UNIVERSITÀ DEGLI STUDI DI NAPOLI
"FEDERICO II"

Scuola Politecnica e delle Scienze di Base

Area Didattica di Scienze Matematiche Fisiche e Naturali

Dipartimento di Fisica "Ettore Pancini"

Laurea Magistrale in Fisica

Information Geometry and Quantum Mechanics

Relatori:
Prof.ssa Vitale Patrizia
Prof. Marmo Giuseppe

Candidato:
Di Nocera Fabio
Matr. $N94000354$

Anno Accademico 2018/2019

Contents

Introduction

Information Geometry is the study of the theory of information using the tools of modern geometry, and it has given many contributions to information theory, due to the works of Chentsov [10] and Amari [4] among others, and is still a theory of great interest for its application to modern problems [3] [31].
This works tries to investigate if there is a possibility of cross-fertilization between this theory and Quantum Mechanics, this idea originates from the following observation: Fisher-Rao metric can be seen as a term of Fubini-Study metric.
Let us justify this statement: Fisher-Rao [37] metric has a privileged role in Information Geometry due to Chentsov theorem [10], it is usually obtained from Shannon relative entropy [39] with an algorithm that we will review in chapter 3, and has the following form:

$$g_{FR} = \sum_{j=1}^{n} q_j d\log q_j \otimes d\log q_j \tag{1}$$

where the q_j are the components of a probability vector.
Now let us write Fubini-Study metric [24] on a Hilbert space $\mathcal{H}$:

$$g_{FS} = \frac{\langle d\psi|d\psi\rangle}{\langle\psi|\psi\rangle} - \frac{\langle d\psi|\psi\rangle \otimes \langle\psi|d\psi\rangle}{\langle\psi|\psi\rangle^2} \tag{2}$$

Introducing an orthonormal basis $(|e_1\rangle, |e_2\rangle, \ldots, |e_n\rangle)$ in $\mathcal{H}$ (so clearly $\mathcal{H}$ has complex dimension n) we have:

$$|\psi\rangle = \sum_{j=1}^{n} z_j |e_j\rangle \qquad z_j \in \mathbb{C} \quad \forall\, j = 1, 2, \ldots, n \tag{3}$$

Then one can take the polar decomposition of the complex coefficients z_j:

$$z_j = \sqrt{p_j} e^{i\phi_j} \tag{4}$$

where $\sqrt{p_j}$ are real positive numbers and the ϕ_j are phases.
Now let us write an explicit expression for Fubini-Study metric, omitting the details of the calculation:

$$\begin{aligned} g_{FS} = \sum_{j,k=1}^{n} \frac{1}{4}(p_j d\log p_j \otimes d\log p_j - p_j p_k d\log p_j \otimes d\log p_k) \\ + p_j d\phi_j \otimes d\phi_j - p_j p_k d\phi_j \otimes d\phi_k \\ + \frac{i}{2}(p_j d\log p_j \wedge d\phi_k - p_j p_k d\log p_j \otimes d\phi_k) \end{aligned} \tag{5}$$

Now we can clearly recognise the first term of this expression as the Fisher-Rao metric, while the other terms that are present in (5) arise because of the phase present in (4).
Fisher-Rao metric is appropriate in Information Geometry, its exclusive use in the classical case is justified by Chentsov theorem [10], is used to obtain relevant results in this subject, like the existence of the Cramer-Rao bound [37] [2] and we have written it in terms of the probabilities p_j.
On the other hand Fubini-Study metric [24] is appropriate in Quantum Theory, and we have written it in terms of the amplitudes z_j. Formula (4) is a simple example of a shift from probabilities to amplitudes, and allows us to write Fubini-Study metric in terms of the components p_j of a probability vector, and when we do so we obtain, together with other terms depending on the phase, Fisher-Rao metric as a term of Fubini-Study metric.
This simple consideration strongly suggests that there may exist a deep link between Information Geometry and Quantum Theory, and that such a link can be seen from a perspective indicated by this transition from probabilities to amplitudes [23].
In mathematical terms, this transition may be translated into a bundle language, where probabilities constitute the base manifold while the total space is made of

probability amplitudes.

The work is structured as follows:

- In the first chapter is given a bundle picture for the description of (pure and mixed) quantum states [6], this description allows to discuss the procedure of lifting from probabilities to amplitudes in a geometric fashion. Moreover we will see how the problem of the purification of mixed states [44] can be formulated in this same setting;
- In the second chapter there is a description of further geometric structure that one can give in Quantum Mechanics [22], this is needed in order to better understand the role of quantum metrics in Quantum Mechanics, and how they may be obtained from relative entropies or their generalization;
- In the third chapter a brief review of Information Geometry is given [3], and are discussed some fundamental result of the application of Information Geometry to Classical Information Theory;
- In the fourth chapter we describe the difficulties and the problems that arise when trying to use the methods developed in chapter 3 in the quantum setting;
- In the fifth and final chapter, by using recent proposals [12–14, 28] we will argue, that the bundle picture may be considered as an overall description in terms of groupoids, which are the mathematical translation of Schwinger approach to quantum theory.

In some recent works [25–27] are derived Schrödinger-Robertson indetermination relations from Cramer-Rao inequalities, extending the possibility of application of Information Geometry to Quantum Metrology and to Foundations of Quantum Mechanics.

