{j}(X(\mathrm{e})). \end{aligned} \tag{5.46}$$

Now, we note that $Ad_{\pi(\mathrm{g})}$ can be thought of as the dual of a norm-continuous, ultra-weakly continuous automorphism of $\mathscr{A} = \mathcal{B}(\mathcal{H})$, and thus, according to remark 20, it holds

$$G_{\mathrm{j(g)}}(T_{\mathrm{g}}\mathrm{j}(X(\mathrm{g})), T_{\mathrm{g}}\mathrm{j}(Y(\mathrm{g}))) = G_{\mathrm{j(e)}}(T_{\mathrm{e}}\mathrm{j}(X(\mathrm{e})), T_{\mathrm{e}}\mathrm{j}(Y(\mathrm{e}))) \tag{5.47}$$

for all left-invariant vector fields X, Y on G. Equation (5.47) allows us to claim that, in order to get an expression of the J-metric $\mathcal{G}$, it is enough to compute its value at the identity element e of G.

Let now $\{\phi_j\}_{j\in\mathbb{N}}$ be an orthonormal basis of the separable Hilbert space $\mathcal{H}$ such that $\varphi = A\,\phi_0$, with $A \in \mathbb{C}$, and recall equation (5.40), we have that for every left-invariant vector field X on G there is a vector $v_e \in \mathcal{H}$ such that

$$T_{\mathrm{e}}\mathrm{j}(X(\mathrm{e})) = |\varphi\rangle\langle v_e| + |v_e\rangle\langle\varphi|, \tag{5.48}$$

and we can then write

$$v_e = v_e^0\phi_0 + \sum_{k=1}^{\infty} v_e^k\phi_k = v_e^0\phi_0 + v_e^{\perp}, \tag{5.49}$$

where we implicitly defined

$$\sum_{k=1}^{\infty} v_e^k\phi_k := v_e^{\perp}. \tag{5.50}$$

Then let us define $\mathbf{X}_{\mathrm{e}} \in \mathscr{A}_{sa} = (\mathcal{B}(\mathcal{H}))_{sa}$ as

$$\mathbf{X}_{\mathrm{e}} = \frac{\overline{A}v_e^0 + A\overline{v_e^0}}{|A|^2}\,|\phi_0\rangle\langle\phi_0| + \frac{1}{2|A|^2}\left(|\varphi\rangle\langle v_e^{\perp}| + |v_e^{\perp}\rangle\langle\varphi|\right), \tag{5.51}$$

then, noticing that $\mathrm{j}(\mathrm{e}) = |\varphi\rangle\langle\varphi|$, it is a matter of straightforward computation to check that

$$|\varphi\rangle\langle v_e| + |v_e\rangle\langle\varphi| = T_{\mathrm{e}}\mathrm{j}(X(\mathrm{e})) = \{\mathrm{j}(\mathrm{e}), \mathbf{X}_{\mathrm{e}}\} \tag{5.52}$$

from which we get

$$Tr(T_{\mathrm{e}}\mathrm{j}(X(\mathrm{e}))a) = Tr(\mathrm{j}(\mathrm{e})\{\mathbf{X}_{\mathrm{e}}, a\}) \tag{5.53}$$

for all $a \in \mathscr{A} = \mathcal{B}(\mathcal{H})$. Now let Y be a left-invariant vector field on G, $\mathbf{Y}_{\mathrm{e}}$ be defined in an analogous way as $\mathbf{X}_{\mathrm{e}}$ and recall equation (3.20), we have

$$G_{\mathrm{j}(\mathrm{e})}(T_{\mathrm{e}}\mathrm{j}(X(\mathrm{e})), T_{\mathrm{e}}\mathrm{j}(Y(\mathrm{e}))) = Tr(\mathrm{j}(\mathrm{e})\{\mathbf{X}_{\mathrm{e}}, \mathbf{Y}_{\mathrm{e}}\}). \tag{5.54}$$

Since $\mathbf{X}_{\mathrm{e}} \in \mathscr{A}_{sa} = (\mathcal{B}(\mathcal{H}))_{sa}$, we can evaluate Equation (5.53) for $a = \mathbf{X}_{\mathrm{e}}$ and replacing $X(\mathrm{e})$ with $Y(\mathrm{e})$. This, together with Equation (5.40) and recalling that $C_{\varphi} = \langle\varphi|\varphi\rangle$, allows us to compute

$$\begin{aligned} G_{\mathrm{j}(\mathrm{e})}(T_{\mathrm{e}}\mathrm{j}(X(\mathrm{e})), T_{\mathrm{e}}\mathrm{j}(Y(\mathrm{e}))) &= Tr(T_{\mathrm{e}}\mathrm{j}(Y(\mathrm{e}))\mathbf{X}_{\mathrm{e}}) \\ &= C_{\varphi}(\langle\varphi|\mathbf{X}_{\mathrm{e}}|Y(\mathrm{e})\rangle + \langle Y(\mathrm{e})|\mathbf{X}_{\mathrm{e}}|\varphi\rangle). \end{aligned} \tag{5.55}$$

Let us write the left-hand side of Equation (5.53) for

$$a = |Y(\mathrm{e})\rangle\langle\varphi| + |\varphi\rangle\langle Y(\mathrm{e})| \tag{5.56}$$

and using Equation (5.41), a direct computation shows that we have

$$\begin{aligned} Tr(T_{\mathrm{e}}\mathrm{j}(X(\mathrm{e}))a) = C_{\varphi}(&\langle X(\mathrm{e})|Y(\mathrm{e})\rangle + \langle Y(\mathrm{e})|X(\mathrm{e})\rangle) \\ &+ 2\langle X(\mathrm{e})|\varphi\rangle\langle Y(\mathrm{e})|\varphi\rangle. \end{aligned} \tag{5.57}$$

The right-hand side, with the same position and using again Equation (5.41), becomes

$$Tr(\mathrm{j}(\mathrm{g})\{\mathbf{X}_{\mathrm{e}}, \mathbf{a}\}) = \frac{C_{\varphi}}{2}\left(\langle\varphi|\mathbf{X}_{\mathrm{e}}|Y(\mathrm{e})\rangle + \langle Y(\mathrm{e})|\mathbf{X}_{\mathrm{e}}|\varphi\rangle\right) \tag{5.58}$$

Eventually, we obtain

$$\begin{aligned} G_{\mathrm{j(e)}}(T_{\mathrm{e}}\mathrm{j}(X(\mathrm{e})), T_{\mathrm{e}}\mathrm{j}(Y(\mathrm{e}))) = 2C_{\varphi}(\langle X(\mathrm{e})|Y(\mathrm{e})\rangle + \langle Y(\mathrm{e})|X(\mathrm{e})\rangle) \\ + 4\langle X(\mathrm{e})|\varphi\rangle\langle Y(\mathrm{e})|\varphi\rangle. \end{aligned} \tag{5.59}$$

Then recalling Equations (3.33) and 5.47, we have

$$\begin{aligned} (\mathcal{G}(X, Y))\,(\mathrm{g}) = 2C_{\varphi}(\langle X(\mathrm{e})|Y(\mathrm{e})\rangle + \langle Y(\mathrm{e})|X(\mathrm{e})\rangle) \\ + 4\langle X(\mathrm{e})|\varphi\rangle\langle Y(\mathrm{e})|\varphi\rangle, \end{aligned} \tag{5.60}$$

which shows that $\mathcal{G}$ is indeed a smooth tensor field on G, allowing us to conclude that $(\mathrm{G}, \mathrm{j}, \mathcal{B}(\mathcal{H}))$ is a J-smooth parametric model as claimed. □

Corollary 4. *Let $(G, \mathrm{j}, \mathcal{B}(\mathcal{H}))$ be a J-smooth parametric model and G_{φ} the isotropy subgroup of G at $\varphi \in \mathcal{H}$ with respect to π,. Let also $G/G_{\varphi} \cong \mathcal{M}$ be a smooth manifold for which there exists a smooth section $\sigma\colon \mathcal{M} \to G$.*

Then $(\mathcal{M}, \mathrm{j}\circ\sigma, \mathcal{B}(\mathcal{H}))$ is a J-smooth parametric model whose J-metric $\mathcal{G}$ is invariant with respect to the canonical action of G on $\mathcal{M}$.

Proof. See Equation (5.47) in the proof of Proposition 36. □

Proposition 36, together with Corollary 4 allow us to claim that this procedure furnishes the complex projective space $\mathbb{CP}(\mathcal{H}) \cong \mathcal{U}(\mathcal{H})/U(1)$ with a Riemannian metric tensor $\mathcal{G}$ which is invariant under the canonical action of $\mathcal{U}(\mathcal{H})$ on $\mathbb{CP}(\mathcal{H})$. This immediately implies that the J-metric $\mathcal{G}$ of our models of rank-one, strongly continuous unitary model has to be a constant multiple of the Fubini-Study metric tensor. We have thus obtained the standard Riemannian structure of the space of normal pure states that is well-known in the context of geometric Quantum Mechanics [10, 55, 77, 118].

Moreover, if G is a finite-dimensional Lie group and π a continuous unitary representation for which both proposition 36 and corollary 4 are valid, then we obtain parametric models of *n.p.l.f.s* that are basically the well-known and widely used coherent states [4, 145].

5.3 Singular models

In Classical Information Geometry it is common to restrict the attention to the interior Δ_n^o of the n-simplex Δ_n, this is done to ensure that the Fisher-Rao metric, as defined in (1.1), is always a well-defined metric tensor, since the probabilities p^j are all strictly positive. This will of course be the case whenever the image of a model is an open in Δ_n^o.

Let us now consider the case in which the image of our model lies completely on the boundary of Δ_n, and more in particular lies completely on the interior of a k-dimensional face F_n^k of Δ_n. This poses no problems at all, since any face of the simplex is isomorphic to a lower-dimensional simplex, so we can always think of F_n^k as the k-dimensional simplex Δ_k and proceed as before, provided that the image of our model is an open set of the face F_n^k.

This brings us to the following consideration, if one takes the approach of defining directly the Fisher-Rao metric on the interior of the simplex, then the models that can be considered are forced to be some open sets of some simplices. What this actually means is that one is already silently assuming that all the probability distributions in his models have the property that if one assigns probability zero to some events, then all the others distributions assign probability zero to those same events.

An analogous requirement can be found, in the infinite-dimensional case, in the work [153] by Giovanni Pistone and Carlo Sempi, where they construct a geometric picture of finite measures that are, in the notation used there, *compatible with a fixed measure* μ_0, which means that they have the same null sets.

Something similar happens in Quantum Information Geometry, where it is a common approach to define the quantum metrics on the space $\mathscr{S}_+$ of faithful states [42, 46]. This is analogous to restrict the attention to the interior of the simplex, since the space $\mathscr{S}_+$ is the interior of the space of faithful states. Another possibility is to restrict one's attention to the space of pure states, and once again what is happening is that one is somehow fixing the rank of the possible states on the model. Let us stress, as an example, that the orbits of $GL(A)$ considered in Section 2.3 and the models that make use of the unfolding presented in Subsection 5.2.1 are all models that have this kind of regularity.

We then say that a model is ***singular*** whenever it is not regular in the sense just specified.

One advantage of our approach is that the metric tensor is not defined on the space of states of the $W^\star$-algebra considered, but there we have some structure that allows then to recover a metric on the parameter manifold $\mathcal{M}$, for this reason, we think that it can be successfully used to describe singular models.

Even if no complete analysis has been carried out for any of the following example, we will now present a small list of examples of singular models that we plan to investigate better in the future.

Example 10. *This first example is a classical, finite-dimensional example of a singular model. Let S^1 be the unit circle, it can be immersed, let us denote be* j *the immersion map, in the 2-simplex (see Example 2) in such a way that three points of S^1 touch each one of the 1-dimensional faces of the simplex, see Figure 5.1. It is then immediate to see that the triple* $(S^1, \mathrm{j}, \mathbb{C}^2)$ *is a smooth parametric model of n.p.l.f.s.*

Example 11. *This example is a quantum, finite-dimensional singular model somehow inspired by the one presented in Example 10.*

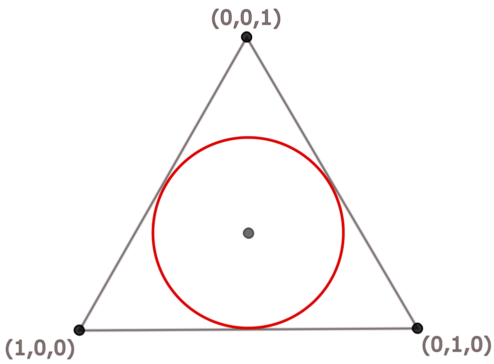


Figure 5.1: Pictorial representations of the model presented in Example 10. The red circle represents $\mathrm{j}(S^1)$.

Let us consider the case of $A = \mathcal{B}(\mathcal{H})$ with $\mathcal{H}$ finite-dimensional (of complex dimension n) as in Example 3. In particular consider the qubit case, i.e. let $n = 2$. We can parametrize the space of states of this algebra, i.e. the space of density matrices, as

$$\rho(x_1, x_2, x_3) = \frac{1}{2}\mathbb{I} + \sum_{j=1}^{3} x_j \sigma_j, \tag{5.61}$$

where the matrices σ_j are the Pauli matrices, defined as

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.62}$$

It is easily seen that this parametrization gives rise to the interior of the Bloch sphere (i.e. the space of faithful states) if x_1, x_2 and x_3 satisfy

$$x_1^2 + x_2^2 + x_3^2 < \frac{1}{4}, \tag{5.63}$$

and a parametrization of the boundary of the Bloch sphere when they satisfy

$$x_1^2 + x_2^2 + x_3^2 = \frac{1}{4}. \tag{5.64}$$

Let us now consider the map j *defined as*

$$\mathrm{j} : [0, \pi] \times [0, 2\pi) \ni (\theta, \phi) \mapsto \rho\left(\frac{1}{2}\sin(\theta)\cos(\phi), \frac{1}{2}\sin(\theta)\sin(\phi), k\frac{1}{2}\cos(\theta)\right), \tag{5.65}$$

with $k \in \left(0, \frac{1}{2}\right)$. The image of this map is an ellipsoid contained in the Bloch sphere that has the property of being contained in the open interior of the Bloch sphere for

$\theta \neq \pi/2$, while the circle given by the intersection of the ellipsoid together with the constraint $\theta = \pi/2$ is a subset of the boundary.

The triple $([0,\pi]\times[0,2\pi), \mathrm{j}, \mathcal{B}(\mathcal{H}))$ is then a singular model, since it contains both rank-one (i.e. pure) and rank-two (i.e., in this case, faithful) states.