Chapter 1

A bundle picture for quantum states

In the standard approach to Quantum Mechanics [18] with every physical quantum mechanical system we associate an Hilbert space $\mathcal{H}$, and a pure state of the physical system is associated to a (non null) vector $|\psi\rangle$ of the Hilbert space.
On an Hilbert space $\mathcal{H}$ we have an Hermitian structure:

$$h : \mathcal{H} \times \mathcal{H} \ni (\psi, \phi) \mapsto \langle\psi|\phi\rangle \in \mathbb{C} \tag{1.1}$$

i.e. a positive-definite, non-degenerate form that is linear in the second argument and anti-linear in the first. It induces a norm on $\mathcal{H}$ given by:

$$\mathcal{H} \ni |\psi\rangle \mapsto \langle\psi|\psi\rangle \in \mathbb{R}^{+} \tag{1.2}$$

The Copenaghen interpretation of Quantum Mechanics consists of interpreting the quantity

$$P(\psi, \phi) = \frac{|\langle\psi|\phi\rangle|^{2}}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle} \tag{1.3}$$

as the transition probability between the states $|\psi\rangle$ and $|\phi\rangle$; we will call the quantity $\langle\psi|\phi\rangle$ transition amplitude between these two states. Then usually one defines the (normalised) expectation value of an observable A, that is an Hermitian operator on $\mathcal{H}$ over the state $|\psi\rangle$ in the following way:

$$e_A(\psi) = \frac{\langle\psi| A |\psi\rangle}{\langle\psi|\psi\rangle} \tag{1.4}$$

What is usually done is using only normalised vectors, that is restrict oneself in the space $\mathcal{S}(\mathcal{H})$ defined as the space of the vectors of $\mathcal{H}_0$ that have norm one:

$$\mathcal{S}(\mathcal{H}) = \left\{ |\psi\rangle \in \mathcal{H}_0 \quad s.t. \quad \langle\psi|\psi\rangle = 1 \right\} \tag{1.5}$$

And one can forget about the denominators in the quantities (1.3) and (1.4).

1.1 A bundle picture for pure states

Notice that neither the transition probabilities (1.3) nor the expectation values of the observables (1.4) are affected if we multiply these vectors by an overall phase or if we change the normalization of our vectors (or if we don't normalize them at all). Thus we find that all the information about the physical state encoded in the vector $|\psi\rangle$ is equally encoded in every vector that one could get from $|\psi\rangle$ by multiplying it by a real (non zero) positive number or by a pure phase, in this spirit we define the following equivalence relation:

$$|\psi\rangle \equiv \lambda\,|\psi\rangle \qquad \forall\lambda \in \mathbb{C}_0 \cong \mathbb{R}_+ \times U(1) \tag{1.6}$$

The equivalence classes defined in $\mathcal{H}_0$ by this equivalence relation will be called the rays of the Hilbert space $\mathcal{H}$, the space of all the rays will be called $\mathcal{P}(\mathcal{H})$, the complex projective space. Let us define a normalised rank-one projector:

$$\rho_\psi = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle} \tag{1.7}$$

This is clearly a projector:

$$\rho_\psi^2 = \rho_\psi \tag{1.8}$$

And it is rank-one since it projects a vector of $\mathcal{H}$ on a subspace of $\mathcal{H}$ that has (complex) dimension 1. Notice that these objects are invariant under multiplication by a real positive number and by a pure phase, and are in a one to one correspondence with the elements of $\mathcal{P}(\mathcal{H})$. Thus rank-one projectors are in a one to one correspondence with physically distinguishable pure states and they will be the mathematical objects that represent pure states in our formalism.

Remark 1. *It should be noticed that the space of rays does not depend on the Hermitian structure we use for the definition of Hilbert space, however the parametrization by means of rank-one projectors depends on the Hermitian product.*

We can define the following map:

$$\pi : |\psi\rangle \longmapsto \rho_\psi \tag{1.9}$$

This map is a projection map from vectors in $\mathcal{H}_0$ to rank-one projector that are in bijection with the elements of $\mathcal{P}(\mathcal{H})$. The geometric structure that emerges is a principal bundle with total space $\mathcal{H}_0$, base space $\mathcal{P}(\mathcal{H})$ and with fibers that are isomorphic to the 2-dimensional real Abelian Lie group $\mathbb{C}_0$. So we have the following structure:

$$\begin{array}{ccc} \mathbb{C}_0 & \longrightarrow & \mathcal{H}_0 \\ & & \downarrow \\ & & \mathcal{P}(\mathcal{H}) \end{array}$$

We could also first quotient w.r.t. the action of $\mathbb{R}_+$ and then w.r.t. the action of $U(1)$ or viceversa:

$$\begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathcal{H}_0 \\ & & \downarrow \\ U(1) & \longrightarrow & \mathcal{S}(\mathcal{H}) \\ & & \downarrow \\ & & \mathcal{P}(\mathcal{H}) \end{array}$$

We already said that quantities (1.3) and (1.4) are invariant under multiplication of vectors by a phase and under multiplication by a real (non zero) positive number (i.e. are constant along the fibers) so we can write in terms of rank-one projectors:

$$e_A(\psi) = \frac{\langle\psi|\, A\, |\psi\rangle}{\langle\psi|\psi\rangle} = \mathrm{Tr}\{\rho_\psi A\} \tag{1.10}$$

$$P(\psi,\phi) = \frac{|\langle\psi|\phi\rangle|^2}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle} = \mathrm{Tr}\{\rho_\psi\rho_\phi\} \tag{1.11}$$

Example 1. *A qubit is a two level quantum mechanical system, so the Hilbert space is $\mathbb{C}^2$. Let $(|e_1\rangle, |e_2\rangle)$ be a basis of $\mathcal{H}$, the generic vector can be written* *in that base:*

$$|\psi\rangle = z_1 |e_1\rangle + z_2 |e_2\rangle \tag{1.12}$$

$$z_j = x_j + iy_j \quad with \quad x_j, y_j \in \mathbb{R} \quad for \quad j = 1, 2 \tag{1.13}$$