Let us denote with E the image of the model, let $\rho \in E$ be given by the coordinates x_1, x_2, x_3 and notice that a vector $A \in T_\rho E$ can be identified with a traceless self-adjoint operator

$$a = \sum_{j=1}^{3} a_j \sigma_j \tag{5.66}$$

satisfying the property

$$a_1 x_1 + a_2 x_2 + \frac{a_3 x_3}{k} = 0. \tag{5.67}$$

The expression for the J-metric of this model can be easily computed using the properties of the trace and of the Pauli matrices, giving

$$\mathcal{G}_{(\theta,\phi)}(A,B) = \sum_j a_j b_j, \tag{5.68}$$

where a and b are related to A and B as before, and considering the fact that A and B must be tangent to $\mathrm{j}(\theta,\phi)$ we have that the components of a and b must be related to θ and ϕ as

$$\begin{aligned}\frac{a_1}{2}a_1\sin(\theta)\cos(\phi) + \frac{a_2}{2}\sin(\theta)\sin(\phi) + \frac{a_3}{2k}\cos(\theta) = 0\\ \frac{b_1}{2}a_1\sin(\theta)\cos(\phi) + \frac{b_2}{2}\sin(\theta)\sin(\phi) + \frac{b_3}{2k}\cos(\theta) = 0\end{aligned} \tag{5.69}$$

Example 12. *Let us consider again the Bloch sphere as in the previous example and let the vector field X_b be defined on the Bloch sphere as in Equation (2.91). As seen in Section 2.3, these vector fields are tangent to the orbits of the action ζ defined in that same section. For this reason, if we consider the dynamics governed byX_b starting at a pure state ρ_0, the trajectory will be contained in the space of pure states.*

This is not true for the vector fields Y_a, see again Equation (2.91), and actually it can be related to the so called ***Gorini–Kossakowski–Sudarshan–Lindblad equation*** *[43, 94], which governs the dynamics of open quantum systems. Let then γ be the trajectory on the space of quantum states of the state ρ_0 subject to a dynamics given by the vector field*

$$V_{ab}(\gamma(t)) = f(t)Y_a(\gamma(t)) + X_b(\gamma(t)) \tag{5.70}$$

where $f(t)$ is the function

$$f(t) = \begin{cases} e^{-1/t} & \text{if } \; 0 < t < T \\ 0 & \text{if } -T < t \leq 0 \end{cases} \tag{5.71}$$

which is a well-known example of smooth but non-analytic function.

Then the trajectory given by the dynamics governed by these vector field starting at ρ_0 *will have the property of belonging to the space of pure states for* $-T < t \leq 0$ *and then transitioning to the interior of the Bloch sphere, i.e. to faithful states, for* $0 < t < T$. *Hence the map* γ *realizes the singular model* $((-T, T), \gamma, \mathcal{B}(\mathcal{H}))$.

This model can be physically interpreted as follows. The dynamics given by X_b *can be seen as given by a conservative Hamiltonian* H_0, *while the one governed by* Y_a *can be seen as given by a non-conservative Hamiltonian* $\widetilde{H}$, *for example associated to some interaction with the environment. Then the dynamics given by* V_{ab} *will be associated to an Hamiltonian*

$$H = H_0 + f(t)\widetilde{H}, \tag{5.72}$$

i.e. it can be seen as some conservative dynamics where a non-conservative perturbation is turned on at the time $t = 0$ *in a smooth way.*

Chapter 6

Conclusions

Let us end this work with a list of conclusions and outlooks. We decided not to mark a definite separation between conclusions and outlooks, but, rather, merge them in a way that highlights the possible connections between them.

1. We have showed (see Examples 2 and 3, Section 2.3 and Chapter 5) that it is possible to reconstruct the background space of most of the examples appearing in both Classical and Quantum Information Geometry in terms of models of *n.p.l.f.s* on a $W^{\star}$-algebra. This is for sure a needed first step in the analysis, but in order to claim that one recovers structures that are related to Information Geometry one must be able to recover also the appropriate Riemannian structure. Regarding this, we saw that a Riemannian structure coming from the Jordan product of the $W^{\star}$-algebra allows to reconstruct the only metric that is relevant in Classical Information Geometry, i.e. Fisher-Rao metric, and some metrics that are widely used in Quantum Information Geometry, such as the Bures-Helstrom metric and the Fubini-Study metric.

   However, in the quantum case, there is a whole family of metrics, characterized by Petz, that can be constructed, see the Introduction of this work and the works [149, 151]. Some of these metrics are widely used in Quantum Information Geometry, without claiming to be exhaustive, let us just mention the ***Wigner-Yanase metric tensor*** [91, 92, 102], the ***Bogolubov-Kubo-Mori metric tensor*** [138, 148] and the ***quantum Tsallis metric tensors*** [132]. So one is naturally brought to wonder whether it is possible to obtain all these other metrics in our framework. Let us give two tentative answers to this question.

   One possibility is that one can define alternative Jordan products on the self-adjoint elements of the $W^{\star}$-algebra considered to obtain different metrics, thus establishing some sort of correspondence between Jordan products on the algebras and metric structures. Since this instance only arises in the quantum case, it is only reasonable to think that this should be connected to the non-commutativity of the product, since in the classical case the usual product of the algebra is already a Jordan product and represents somehow a natural choice.

The other possibility would be to impose the monotonicity condition (see Proposition 30) and look for all the possible metric tensors satisfying it, trying to reproduce in this framework what done by Petz in [149]. At this purpose, let ρ be a quantum state, Petz' result is expressed in terms of the superoperator $L_\rho R_\rho^{-1}$, which is precisely the form of the modular operator for $\mathcal{B}(\mathcal{H})$ with $\mathcal{H}$ finite-dimensional. It would then be reasonable to think that the modular operator Δ_ρ associated with each state will play a prominent role in this discussion. This approach would be more in line with the one advocated in [124].

Two crucial results, respectively in Classical and Quantum Information Geometry, are Čencov theorem and the Petz characterization of quantum metrics (see [34], [149] and the Introduction of this work). And one could wonder how do these results fit in our approach.

We would claim that Čencov theorem is just in some sense built in the framework. In fact, Čencov theorem selects the Fisher-Rao metric as the only metric that is invariant under ***sufficient statistics*** and the metric one constructs in the Classical case exploiting the Jordan product on the algebra is precisely Fisher-Rao metric.

However, it would be more satisfactory to be able to obtain the statement of the theorem in our framework. We believe that it is reasonable to assume that the notion of invariance under congruent embeddings used by Čencov in the proof of the result can be somehow reformulated in terms of the monotonicity and invariance properties stated in Subsection 3.3.1

Regarding Petz characterization, a big obstacle in obtaining it is the aspect that is highlighted in Point 1 of this list. The *plethora* of quantum metrics selected by monotonicity under CPTP maps is just not present in our framework at the moment. Whether the characterization of Petz can be reproduced in our framework thus heavily depends on whether we are able to obtain other quantum metrics from the Jordan product.

However, the characterization of quantum metrics by Petz characterization is a result that depends on the quantum version of the ***data processing inequality*** [16, 29, 59], and Proposition 30 is an inequality that is reminiscent of this inequality, so it is reasonable to think that this Proposition will be an essential ingredient if one wants to prove this result in this framework.

2. Referring to the notion of ***statistical manifold***, as introduced in [126], we have that the dually flat Riemannian structure of Classical Information Geometry is given in terms of a three times covariant tensor, that goes under the name of ***Amari-Čencov tensor***.

   Since we introduced a two times covariant tensor $\mathcal{G}_m$ (see Equation (3.33)) in terms of the Jordan product, one can proceed in analogy with this instance and wonder whether the Amari-Čencov tensor can be reconstructed from the ***Jordan triple product*** of the self-adjoint part of the $W^{\star}$-algebra considered.

In fact, any Jordan algebra allows to define such a triple product $\{\cdot,\cdot,\cdot\}$ [111], and in the case that the Jordan product is given by the anti-commutator product $\{\cdot,\cdot\}$ defined on the $W^{\star}$-algebra $\mathscr{A}$, we have

$$\{a,b,c\} = \{\{a,b\},c\} + \{a,\{b,c\}\} - \{b,\{a,c\}\} = \frac{1}{2}(abc + cba), \tag{6.1}$$

with $a,b,c \in \mathscr{A}$. This product is trilinear, and it satisfies

$$\begin{aligned} &\{a,b,c\} = \{c,b,a\}, \\ &\{a,b,\{c,d,e\}\} = \{\{a,b,c\},d,e\} - \{c,\{b,a,d\},e\} + \{c,d,\{a,b,e\}\}, \end{aligned} \tag{6.2}$$

where d and e are clearly also elements of $\mathscr{A}$. Let now ω be a faithful *n.p.l.f.*and recall Equation (3.7), the idea is to define a map

$$T_{\omega} : J_{\omega} \times J_{\omega} \times J_{\omega} \to \mathbb{R} \tag{6.3}$$

as

$$T_{\omega}(\eta_a,\eta_b,\eta_c) = \omega(\{a,b,c\}) = \frac{1}{2}\omega(abc + bca) \tag{6.4}$$

where $\eta_a,\eta_b,\eta_c \in J_{\omega}$ are related to a, b and c as in (3.9). Since ω is a faithful functional,

$$\eta_a = 0 \implies a = 0, \tag{6.5}$$

so that

$$T_{\omega}(\eta_a,\eta_b,\eta_c) = 0 \quad \forall \eta_b,\eta_c \in J_{\omega}. \tag{6.6}$$

We thus have that is well-defined on $J_{\omega} \times J_{\omega} \times J_{\omega}$ and it is clearly multilinear and symmetric in the second and third variable.