Now we can impose the normalization condition:

$$x_1^2 + y_1^2 + x_2^2 + y_2^2 = 1 \tag{1.14}$$

This is just the equation of S^3 embedded in $\mathbb{R}^4 \cong \mathbb{C}^2$.
It is a well known result, due to Hopf [1], that S^3 can be seen as the total space of a principal bundle with base S^2 and structure group $\mathcal{U}(1)$.
So, going to the quotient with respect to the action of $\mathbb{R}_+$ takes us to S^3, then going to the quotient with respect to the action of $\mathcal{U}(1)$ takes us on S^2.

$$\begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathcal{H}_0 \cong \mathbb{C}_0^2 \\ & & \downarrow \\ U(1) & \longrightarrow & S^3 \\ & & \downarrow \\ & & \mathcal{P}(\mathbb{C}_0^2) \cong S^2 \end{array}$$

What we are left with is the boundary of what is usually called the Bloch sphere, we are going to "fill" the sphere when we will discuss mixed state.

We can consider the action of $GL(\mathcal{H})$ on $\mathcal{H}$:

$$GL(\mathcal{H}) \ni T : |\psi\rangle \mapsto T|\psi\rangle \tag{1.15}$$

This is clearly a linear action, consider now the projection via π of the transformed vector:

$$\pi(T|\psi\rangle) = \frac{T|\psi\rangle\langle\psi| T^\dagger}{\langle\psi| T^\dagger T |\psi\rangle} \tag{1.16}$$

So we can state the following remark:

Remark 2. *$GL(\mathcal{H})$ acts linearly on $\mathcal{H}$ and the action "descends" to $\mathcal{P}(\mathcal{H})$ with a non linear action*

Notice now that the projections cancels the effect of a multiplication by a complex non-zero number:

$$\pi(|zT\psi\rangle) = \pi(T|\psi\rangle) \quad \forall z \in \mathbb{C}_0 \tag{1.17}$$

Thus the effective action on $\mathcal{P}(\mathcal{H})$ is that of the quotient of $GL(\mathcal{H})$ with respect to his center:

$$\mathcal{Z} = \{Z \in GL(\mathcal{H}) : Z = z\mathbb{I} \quad z \in \mathbb{C}_0\} \tag{1.18}$$

Being the center a normal subgroup this quotient gives rise to a group, and this group is the *special linear group* $SL(\mathcal{H})$. Moreover, being $\mathcal{H}$ an orbit of $GL(\mathcal{H})$ we can state the following remark:

Remark 3. *$\mathcal{P}(\mathcal{H})$ is an orbit of the action of $SL(\mathcal{H})$ defined by relation* (1.16)

If we take an element U in the subgroup $\mathcal{U}(\mathcal{H})$ of $GL(\mathcal{H})$, we get:

$$\pi(U|\psi\rangle) = \frac{U|\psi\rangle\langle\psi|U^\dagger}{\langle\psi|U^\dagger U|\psi\rangle} = \frac{U|\psi\rangle\langle\psi|U^\dagger}{\langle\psi|\psi\rangle} = U\rho_\psi U^\dagger \tag{1.19}$$

That is just the coadjoint action of $U(\mathcal{H})$ on the space of Hermitian operators. Repeating the argument we exposed for $GL(\mathcal{H})$ one gets that the effective action on $\mathcal{P}(\mathcal{H})$ is the action of $\mathcal{SU}(\mathcal{H})$. It is a well known fact [33] that $U(\mathcal{H})$ acts transitively on $\mathcal{P}(\mathcal{H})$ and thus we have:

Remark 4. *$\mathcal{P}(\mathcal{H})$ is an orbit of the action of $U(\mathcal{H})$ defined by relation* (1.19)*, and it is also an orbit of SU(H) and of SL(H) at the same time.*

Now let us note that with rank-one projectors one can construct transition probability:

$$P(\phi, \psi) = \mathrm{Tr}\{\rho_\phi \rho_\psi\} \tag{1.20}$$

But one cannot construct amplitudes. While using elements of the total space, that is vectors in the Hilbert space $\mathcal{H}_0$, we can.
So in this setting the idea of going from probabilities to amplitudes means going from $\mathcal{P}(\mathcal{H})$ to $\mathcal{H}_0$, and this has an ambiguity given by an element of $\mathbb{C}_0$.
In this section we saw that the most natural setting for Quantum Mechanics is not the Hilbert space itself but rather the complex projective space. Nonetheless we need to recover the superposition principle in order to describe interference phenomena, and the previous observation suggests that rank-one projectors are not suitable to describe such phenomena.
Now we will first give an example that makes clear what is the problem with describing interference with rank-one projectors, and then give a procedure to overcome this difficulty.
So we want to consider a superposition between the states represented by two rank-one projectors ρ_1 and ρ_2:

$$\rho_1 = \frac{|\psi_1\rangle\langle\psi_1|}{\langle\psi_1|\psi_1\rangle} \qquad \rho_2 = \frac{|\psi_2\rangle\langle\psi_2|}{\langle\psi_2|\psi_2\rangle} \tag{1.21}$$

So what one can do is going back to the total space with the following section:

$$\sigma_0 : \rho_{\psi_j} \mapsto |\psi_j\rangle \qquad for \quad j = 1, 2 \tag{1.22}$$