A straightforward computation for the case that $\mathscr{A}$ is a finite-dimensional, Abelian $W^{\star}$-algebra shows that T_{ω} reduces to the usual Amari-Čencov tensor on the manifold of strictly positive probability vectors. We plan to investigate whether this same instance arises even in other cases, i.e. if we allow $\mathscr{A}$ to be either non-Abelian or infinite-dimensional, or both.

3. Another procedure that is typical of Information Geometry is to construct the metric structure from a two-point function, called a ***divergence function***. More specifically, one constructs the metric tensor from second order derivatives of these functions, that typically have the information-theoretical interpretation of a measure of difference of entropy [7, 8, 11].

   In Classical Information Geometry, we have that the only relevant metric, i.e. the Fisher-Rao metric, can be constructed in this way from ***Shannon relative entropy*** [7, 163]. In Quantum Information Geometry, we have a *plethora* of divergence functions that can be used, reflecting the non-existence of a unique (or privileged) metric structure. Two examples of this instance are the fact that the Bogolubov-Kubo-Mori metric tensor can be constructed from the ***Von Neumann relative entropy*** and the Bures-Helstrom metric can be constructed from the ***root fidelity*** [46, 132].

A possible further development would be to introduce in this picture an analogue of divergence functions that can be defined on $W^\star$-algebras, i.e. the quasi-entropies introduced by Dénes Petz in [147] and try to replicate this procedure in our framework.

Referring to Point 1 of this list, it would be reasonable to assume that the existence of a privileged metric structure, and hence of a privileged divergence function in Classical Information Geometry should arise from the commutativity of the product on the underlying $W^\star$-algebra. In the quantum, finite-dimensional case, on the other hand, one expects to obtain results that are in line with what done in [127].

4. The ideas at Points 2 and 3 should lead to a better understanding of the role of the Amari-Čencov tensor and of divergence functions in relation with the Jordan product of the underlying $W^\star$-algebra.

   Let us now add to the picture the fact that the Amari-Čencov tensor can be constructed from third order derivatives of divergence functions [7, Chapter 6], it is then our hope to succeed in replicating this procedure in our framework.

   Let us notice that the 3-tensor extracted from a divergence function will always be totally symmetric, essentially because of equality of mixed partial derivatives. While, since the triple product is not totally symmetric, the tensor associated with the Jordan triple product, as proposed in Point 2, for a non-associative Jordan algebra is in general not totally symmetric. Indeed, the connections it gives rise to have torsion, unlike the ones usually employed in Information Geometry.

   This consideration seems to hint to the following picture for the non-Abelian case: it seems that divergence functions and Jordan product lead to the same type of geometry, i.e. to the same metric tensors, while triple derivatives of divergence functions and triple products give rise to qualitatively different objects, i.e. connections that in one case are torsionless and in the other they are not. It would then be interesting to investigate how should this instance be interpreted and what are the kind of information-theoretic tasks that dictate the use of one geometry instead of the other.

   Finally, another question that naturally arises is whether higher order tensors can be defined, and what is their role in this framework.

5. In Subsection 4.2.1, Proposition 33 is proved. We discussed how this can be seen as a preliminary result in obtaining the Cramér-Rao bound in our framework. Regarding this instance, there are two possible (and non-exclusive) ways one can proceed.

The first possibility is just to try to find a proof of the Cramér-Rao bound that does not rely on a maximization process as in Equation (4.40). In fact, as we have seen, that is the part of the proof whose generalization to the infinite-dimensional setting is not possible in general.

Another possibility is to conduct a case-by-case analysis and produce explicit infinite-dimensional examples of models and see whether one is able to obtain the bound for the particular model under investigation.

6. Basically the same line of reasoning as in Point 5 of this list can be applied to the Helstrom bound (see Subsection 4.2.2).

   If one is able to prove the Cramér-Rao bound, then the Helstrom bound is immediately recovered using Proposition 35. If the case-by-case analysis allows to reveal a classical model $(\mathcal{M}, \mathrm{j}^c, \mathscr{A}^c)$ that satisfies the Cramér-Rao bound, then one can construct some quantum model $(\mathcal{M}, \mathrm{j}, \mathscr{A}^q)$ and a measurement procedure $\mathfrak{m}$ such that $\mathrm{j}^c = \mathfrak{m}^* \circ j$ and the Helstrom bound will apply to the model $(\mathcal{M}, \mathrm{j}, \mathscr{A}^q)$.

7. In Section 5.3 the concept of ***singular model*** is introduced, there we also give the reasons why we believe that our approach is suitable to describe such models; moreover, a list of possible models is presented. A plan for the near future would be to investigate in detail the examples presented.

Bibliography

[1] M. C. Abbati, R. Cirelli, P. Lanzavecchia, and A. Mania. Pure states of general quantum-mechanical systems as kähler bundles. *Nuovo Cimento B;(Italy)*. ↓21

[2] R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, tensor analysis, and applications*. Springer-Verlag, New York, third edition, 2012. ↓31

[3] A. C. Aitken and H. Silverstone. On the estimation of statistical parameters. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 61(2):186–194, 1942. ↓7

[4] T. S. Ali, J.P. Antoine, and J.P. Gazeau. *Coherent states, wavelets, and their geeralizations*. Springer-Verlag, New York, 1999. ↓96

[5] S. I. Amari. A foundation of information geometry. *Electronics and Communications in Japan (Part I: Communications)*, 66(6):1–10, 1983. ↓1

[6] S. I. Amari. *Differential-Geometrical Methods in Statistics*, volume 28 of *Lecture Notes in Statistics*. Springer-Verlag, Berlin Heidelberg, 1985. DOI: 10.1007/978-1-4612-5056-2. ↓1

[7] S. I. Amari. *Information Geometry and its Application*. Springer, Japan, 2016. ↓1, 103, 104

[8] S. I. Amari and H. Nagaoka. *Methods of Information Geometry*. American Mathematical Society, Providence, RI, 2000. DOI: 10.1090/mmono/191. ↓1, 103

[9] W. Arveson. *An invitation to C*-algebras*, volume 39. Springer Science & Business Media, 2012. ↓11, 13

[10] A. Ashtekar and T. A. Schilling. Geometrical formulation of quantum mechanics. In A. Harvey, editor, *On Einstein's Path: Essays in Honor of Engelbert Schucking*, pages 23 – 65. Springer-Verlag, New York, 1999. ↓21, 96

[11] N. Ay and S. I. Amari. A novel approach to canonical divergences within information geometry. *Entropy*, 17(12):8111 – 8129, 2015. ↓103

[12] N. Ay, J. Jost, H. V. Le, and L. Schwachhöfer. Information geometry and sufficient statistics. *Probability Theory and Related Fields*, 162(1):327– 364, 2015. ↓2, 85, 86

[13] N. Ay, J. Jost, H. V. Le, and L. Schwachhöfer. *Information Geometry*. Springer International Publishing, 2017. DOI: 10.1007/978-3-319-56478-4. ↓2, 6, 9, 54, 75, 85, 86

[14] N. Ay, J. Jost, H. V. Le, and L. Schwachhöfer. Parametrized measure models. *Bernoulli*, 24(3):1692 – 1725, 2018. ↓85, 86

[15] O. E. Barndorff-Nielsen, R. D. Gill, and P. E. Jupp. On quantum statistical inference. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 65(4):775–816, 2003. ↓7, 70, 81

[16] N. J. Beaudry and R. Renner. An intuitive proof of the data processing inequality. *arXiv preprint arXiv:1107.0740*, 2011. ↓6, 102

[17] D. Beltita. *Smooth homogeneous structures in operator theory*. CRC Press, 2005. ↓12, 34

[18] D. Beltita and T. S. Ratiu. Symplectic leaves in real Banach Lie-Poisson spaces. *Geometric and Functional Analysis*, 15:753 – 779, 2005. ↓89

[19] I. Bengtsson and K. Życzkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, New York, 2006. DOI: 10.1017/cbo9780511535048. ↓3

[20] R. Bhatia. *Positive Definite Matrices*. Princeton University Press, 2007. ↓6

[21] P.K. Bhattacharya and Prabir Burman. 4 - basic concepts of statistical inference. In P.K. Bhattacharya and Prabir Burman, editors, *Theory and Methods of Statistics*, pages 69–88. Academic Press, 2016. ↓50

[22] P. Billingsley. *Probability and measure*. John Wiley & Sons, 2017. ↓68

[23] B. Blackadar. *Operator Algebras: Theory of C^*-algebras and von Neumann Algebras*. Springer-Verlag, Berlin, 2006. ↓11, 18

[24] V. Bogachev. *Measure Theory Volume I*. Springer, Berlin, 2007. ↓54

[25] M. Born. Quantenmechanik der stoßvorgänge. *Zeitschrift für physik*, 38(11-12):803–827, 1926. ↓3

[26] O. Bratteli and D. W. Robinson. *Operator Algebras and Quantum Statistical Mechanics I*. Springer-Verlag, Berlin, second edition, 1987. DOI: 10.1007/978-3-662-03444-6. ↓11, 16, 17, 19, 23

[27] S. L. Braunstein and C. M. Caves. Statistical distance and the geometry of quantum states. *Physical Review Letters*, 72(22):3439 – 3443, 1994. ↓82

[28] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer, 2011. ↓16, 20, 27

[29] L. Brillouin. *Science and information theory*. Courier Corporation, 2013. ↓6, 102

[30] D. Bures. An Extension of Kakutani's Theorem on Infinite Product Measures to the Tensor Product of Semifinite W*-Algebras. *Transactions of the American Mathematical Society*, 135:199–212, 1969. DOI: 10.2307/1995012. ↓39

[31] J. F. Cariñena, J. Clemente-Gallardo, J. A. Jover-Galtier, and G. Marmo. Tensorial dynamics on the space of quantum states. *Journal of Physics A*, 50(36):365301–30, 2017. DOI: 10.1088/1751-8121/aa8182. ↓41, 42

[32] J. F. Cariñena, J. Clemente-Gallardo, and G. Marmo. Geometrization of quantum mechanics. *Theoretical and Mathematical Physics*, 152(1):894–903, 2007. DOI: 10.1007/s11232-007-0075-3. ↓41, 42

[33] P. Castiglione, M. Falcioni, A. Lesne, and A. Vulpiani. Chaos and coarse graining in statistical mechanics. *Chaos and Coarse Graining in Statistical Mechanics*, 2008. ↓82

[34] N. N. Cencov. *Statistical Decision Rules and Optimal Inference*. American Mathematical Society, Providence, RI, 1982. ↓1, 6, 50, 64, 102

[35] M. Choi. A schwarz inequality for positive linear maps on c^*-algebras. *Illinois Journal of Mathematics*, 18(4):565–574, 1974. ↓23

[36] M. Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975. DOI: 10.1016/0024-3795(75)90075-0. ↓23

[37] D. Chruściński, F. M. Ciaglia, A. Ibort, G. Marmo, and F. Ventriglia. Stratified manifold of quantum states, actions of the complex special linear group. *Annals of Physics*, 400:221–245, 2019. DOI: 10.1016/j.aop.2018.11.015. ↓12, 41, 42

[38] D. Chruściński and G. Marmo. Remarks on the GNS Representation and the Geometry of Quantum States. *Open Systems & Information Dynamics*, 16(2 and 3):155 – 177, 2009. ↓11, 21

[39] C. Chu. *Jordan Structures in Geometry and Analysis*. Cambridge University press, Cambridge, UK, 2012. ISBN: 9781107016170. ↓12, 31, 33, 35

[40] F. M. Ciaglia. Quantum states, groups and monotone metric tensors. *European Physical Journal Plus*, 135:530–16pp, 2020. DOI: 10.1140/epjp/s13360-020-00537-y. ↓12

[41] F. M. Ciaglia, H. Cruz, and G. Marmo. Contact manifolds and dissipation, classical and quantum. *Annals of Physics*, 398:159–179, 2018. DOI: 10.1016/j.aop.2018.09.012. ↓41, 42

[42] F. M. Ciaglia, F. Di Cosmo, F. Di Nocera, and P. Vitale. Monotone metric tensors in quantum information geometry. *arXiv preprint arXiv:2203.10857*, 2022. ↓v, 19, 30, 85, 87, 89, 90, 97

[43] F. M. Ciaglia, F. Di Cosmo, A. Ibort, M. Laudato, and G. Marmo. Dynamical vector fields on the manifold of quantum states. *Open Systems & Information dynamics*, 24(3):1740003–38, 2017. DOI: 10.1142/S1230161217400030. ↓29, 41, 42, 99

[44] F. M. Ciaglia, F. Di Cosmo, A. Ibort, and G. Marmo. Dynamical aspects in the quantizer-dequantizer formalism. *Annals of Physics*, 385:769–781, 2017. DOI: 10.1016/j.aop.2017.08.025. ↓41, 42

[45] F. M. Ciaglia, F. Di Cosmo, M. Laudato, and G. Marmo. Differential Calculus on Manifolds with Boundary. Applications. *International Journal of Geometrical Methods in Modern Physics*, 14(8):1740003–39, 2017. DOI: 10.1142/s0219887817400035. ↓41, 42

[46] F. M. Ciaglia, F. Di Cosmo, M. Laudato, G. Marmo, G. Mele, F. Ventriglia, and P. Vitale. A Pedagogical Intrinsic Approach to Relative Entropies as Potential Functions of Quantum Metrics: the q-z family. *Annals of Physics*, 395:238–274, 2018. DOI: 10.1016/j.aop.2018.05.015. ↓19, 30, 78, 85, 87, 89, 97, 104

[47] F. M. Ciaglia and F. Di Nocera. Group Actions and Monotone Metric Tensors: The Qubit Case. In Frank Nielsen and Frédéric Barbaresco, editors, *Geometric Science of Information 2021*, volume 12829 of *Lecture Notes in Computer Science*, pages 145–153. Springer International Publishing, 2021. DOI: 10.1007/978-3-030-80209-717. ↓42

[48] F. M. Ciaglia and F. Di Nocera. Group actions and monotone quantum metric tensors. *Mathematics*, 10(15):2613, 2022. ↓42

[49] F. M. Ciaglia, F. Di Nocera, J. Jost, and L. Schwachhöfer. Parametric models and information geometry on w*-algebras. *Information Geometry*, pages 1–26, 2023. ↓v, 49

[50] F. M. Ciaglia, A. Ibort, J. Jost, and G. Marmo. Manifolds of classical probability distributions and quantum density operators in infinite dimensions. *Information Geometry*, 2(2):231–271, 2019. DOI: 10.1007/s41884-019-00022-1. ↓12, 34, 36, 37, 38

[51] F. M. Ciaglia, A. Ibort, and G. Marmo. Geometrical structures for classical and quantum probability spaces. *International Journal of Quantum Information*, 15(8):1740007–14, 2017. ↓41, 42