Now clearly we can take a linear superposition of the two states $|\psi_1\rangle$ and $|\psi_2\rangle$:

$$|\psi\rangle = c_1 |\psi_1\rangle + c_2 |\psi_2\rangle \qquad |\psi_1\rangle , |\psi_2\rangle \in \mathcal{H}_0; \quad c_1, c_2 \in \mathbb{C}_0 \tag{1.23}$$

to obtain another vector of our Hilbert space.
Then consider the rank-one projector associated to $|\psi\rangle$:

$$\rho = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle} \tag{1.24}$$

Clearly we have that:

$$\rho = \frac{(c_1 |\psi_1\rangle + c_2 |\psi_2\rangle)(c_1^* \langle\psi_1| + c_2^* \langle\psi_2|)}{(c_1^* \langle\psi_1| + c_2^* \langle\psi_2|)(c_1 |\psi_1\rangle + c_2 |\psi_2\rangle}$$
$$= \frac{|c_1|^2\rho_1 + |c_2|^2\rho_2 + c_1c_2^* |\psi_1\rangle\langle\psi_2| + c_1^*c_2 |\psi_2\rangle\langle\psi_1|}{|c_1|^2 \langle\psi_1|\psi_1\rangle + |c_2|^2 \langle\psi_2|\psi_2\rangle + c_1^*c_2 \langle\psi_1|\psi_2\rangle + c_1c_2^* \langle\psi_2|\psi_1\rangle} \quad (1.25)$$

where the terms containing the two different vectors cannot be written in terms of the rank-one projectors ρ_1 and ρ_2.

So what we have done is going back to the total space with the section given by (1.22), then do the superposition on the total space and project back on the base space. The problem we encountered is that there is no simple way of making explicit the relation between the initial rank one projectors ρ_1 and ρ_2 and the result of this procedure.

Now, to overcome this problem [22], we can consider a fiducial vector $|\psi_0\rangle$, which is not orthogonal neither to $|\psi_1\rangle$ nor to $|\psi_2\rangle$. Let us define:

$$\rho_0 = \frac{|\psi_0\rangle\langle\psi_0|}{\langle\psi_0|\psi_0\rangle} \quad (1.26)$$

And let us introduce the following section:

$$\sigma : \rho_j \mapsto |\phi_j\rangle = \rho_j |\psi_0\rangle \qquad for \quad j = 1, 2 \quad (1.27)$$

Notice that this vector could be obtained multiplying $|\psi_j\rangle$ by a complex number, so they belong to the same fiber.

And now we take the linear superposition $|\phi\rangle = c_1 |\phi_1\rangle + c_2 |\phi_2\rangle$ and construct the rank-one projector:

$$\rho = \frac{|\phi\rangle\langle\phi|}{\langle\phi|\phi\rangle} \quad (1.28)$$

Clearly we have:

$$|\phi_j\rangle\langle\phi_k| = \rho_j\rho_0\rho_k \qquad \langle\phi_j|\phi_k\rangle = \mathrm{Tr}\,\rho_0\rho_j\rho_k \quad (1.29)$$

And then:

$$\rho = \sum_{j,k=1}^{2} \frac{c_j c_k^* \rho_j \rho_0 \rho_k}{c_j^* c_k \operatorname{Tr}\{\rho_0 \rho_j \rho_k\}} \tag{1.30}$$

This expression can be simplified once one considers that:

$$\frac{c_j c_j^* \rho_j \rho_0 \rho_j}{c_j^* c_j \operatorname{Tr}\{\rho_0 \rho_j \rho_j\}} = \frac{|\psi_j\rangle\langle\psi_j|\psi_0\rangle\langle\psi_0|\psi_j\rangle\langle\psi_j|}{\langle\psi_0|\psi_j\rangle\langle\psi_j|\psi_j\rangle\langle\psi_j|\psi_0\rangle} = \rho_j \tag{1.31}$$

So we have that (1.30) becomes:

$$\rho = \rho_1 + \rho_2 + \frac{c_1 c_2^* \rho_1 \rho_0 \rho_2}{c_1^* c_2 \operatorname{Tr}\{\rho_0 \rho_1 \rho_2\}} + \frac{c_1^* c_2 \rho_2 \rho_0 \rho_1}{c_2^* c_1 \operatorname{Tr}\{\rho_0 \rho_2 \rho_1\}} \tag{1.32}$$

The advantage of the latter procedure is that the result is written in terms of rank-one projectors. In this way, even if we constructed the result using the section (1.27), one can operate the superposition using formula (1.32) without having to explicitly choose a section.

1.2 A bundle picture for mixed states

Density matrices are used to describe mixed states, i.e. to describe statistical mixtures of different quantum states. A typical example is when we have to describe a beam of particles that is a mixture of m fractions of particles in different quantum states. In that case one writes:

$$\rho = \sum_{j=1}^{m} w_j \rho_{\psi_j} \tag{1.33}$$

Where ρ_{ψ_j} are rank-one projectors that represent the state of the j-th fraction of the beam and the coefficients w_j are the weights of the mixture. So the result will be a convex combination of the ρ_{ψ_j}, that is:

$$\sum_{j=1}^{n} w_j = 1; \qquad 0 \leqslant w_j \leqslant 1 \quad for \quad j = 1, 2, \ldots, m \tag{1.34}$$

It is easily seen that ρ is a semi-positive definite, Hermitian and trace one operator on $\mathcal{H}$. We will assume that $\mathcal{H}$ is finite dimensional (of complex dimension n) because in this way the space of linear operators over $\mathcal{H}$ coincides with the space of trace-class operators.

The space of semi-positive definite operators $\mathbf{P}(\mathcal{H})$ is a cone in the space of trace class operators $\mathcal{B}(\mathcal{H})$, while the trace one operators form an affine subspace, the intersection between this subspace and $\mathbf{P}(\mathcal{H})$ is the space of density state, called $\mathcal{D}(\mathcal{H})$. With $\mathbf{P}^k(\mathcal{H})$ and $\mathcal{D}^k(\mathcal{H})$ we will denote the spaces of positive operators or density states of rank k with $k = 1, 2, \ldots, n$.