[52] F. M. Ciaglia, J. Jost, and L. Schwachhöfer. Differential geometric aspects of parametric estimation theory for states on finite-dimensional C*-algebras. *Entropy*, 22(11):1332, 2020. DOI: 10.3390/e22111332. ↓80

[53] F. M. Ciaglia, J. Jost, and L. Schwachhöfer. From the Jordan product to Riemannian geometries on classical and quantum states. *Entropy*, 22(06):637–27, 2020. DOI: 10.3390/e22060637. ↓8, 12, 26, 39, 41, 43, 44

[54] F. M. Ciaglia, J. Jost, and L. Schwachhöfer. Information geometry, jordan algebras, and a coadjoint orbit-like construction. arxiv. *arXiv preprint arXiv:2112.09781*, 1, 2021. ↓88, 89

[55] R. Cirelli, A. Mania, and L. Pizzocchero. Quantum mechanics as an infinite dimensional Hamiltonian system with uncertainty structure. *Journal of Mathematical Physics*, 31(12):2891 – 2903, 1990. ↓96

[56] C. Cohen-Tannoudji, B. Diu, and F. Laloe. *Quantum Mechanics, Volume 1*, volume 1. 1986. ↓70

[57] A. Connes. Classification of injective factors. cases ii_1, ii_∞, iii_λ, $\lambda \neq 1$. *Annals of Mathematics*, 104(1):73–115, 1976. ↓37

[58] A. Connes and E. Stormer. Homogeneity of the State Space of Factors of Type III_1. *Journal of Functional Analysis*, 28:187 – 196, 1978. ↓39

[59] T. M. Cover. *Elements of information theory.* John Wiley & Sons, 1999. ↓6, 102

[60] D. R. Cox and D. V. Hinkley. *Theoretical statistics.* CRC Press, 1979. ↓67

[61] H. S. M. Coxeter. *Introduction to geometry.* John Wiley & Sons, Inc., 1969. ↓2

[62] H. Cramér. *Mathematical methods of statistics*, volume 43. Princeton university press, 1999. ↓7, 67, 69

[63] F. D'Andrea and D. Franco. On the pseudo-manifold of quantum states. *Differential Geometry and its Applications*, 78:101800, 2021. DOI: 10.1016/j.difgeo.2021.101800. ↓88

[64] G. Darmois. Sur les limites de la dispersion de certaines estimations. *Revue de l'Institut International de Statistique / Review of the International Statistical Institute*, 13(1/4):9–15, 1945. ↓7

[65] K. R. Davidson. *C^*-algebras by example.* American Mathematical Society, Providence, RI, 1996. ↓11

[66] C. Davisson and L. H. Germer. Diffraction of electrons by a crystal of nickel. *Physical review*, 30(6):705, 1927. ↓72

[67] C. Davisson and L. H. Germer. The scattering of electrons by a single crystal of nickel. *Nature*, 119(2998):558–560, 1927. ↓72

[68] C. J. Davisson and L. H. Germer. Reflection and refraction of electrons by a crystal of nickel. *Proceedings of the National Academy of Sciences*, 14(8):619–627, 1928. ↓72

[69] L. De Broglie. *Recherches sur la théorie des quanta*. PhD thesis, Migration-université en cours d'affectation, 1924. ↓72

[70] M. H. DeGroot. *Optimal statistical decisions*. John Wiley & Sons, 2005. ↓68

[71] F. Di Nocera. Unfolding of relative g-entropies and monotone metrics. In *Physical Sciences Forum*, volume 5, page 34. MDPI, 2022. ↓v, 89

[72] J. Dittmann. On the Riemannian Geometry of Finite Dimensional Mixed States. *Seminar Sophus Lie*, 3:73–87, 1993. Web source. ↓47

[73] J. Dittmann. On the Riemannian metric on the space of density matrices. *Reports on Mathematical Physics*, 36(3):309–315, 1995. DOI: 10.1016/0034-4877(96)83627-5. ↓47

[74] J. Dixmier. *C^*-algebras*. North Holland, Amsterdam, 1977. ↓11, 16

[75] N. Dunford and J. T. Schwarz. Linear operators i (pure and appl. math. 7). *Inter-science Publishers, New York*, 1958. ↓22

[76] H. A. Dye. The radon-nikodỳm theorem for finite rings of operators. *Transactions of the American Mathematical Society*, 72(2):243–280, 1952. ↓49

[77] E. Ercolessi, G. Marmo, and G. Morandi. From the equations of motion to the canonical commutation relations. *Rivista del Nuovo Cimento*, 33:401–590, 2010. DOI: 10.1393/ncr/i2010-10057-x. ↓46, 96

[78] R. A. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A*, 222:309 – 368, 1922. ↓1, 5, 39, 69

[79] M. Fréchet. Sur l'extension de certaines evaluations statistiques au cas de petits echantillons. *Revue de l'Institut International de Statistique / Review of the International Statistical Institute*, 11(3/4):182–205, 1943. ↓7

[80] B. E. Fristedt and L. F. Gray. *A modern approach to probability theory*. Springer Science & Business Media, 2013. ↓68

[81] T. Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. *Advances in Mathematics*, 370:107239, 2020. ↓5

[82] T. Fritz, T. Gonda, P. Perrone, and E. F. Rischel. Representable Markov Categories and Comparison of Statistical Experiments in Categorical Probability. *arXiv math-st: 2010.07416*, 2020. ↓5

[83] T. Fritz and P. Perrone. A probability monad as the colimit of spaces of finite samples. *arXiv preprint arXiv:1712.05363*, 2017. ↓5

[84] G. Fubini. *Sulle metriche definite da una forma hermitiana: nota*. Office graf. C. Ferrari, 1904. ↓39, 46, 92

[85] A. Fujiwara. Geometry of quantum information systems. *Geometry in Present Day Science*, pages 35–48, 1999. ↓93

[86] P. A Gagniuc. *Markov chains: from theory to implementation and experimentation*. John Wiley & Sons, 2017. ↓4

[87] L. Gårding. Note on continuous representations of lie groups. *Proceedings of the National Academy of Sciences*, 33(11):331–332, 1947. ↓93

[88] S. Geisser and W. O. Johnson. *Modes of parametric statistical inference*. John Wiley & Sons, 2006. ↓50

[89] I. M. Gelfand and M. A. Neumark. On the imbedding of normed rings into the ring of operators in hilbert space. *Mat. Sbornik*, 12(54):197–217, 1943. ↓13, 26

[90] W. Gerlach and O. Stern. Das magnetische moment des silberatoms. *Zeitschrift für Physik*, 9(1):353–355, 1922. ↓70

[91] P. Gibilisco and T. Isola. A characterization of Wigner-Yanase skew information among statistically monotone metrics. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 4(4):553 – 557, 2001. ↓6, 101

[92] P. Gibilisco and T. Isola. Wigner-Yanase information on quantum state space: the geometric approach. *Journal of Mathematical Physics*, 44(9):3752 – 3762, 2003. ↓6, 101

[93] J. P. Gordon. Quantum effects in communications systems. *Proceedings of the IRE*, 50(9):1898–1908, 1962. ↓1

[94] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan. Completely positive dynamical semigroups of N-level systems. *Journal of Mathematical Physics*, 17(5):821 – 825, 1976. ↓42, 99

[95] J. Grabowski, M. Kuś, and G. Marmo. Geometry of quantum systems: density states and entanglement. *Journal of Physics A: Mathematical and General*, 38(47):10217–10244, 2005. DOI: 10.1088/0305-4470/38/47/011. ↓20, 29

[96] J. Grabowski, M. Kuś, and G. Marmo. Symmetries, group actions, and entanglement. *Open Systems & Information Dynamics*, 13(04):343 – 362, 2006. ↓12, 29

[97] R. Haag. *Local quantum physics: Fields, particles, algebras*. Springer-Verlag, Berlin, 1996. ↓39

[98] B. C. Hall. Lie groups, lie algebras, and representations. In *Quantum Theory for Mathematicians*, pages 333–366. Springer, 2013. ↓39, 40

[99] H. Halvorson and M. Müger. Algebraic quantum field theory. *arXiv preprint math-ph/0602036*, 2006. ↓39

[100] L. A. Harris and W. Kaup. Linear algebraic groups in infinite dimensions. *Illinois Journal of Mathematics*, 21(3):666 – 674, 1977. ↓36

[101] R. V. L. Hartley. Transmission of information 1. *Bell System technical journal*, 7(3):535–563, 1928. ↓1

[102] H. Hasegawa. Dual geometry of the Wigner-Yanase-Dyson information content. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 6(3):413 – 430, 2003. ↓6, 101

[103] C. W. Helstrom. Minimum mean-squared error of estimates in quantum statistics. *Physics Letters A*, 25(2):101–102, 1967. DOI: 10.1016/0375-9601(67)90366-0. ↓7, 46, 70, 81

[104] C. W. Helstrom. The minimum variance of estimates in quantum signal detection. *IEEE Transactions on Information Theory*, 14(2):234–242, 1968. DOI: 10.1109/TIT.1968.1054108. ↓7, 70, 81

[105] C. W. Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1(2):231–252, 1969. DOI: 10.1007/BF01007479. ↓1, 7, 70, 81

[106] H. Hendriks. A Cramer-Rao Type Lower Bound for Estimators with Values in a Manifold. *Journal of Multivariate Analysis*, 38(2):245 – 261, 1991. ↓77, 80, 81

[107] A. S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973. ↓1

[108] A. S. Holevo. On capacity of a quantum communications channel. *Problemy Peredachi Informatsii*, 15(4):3–11, 1979. ↓1

[109] A. S. Holevo. *Statistical Structure of Quantum Theory*. Springer-Verlag, Berlin, 2001. ↓7, 70, 81, 93

[110] A. S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. Edizioni della Normale, 2011. ↓7, 70, 81, 93

[111] N. Jacobson. Lie and jordan triple systems. In *Nathan Jacobson Collected Mathematical Papers: Volume 2 (1947–1965)*, pages 17–38. Springer, 1949. ↓103

[112] P. Jordan. Über die multiplikation quantenmechanischer größen. *Zeitschrift für Physik*, 80(5-6):285–291, 1933. ↓12

[113] P. Jordan, J. von Neumann, and E. P. Wigner. On an algebraic generalization of the quantum mechanical formalism. *Annals of Mathematics*, 35:29–64, 1934. DOI: 10.1007/978-3-662-02781-321. ↓12

[114] L. P. Kadanoff. *Statistical physics: statics, dynamics and renormalization*. World Scientific, 2000. ↓82

[115] R. V. Kadison. A generalized schwarz inequality and algebraic invariants for operator algebras. *Annals of Mathematics*, pages 494–503, 1952. ↓23

[116] R. V. Kadison and J. R. Ringrose. *Fundamentals of the Theory of Operator Algebras Volume I: Elementary Theory*. Academic Press, New York, 1983. ↓11

[117] R. V. Kadison and J. R. Ringrose. *Fundamentals of the Theory of Operator Algebras Volume II: Advanced Theory*. Academic Press, New York, 1986. ↓11

[118] T. W. B. Kibble. Geometrization of Quantum Mechanics. *Communications in Mathematical Physics*, 65(2):189 – 201, 1979. ↓96

[119] A. A. Kirillov. Unitary representations of nilpotent lie groups. *Russian Mathematical Surveys*, 17(4):53–104, 1962. DOI: 10.1070/RM1962v017n04ABEH004118. ↓8, 41, 87

[120] A. A. Kirillov. *Elements of the Theory of Representations*. Springer-Verlag Berlin Heidelberg, 1976. DOI: 10.1007/978-3-642-66243-0. ↓8, 41, 87

[121] A. Klenke. *Probability theory: a comprehensive course*. Springer Science & Business Media, 2013. ↓4

[122] B. O. Koopman. Hamiltonian Systems and Transformations in Hilbert Space. *Proceedings of the National Academy of Sciences*, 17(5):315 – 318, 1931. ↓15

[123] B. Kostant. Quantization and unitary representations. In *Lectures in Modern Analysis and Applications III*, volume 170 of *Lecture Notes in Mathematics*, pages 87–208. Springer, Berlin, Heidelberg, 1970. DOI: 10.1007/BFb0079068. ↓8, 41, 87

[124] R. P. Kostecki. Local quantum information dynamics. *arXiv preprint arXiv:1605.02063*, 2016. ↓102

[125] S. Lang. *Fundamentals of Differential Geometry*. Springer-Verlag, Berlin, 1999. ↓31

[126] S. Lauritzen. Statistical Manifolds. In S. I. Amari, O. E. Barndorff-Nielsen, R. E. Kass, S. L. Lauritzen, and C. R. Rao, editors, *Differential geometry in statistical inference*, volume 10, pages 163–216. Institute of Mathematical Statistics, 1987. DOI: 10.1214/lnms/1215467061. ↓102

[127] A. Lesniewski and M. B. Ruskai. Monotone riemannian metrics and relative entropy on noncommutative probability spaces. *Journal of Mathematical Physics*, 40(11):5702 – 5724, 1999. ↓104

[128] E. H. Lieb and M. B. Ruskai. Proof of the strong subadditivity of quantum-mechanical entropy. *Les rencontres physiciens-mathématiciens de Strasbourg-RCP25*, 19:36–55, 1973. ↓6

[129] G. Lindblad. On the Generators of Quantum Dynamical Semigroups. *Communications in Mathematical Physics*, 48:119 – 130, 1976. ↓42

[130] P. C. Mahalanobis. On the generalized distance in Statistics. *Proceedings of the National Institute of Sciences of India*, II(1):49–55, 1936. ↓1

[131] G. Maltese and G. Niestegge. A linear Radon-Nikodym type theorem for C^*-algebraswith applications to measure theory. *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 4*, 14(2):345–354, 1987. ↓49, 55

[132] V. I. Man'ko, G. Marmo, F. Ventriglia, and P. Vitale. Metric on the space of quantum states from relative entropy. Tomographic reconstruction. *Journal of Physics A: Mathematical and Theorerical*, 50(33):335302, 2017. DOI: 10.1088/1751-8121/aa7d7d. ↓6, 19, 30, 85, 89, 101, 104

[133] E. A. Morozowa and N. N. Cencov. Markov invariant geometry on state manifolds. *Journal of Soviet Mathematics*, 56(5):2648–2669, 1991. DOI: 10.1007/BF01095975. ↓2, 6

[134] C. N. Morris. Parametric empirical bayes inference: theory and applications. *Journal of the American statistical Association*, 78(381):47–55, 1983. ↓50

[135] F. J. Murray and J. v. Neumann. On rings of operators. *Annals of Mathematics*, 37(1):116–229, 1936. ↓12, 39

[136] F. J. Murray and J. von Neumann. On rings of operators. ii. *Transactions of the American Mathematical Society*, 41(2):208–248, 1937. ↓12

[137] F. J. Murray and J. von Neumann. On rings of operators. iv. *Annals of Mathematics*, 44(4):716–808, 1943. ↓12

[138] J. Naudts, A. Verbeure, and R. Weder. Linear Response Theory and the KMS Condition. *Communications in Mathematical Physics*, 44:87–99, 1975. DOI: 10.1007/BF01609060. ↓6, 101

[139] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, New York, NY, 2011. ↓1, 3, 5, 20, 60, 70

[140] G. Niestegge. Absolute continuity for linear forms on B^*-algebras and a Radon-Nikodym type theorem (quadratic version). *Rendiconti del Circolo Matematico di Palermo*, 32(2):358–376, 1983. ↓49, 54, 55

[141] O. Nikodym. Sur une généralisation des intégrales de mj radon. *Fundamenta Mathematicae*, 15(1):131–179, 1930. ↓27, 54