It can be proven [29] that $\partial\mathcal{D}(\mathcal{H}) = \bigcup_{k=1}^{n-1} \mathcal{D}_k(\mathcal{H})$, this means that density states of non maximal rank are in the boundary of $\mathcal{D}(\mathcal{H})$, while the space of density states of maximal rank $\mathcal{D}_n(\mathcal{H})$ is the bulk of $\mathcal{D}(\mathcal{H})$, it also holds the following result:

Theorem 1. *The spaces $\mathcal{D}_k(\mathcal{H})$ are smooth and connected submanifolds of $\mathcal{B}(\mathcal{H})$ of (real) dimension $2nk - k^2 - 1$, while the whole $\mathcal{D}(\mathcal{H})$ and its boundary $\partial\mathcal{D}(\mathcal{H}) = \bigcup_{k=1}^{n-1} \mathcal{D}_k(\mathcal{H})$ are not smooth manifolds*

This theorem allows us to say that $\mathcal{D}(\mathcal{H})$ is a stratified manifold, with n strata,

given by the smooth manifolds $\mathcal{D}_k(\mathcal{H})$.
The spectral theorem [8] states that any Hermitian matrix can be diagonalized by a unitary matrix, and that the resulting diagonal matrix has only real entries i.e. it has real eigenvalues.
The semi-positive definiteness implies that the eigenvalues are all non-negative, imposing also the trace one condition one has that the sum of the eigenvalues (each taken with his algebraic multiplicity) have to be one.
So we have that every mixed state can be written in following form:

$$\rho = U \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} U^\dagger \tag{1.35}$$

With:

$$\sum_{j=1}^{n} \lambda_j = 1; \qquad 0 \leqslant \lambda_j \leqslant 1 \quad for \quad j = 1, 2, \dots, n \tag{1.36}$$

$$UU^\dagger = U^\dagger U = \mathbb{I} \tag{1.37}$$

So we see that every mixed state can be parameterized by giving a diagonal matrix satisfying the constraints (1.36) and an element of $U(n)$. Diagonal matrices of this kind can be put in a one to one correspondence with points in an $n-1$ dimensional simplex.
The vertices, or 0-faces, are associated to rank-one projectors, that is pure states, while the $k-1$-faces with $1 < k < n+1$ are rank k matrices, with the $(n-1)$-face (that is the bulk of the simplex) made of maximal rank states, that is invertible states. The barycenter of this simplex will be the so called *maximally mixed state*, that is the state proportional to the identity.
There is however an issue with this decomposition, one can swap two eigenvalues, while leaving unaltered all the others, with a unitary transformation. This means that for every ρ we have $n!$ different decompositions, one for every permutation

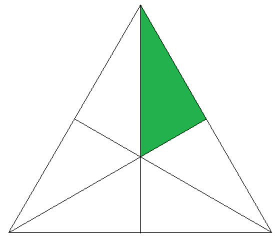


Figure 1.1: A pictorial representation of a bidimensional simplex, in green one of the Weyl chambers

of the eigenvalues.

So in order to uniquely decompose a density matrix via (1.35) we should not consider elements in the $n-1$ dimensional simplex but in its quotient with respect to the action of the permutation group of degree n, the result of this quotient is what is usually called the *Weyl chamber*, which can be then immersed in S_{n-1}, see figure 1.1.

Equation (1.35) also tells us that the coadjoint action of the unitary group $U(n)$ connects density matrices that have the same spectrum, a stronger result actually holds:

Remark 5. *The orbits of the coadjoint action of the unitary group $U(n)$ on $\mathcal{D}(\mathcal{H})$ are smooth submanifolds of $\mathcal{D}(\mathcal{H})$ that are made of operators that have the same spectrum*

Example 2. *Every Hermitian 2×2 matrix can be written in the following form:*

$$\rho = \alpha(\mathbb{I} + \mathbf{x} \cdot \boldsymbol{\sigma}) \qquad \forall \mathbf{x} \in \mathbb{R}^3 \qquad \forall \alpha \in \mathbb{R} \tag{1.38}$$

with $\boldsymbol{\sigma}$ being a (tri)vector whose components are the Pauli matrices. Being the Pauli matrices traceless, in order to have a trace one matrix we need to put $\alpha = 1/2$, let us write the eigenvalues of ρ:

$$\lambda_{\pm} = \frac{1}{2}(1 \pm |\mathbf{x}|) \tag{1.39}$$

So if we want ρ to be semi-positive definite we need $0 \leqslant |\mathbf{x}| \leqslant 1$. Every 2×2 density state can be decomposed in the following way:

$$\rho = U \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} U^\dagger \tag{1.40}$$

$$0 \leqslant \lambda_\pm \leqslant 1 \quad \lambda_+ + \lambda_- = 1 \quad U \in U(2) \tag{1.41}$$