[142] H. Nyquist. Certain factors affecting telegraph speed. *Transactions of the American Institute of Electrical Engineers*, 43:412–422, 1924. ↓1

[143] M. G. A. Paris. Quantum Estimation for Quantum Technology. *International Journal of Quantum Information*, 7(1):125–137, 2009. DOI: 10.1142/S0219749909004839. ↓7, 46, 70, 81

[144] A. J. Parzygnat. Inverses, disintegrations, and Bayesian inversion in quantum Markov categories. *arXiv quant-ph: 2001.08375*, 2020. ↓5

[145] A. Perelomov. *Generalized Coherent States and Their Applications*. Springer-Verlag Berlin Heidelberg, 1986. ↓96

[146] A. Peres. *Quantum theory: concepts and methods*. Springer, 1997. ↓70

[147] D. Petz. Quasi-entropies for States of a von Neumann Algebra. *Publications of the RIMS, Kyoto University*, 21:787–800, 1985. DOI: 10.2977/prims/1195178929. ↓104

[148] D. Petz. Geometry of canonical correlation on the state space of a quantum system. *Journal of Mathematical Physics*, 35(2):780–795, 1994. 10.1063/1.530611. ↓6, 101

[149] D. Petz. Monotone metrics on matrix spaces. *Linear Algebra and its Applications*, 244:81–96, 1996. DOI: 10.1016/0024-3795(94)00211-8. ↓2, 6, 50, 64, 101, 102

[150] D. Petz. *Quantum information theory and quantum statistics*. Springer, Berlin, 2007. ↓3, 20

[151] D. Petz and C. Sudar. Geometries of Quantum States. *Journal of Mathematical Physics*, 37, 1996. ↓2, 6, 64, 101

[152] G. Pistone. Nonparametric information geometry. In *International Conference on Geometric Science of Information*, pages 5–36. Springer, 2013. ↓2

[153] G. Pistone and C. Sempi. An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one. *The Annals of Statistics*, 23(5):1543 – 1561, 1995. ↓2, 7, 97

[154] J. Radon. *Theorie und Anwendungen der absolut additiven Mengenfunktionen*. Hölder, 1913. ↓27, 54

[155] A. Raju, B. B. Machta, and J. P. Sethna. Information loss under coarse graining: A geometric approach. *Physical Review E*, 98(5):052112, 2018. ↓82

[156] C. R. Rao. Information and accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37(3):81–91, 1945. ↓1, 5, 7, 69

[157] S. Sakai. A characterization of W^*-algebras. *Pacific Journal of Mathematics*, 6(4):763 – 773, 1956. ↓13, 29

[158] S. Sakai. A radon-nikodym theorem in wˆ*-algebras. 1965. ↓49

[159] S. Sakai. *C*-algebras and W*-algebras*. Springer Science & Business Media, 2012. ↓11, 27, 28, 29, 53, 55

[160] B. Schumacher and M. A. Nielsen. Quantum data processing and error correction. *Physical Review A*, 54(4):2629, 1996. ↓60

[161] I. E. Segal. Postulates for General Quantum Mechanics. *nnals of Mathematics*, 48(4):930 – 948, 1947. ↓13

[162] I. E. Segal. A non-commutative extension of abstract integration. *Annals of mathematics*, pages 401–457, 1953. ↓49

[163] C. E. Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948. ↓1, 60, 103

[164] L. R. Shenton. The so-called cramer-rao inequality. *The American Statistician*, (2):35–39. ↓7

[165] J. M. Souriau. *Structure des systèmes dynamiques*. Dunod, Paris, 1970. Web source. ↓8, 41, 87

[166] W. F. Stinespring. Positive functions on C^*-algebras. *Proceedings of the American Mathematical Society*, 6:211–216, 1955. DOI: 10.2307/2032342. ↓23

[167] E. Study. Kürzeste wege im komplexen gebiet. *Mathematische Annalen*, 60(3):321–378, 1905. ↓39, 46, 92

[168] P. K. Suetin, A. I. Kostrikin, and Y. I. Manin. *Linear algebra and geometry*. CRC Press, 1989. ↓14

[169] J. Suzuki. Information Geometrical Characterization of Quantum Statistical Models in Quantum Estimation Theory. *Entropy*, 21(7):703, 2019. DOI: 10.3390/e21070703. ↓46

[170] M. Takesaki. *Theory of Operator Algebras I*. Encyclopaedia of Mathematical Sciences. Springer Berlin Heidelberg, 2001. ↓11, 16, 17, 18, 22, 24, 26, 34

[171] M. Takesaki. *Theory of Operator Algebra I*. Springer-Verlag, Berlin, 2002. ↓28

[172] M. Tomamichel. *Quantum Information Processing with Finite Resources Mathematical Foundations*. Springer International Publishing, 2016. DOI: 10.1007/978-3-319-21891-5. ↓3, 20

[173] W. K. Tung. *Group theory in physics*, volume 1. World Scientific, 1985. ↓26

[174] A. Uhlmann. The transition probability in the state space of a *-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976. DOI: 10.1016/0034-4877(76)90060-4. ↓39

[175] A. Uhlmann. Parallel transport and "quantum holonomy" along density operators. *Reports on Mathematical Physics*, 24(2):229–240, 1986. DOI: 10.1016/0034-4877(86)90055-8. ↓39

[176] A. Uhlmann. *Groups and related topics*, chapter The Metric of Bures and the Geometric Phase, pages 267–274. Springer, Dordrecht, 1992. DOI: 10.1007/978-94-011-2801-823. ↓39, 47

[177] A. Uhlmann. Transition Probability (Fidelity) and Its Relatives. *Foundations of Physics*, 41(3):288–298, 2011. DOI: 10.1007/s10701-009-9381-y. ↓39

[178] H. Upmeier. *Symmetric Banach manifolds and Jordan C^*-algebras*. Elsevier, Amsterdam, 1985. ISBN: 978-0-444-87651-5. ↓12, 31, 32, 33, 34, 36, 38

[179] J. v. Neumann. Zur algebra der funktionaloperationen und theorie der normalen operatoren. *Mathematische Annalen*, 102(1):370–427, 1930. ↓12, 13

[180] J. v. Neumann. On rings of operators. iii. *Annals of Mathematics*, 41(1):94–161, 1940. ↓12

[181] A. Van Daele. A radon nikodỳm theorem for weights on von neumann algebras. *Pacific Journal of Mathematics*, 61(2):527–542, 1975. ↓49

[182] L. Wasserman. *All of statistics: a concise course in statistical inference*, volume 26. Springer, 2004. ↓68

[183] M. M. Wilde. *Quantum information theory.* Cambridge university press, 2013. ↓5

[184] J. Yngvason. The role of type iii factors in quantum field theory. *Reports on Mathematical Physics*, 55(1):135–147, 2005. ↓39

Bibliographische Daten

A $W^\star$-algebraic formalism for parametric models in Classical and Quantum Information Geometry
(Ein $W^\star$-algebraischer Formalismus für parametrische Modelle in der klassischen und Quanteninformationsgeometrie)
Di Nocera, Fabio
Universität Leipzig, Dissertation, 2023
131 Seiten, 4 Abbildungen, 184 Referenzen

Selbstständigkeitserklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig und ohne unzulässige fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die angeführten Quellen und Hilfsmittel benutzt und sämtliche Textstellen, die wörtlich oder singemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialien oder erbrachten Dienstleistungen als solche gekennzeichnet.

Leipzig, den November 2, 2023

. . . .
(Fabio Di Nocera)

Daten zum Autor

Name:	Fabio Di Nocera
Geburtsdatum:	19.09.1992 in Gragnano (Italien)
10/2011 - 05/2015	B.Sc. in Physik Università degli Studi di Napoli "Federico II"
10/2015 - 10/2019	M.Sc. in Physik Università degli Studi di Napoli "Federico II"
seit 09/2020	Doktorand der Mathematik Max Planck Institut für Mathematik in den Naturwissenschaften