The couples (λ_+, λ_-) can be seen as points in the monodimensional simplex, that is the segment in figure 1.2, and quotienting with respect to the action of the permutation group of degree 2 just means taking only half of it, this means taking $1/2 \leqslant \lambda_+ \leqslant 1$ and $0 \leqslant \lambda_- \leqslant 1/2$. Putting $\lambda_+ = 1$ and $\lambda_- = 0$ means that we are considering a pure state, and by remark 4 we know that every pure state is reached if we consider the coadjoint action of the unitary group on a pure state. But when $\lambda_+ = 1$ and $\lambda_- = 0$ then $|\mathbf{x}| = 1$, that is:

$$x_1^2 + x_2^2 + x_3^2 = 1 \tag{1.42}$$

That is just the equation of the sphere of radius 1 embedded in $\mathbb{R}^3$, so we can conclude that pure states in the qubit case are in a one to one correspondence with points on the surface of the sphere, this is just what we obtained on example 1. Now consider $1/2 < \lambda_+ < 1$ and $0 < \lambda_- < 1/2$, in this case we obtain spheres of radius $|\mathbf{x}|$ with $0 < |\mathbf{x}| < 1$, all these spheres share the same center, that is $\mathbf{x} = \mathbf{0}$. This point is the whole orbit of the coadjoint action of $U(n)$ on the diagonal matrix with: $\lambda_+ = \lambda_- = 1/2$, in fact, being this matrix proportional to the identity, this action is trivial on this point. So we get the following picture, the bulk of the sphere is the stratum of maximal rank (rank two) states, while the boundary is the stratum of pure states. States that belong to the same sphere share the same spectrum, the center of this sphere being the maximally mixed state.

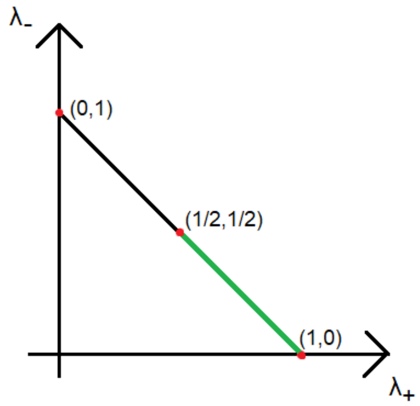


Figure 1.2: A pictorial representation of a monodimensional simplex, in green one of the two Weyl chambers

Now we will describe the principal bundle structure that emerges in this setting.
We can define the following projection map:

$$\pi_1 : \mathcal{B}(\mathcal{H}) \ni A \mapsto \rho = \frac{A^\dagger A}{\mathrm{Tr}\{A^\dagger A\}} \in \mathcal{D}(\mathcal{H}) \tag{1.43}$$

This map associates an operator in $\mathcal{B}(\mathcal{H})$ with an Hermitian, semi-positive definite and trace one operator, that is a mixed state. Clearly one can define another projection:

$$\pi_2 : \mathcal{B}(\mathcal{H}) \ni A \mapsto \rho = \frac{A A^\dagger}{\mathrm{Tr}\{A^\dagger A\}} \in \mathcal{D}(\mathcal{H}) \tag{1.44}$$

The operators at the numerator are already semi-positive definite and Hermitian, while the denominator fixes the trace to one. Notice that:

$$\pi_1(UA) = \pi_1(A) \tag{1.45}$$

$$\pi_1(\lambda A) = \pi_1(A) \tag{1.46}$$

$$\pi_2(AU) = \pi_2(A) \tag{1.47}$$

$$\pi_2(\lambda A) = \pi_2(A) \tag{1.48}$$

$$\forall U \in U(n) \qquad \forall \lambda \in \mathbb{R}_+ \tag{1.49}$$

So we have that π_1 maps all elements connected by the left action of $U(n)$ in the same positive operator, while π_2 maps all elements connected by the right action of $U(n)$ in the same positive operators. Both projections are not sensible to multiplication by a real number, due to their normalization.

In order to define a bundle, the inverse images of this projections have to be isomorphic [6]:

$$\pi_1^{-1}(A) \cong \pi_1^{-1}(A') \tag{1.50}$$

$$\pi_2^{-1}(A) \cong \pi_2^{-1}(A') \tag{1.51}$$

This can be done by restricting the maps to operators with trivial kernel (i.e. invertible operators), so we have:

$$\pi_1 : GL(n, \mathbb{C}) \ni A \mapsto \rho = \frac{A^\dagger A}{\mathrm{Tr}\{A^\dagger A\}} \in \mathcal{D}^n(\mathcal{H}) \tag{1.52}$$

$$\pi_2 : GL(n, \mathbb{C}) \ni A \mapsto \rho = \frac{AA^\dagger}{\mathrm{Tr}\{A^\dagger A\}} \in \mathcal{D}^n(\mathcal{H}) \tag{1.53}$$

And the fibers result all isomorphic to $U(n) \times \mathbb{R}_+$, so we end up with the following structure:

$$\begin{array}{ccc} U(n) & \longrightarrow & GL(n,\mathbb{C}) \\ & & \downarrow \\ \mathbb{R}_+ & \longrightarrow & \mathcal{P}_n(\mathcal{H}) \\ & & \downarrow \\ & & \mathcal{D}_n(\mathcal{H}) \end{array}$$

We can, like we did for the pure states, consider the right action of $GL(n,\mathbb{C})$ on the total space, in this case $GL(n,\mathbb{C})$ itself:

$$GL(n,\mathbb{C}) \ni T \mapsto AT \tag{1.54}$$

And this will induce a non-linear action on the base space:

$$\pi_1(AT) = \frac{T^\dagger A^\dagger AT}{\mathrm{Tr}\{T^\dagger A^\dagger AT\}} \tag{1.55}$$

The effective action is again that of $SL(n,\mathbb{C})$, as we showed in the case of pure states, and it can be shown with analog arguments. One can easily see that this action preserves the rank of ρ, and it actually holds a stronger result [29]:

Remark 6. *The orbit of the action of the special linear group $SL(n,\mathbb{C})$ defined by equation (1.55) on a state ρ of rank k coincides with the stratum $\mathcal{D}^k(\mathcal{H})$ of $\mathcal{D}(\mathcal{H})$*

Clearly all of this can be reproduced for π_2 considering the left action of $GL(n,\mathbb{C})$. Notice that if we use only elements in the subgroup $U(n)$ of $GL(n,\mathbb{C})$ we get again the coadjoint action of the unitary group, that we already discussed. Notice that the bundle we constructed for mixed states is trivial, unlike the one we constructed for pure states.
In fact, every $n \times n$ complex matrix admits the following decomposition:

$$M_n(\mathbb{C}) \ni A = UP \tag{1.56}$$

Where P is a semi-positive definite Hermitian matrix and U is a unitary matrix. This decomposition is called *polar decomposition* of the matrix A.

If one restricts himself to invertible matrices this decomposition is also unique. So in our case we can always write:

$$GL(n, \mathbb{C}) \ni A = U\sqrt{\rho} \tag{1.57}$$

And clearly the point A in the total space projects via π_1 on the point on the base space ρ.

So we can define everywhere the section:

$$\sigma_U : \mathcal{D}_n(\mathcal{H}) \ni \rho \mapsto U\sqrt{\rho} \in GL(n, \mathbb{C}) \tag{1.58}$$

So this will be a global section for our bundle, showing that it is trivial.

1.3 Purification of mixed states

Now we will see how the geometrical approach developed in this chapter can be used to make contact with the idea of purification of mixed states. Let us recall briefly what we mean by purification of mixed states.

Let us consider a mixed state ρ_1 in the space of density states $\mathcal{D}(\mathcal{H}_1)$ associated to the Hilbert space $\mathcal{H}_1$, then what one could ask is: can ρ_1 be seen as the reduction of a pure state of a bigger system?

What we are saying is that we want to find a pure state ρ_{12} in $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, that is indistinguishable from ρ_1 if one makes measurements only on the system associated to $\mathcal{H}_1$. This equals to find a ρ_{12} such that:

$$\rho_1 = \mathrm{Tr}_2\, \rho_{12} \qquad \rho_{12}^2 = \rho_{12} \tag{1.59}$$

Where with Tr_2 we mean the partial trace over the subsystem $\mathcal{H}_2$.

Going from ρ_1 to ρ_{12} is what is usually called *purification of mixed states*, while the inverse process is called *reduction*. Clearly the purification process has an ambiguity, we can see this by noticing that:

$$\rho_1 = \mathrm{Tr}_2\, \rho_1 \otimes \rho_2 \qquad \forall \rho_2 \in \mathcal{D}(\mathcal{H}_2) \tag{1.60}$$

This means that in the reduction process all information about the state ρ_2 is lost.

In order to continue our discussion let us give the following result [6]: any density matrix ρ on a Hilbert space $\mathcal{H}$ can always be purified choosing $\mathcal{H}_2 = \mathcal{H}^*$, this means that we can find a purification of ρ as a vector of the Hilbert-Schmidt space $B(\mathcal{H}) = \mathcal{H} \otimes \mathcal{H}^*$.

So now we will take a slightly broader point of view and say that a purification process is a procedure that, given a density state, associates to it a bounded operator (that means a generic linear operator if we restrict ourselves to finite dimensional Hilbert spaces). The associated procedure of that allows us to "go back" to the density state will be called a reduction procedure.

The bundle picture developed the previous section is perfectly suitable to describe this situation. Recall now that we had two distinct projection maps:

$$\pi_1 : GL(n,\mathbb{C}) \ni A \mapsto \rho = \frac{A^\dagger A}{\mathrm{Tr}\{A^\dagger A\}} \in \mathcal{D}^n(\mathcal{H}) \tag{1.61}$$

$$\pi_2 : GL(n,\mathbb{C}) \ni A \mapsto \rho = \frac{AA^\dagger}{\mathrm{Tr}\{A^\dagger A\}} \in \mathcal{D}^n(\mathcal{H}) \tag{1.62}$$

At this level the choice of π_1 or π_2 remains arbitrary, but in chapter 5 we will give different interpretations of these two maps.

Now we can point out a strong analogy between this procedure and the *leitmotiv* of this work, that is the process of going from probabilities to amplitudes.

In fact we can now see the purification procedure as a section of this principal bundle, while on the other hand the reduction procedure is represented by one of the two projection maps π_1 or π_2.

Now let us study some geometrical aspects of this structure, from now on in this section we will always consider the projection map π_2.

In the first place let us find the vertical vectors of this bundle. As we already said, the fibers of our principal bundle are the orbits of the right action of the unitary group on $GL(n,\mathbb{C})$ (set apart the part relative to the action of $\mathbb{R}_+$). So being A_0 an element of $GL(n,\mathbb{C})$, the following curve:

$$I \in t \mapsto A(t) = A_0 U(t) \quad U(t) \in \mathcal{U}(n) \quad \forall t \in I \tag{1.63}$$

is contained in a fiber.

Provided that this curve is regular, we can find the vector tangent to this curve in A_0:

$$\left.\frac{d}{dt}A(t)\right|_{t=0} = A_0 \left.\frac{dU(t)}{dt}\right|_{t=0} \tag{1.64}$$

Now let us use the fact that anti-Hermitian matrices are the infinitesimal generators of unitary transformation:

$$\left.\frac{d}{dt}A(t)\right|_{t=0} = A_0 K \tag{1.65}$$

With $K = -K^\dagger$.

Being this curve all contained in a fiber, the tangent vectors will be vertical.

So being both the starting point A_0 and the curve generic, we can conclude that:

$$AK \in V_A(GL(n, \mathbb{C})) \quad \forall K = -K^\dagger \tag{1.66}$$

Where $V_A(GL(n, \mathbb{C})$ is the vertical part of the space tangent to $GL(n, \mathbb{C})$ in A.

Vertical vectors are determined once one specifies the bundle structure, this is not true for horizontal vectors: in order to specify horizontal vectors we need to introduce a connection.

Let us recall that on the Hilbert-Schmidt space $\mathcal{B}(\mathcal{H})$ is canonically defined an Hermitian product:

$$\langle A, B \rangle = \mathrm{Tr}\{A^\dagger B\} \quad A, B \in \mathcal{B}(\mathcal{H}) \tag{1.67}$$

we can take its real part to get a metric on $\mathcal{B}(\mathcal{H})$:

$$g(A, B) = \frac{1}{2}\mathrm{Tr}\{A^\dagger B + B^\dagger A\} \tag{1.68}$$

This is what is usually called *Bures metric* [9].

One typical choice for the horizontal spaces is the space of vectors that are orthogonal (w.r.t. the metric we just defined) to vertical vectors. In that case one would get that a vector $X \in T_A GL(n, \mathbb{C})$ is horizontal if:

$$g(AK, X) = 0 \tag{1.69}$$

That is:

$$\begin{aligned}\mathrm{Tr}\{(AK)^\dagger X + X^\dagger AK\} &= \mathrm{Tr}\{-A^\dagger XK + X^\dagger AK\} \\ &= \mathrm{Tr}\{(X^\dagger A - A^\dagger X)K\} = \langle X^\dagger A - A^\dagger X, K\rangle = 0\end{aligned} \tag{1.70}$$

And this is sufficient to conclude that:

$$X^\dagger A - A^\dagger X = 0 \tag{1.71}$$

This is because every element M of $\mathcal{M}_n(\mathbb{C})$ can be written as:

$$M = K_1 + iK_2 \tag{1.72}$$

Where K_1 and K_2 are anti-Hermitian matrices.
Then one could ask if can be found a connection form that has the vectors of the form (1.71) in its kernel.
This problem was addressed by Uhlmann [41] and he concluded that such a connection form $\mathscr{A}$ has to satisfy the following relation:

$$\mathscr{A} A^\dagger A + A^\dagger A \mathscr{A} = A^\dagger dA - (dA)^\dagger A \tag{1.73}$$

This is the approach usually followed if one is interested in geometric phases [7] [40] [20] [19].

Chapter 2

Riemannian and Poisson Geometry on $\mathcal{H}_0$

Let us start this chapter with the definition of Kähler manifold.

Before we say what a Kähler manifold is, we need the notion of *realification* of a complex Hilbert space: given a complex Hilbert space $\mathcal{H}$ with (complex) dimension n, the *realified* $\mathcal{H}_{\mathbb{R}}$ of $\mathcal{H}$ is a $2n$-dimensional real vector space that has the same group structure of $\mathcal{H}$. Being $\mathcal{H}_{\mathbb{R}}$ a real vector space, only multiplication by real scalars is allowed.

A complex structure is an operator that plays the role of the multiplication by the imaginary unit i, so it will satisfy the property $J^2 = -\mathbb{I}$.

Let us construct a complex structure J, given a basis $B = (e_1, e_2, \ldots, e_n)$ in $\mathcal{H}$ and a basis in $B_{\mathbb{R}} = (f_1, f_2, \ldots, f_{2n})$ in $\mathcal{H}_{\mathbb{R}}$, we associate to a vector ψ of $\mathcal{H}$ specified in the base B by the components $(z_1, \ldots, z_n)$, with $z_j = u_j + iv_j$, to a vector in $\mathcal{H}_{\mathbb{R}}$ specified in the base $B_{\mathbb{R}}$ by the components $(u_1, \ldots, u_n, v_1, \ldots, v_n)$.

The multiplication by the imaginary unit i in $\mathcal{H}$ will be represented in $\mathcal{H}_{\mathbb{R}}$ by an operator J that maps the vector $\psi = (u_1, \ldots, u_n, v_1, \ldots, v_n)$ in the vector $J\psi = (-v_1, \ldots, -v_n, u_1, \ldots, u_n)$.

A complex manifold is a manifold M that can be locally modeled on $\mathbb{C}^n$ for some n. Then on the tangent bundle TM one can define the complex structure J via:

$$J : TZ \longrightarrow TZ \quad s.t. \quad J(X) = iX \qquad \forall v \in TZ. \tag{2.1}$$

In a real, even-dimensional manifold, let us call it $\mathcal{K}$, with a complex structure and a closed two form satisfying the property:

$$\omega(JX, JY) = \omega(X, Y) \qquad \forall X, Y \in T\mathcal{K} \tag{2.2}$$

One can define a type $(0, 2)$ tensor in the following way:

$$g(X, Y) = \omega(X, JY) \qquad \forall X, Y \in T\mathcal{K} \tag{2.3}$$

And this will be a symmetric tensor, also it will be non degenerate if ω is non degenerate. If g is also positive, then $\mathcal{K}$ is a Kähler manifold, notice that (2.2) and (2.3) imply:

$$g(JX, JY) = g(X, Y) \qquad \forall X, Y \in T\mathcal{K} \tag{2.4}$$

The triple (g, J, ω) will be called an *admissible triple*.
Notice that (2.2), (2.3) and (2.4) show that J is a generator for both finite and infinitesimal orthogonal and symplectic transformations.
We can also notice that it is possible to have alternative symplectic structures and alternative complex structures which in some combinations give positive definite symmetric tensors while in some other combinations give symmetric tensors which are not positive definite.

2.1 $\mathcal{H}_0$ as a Kähler manifold

Let us now take a closer look at the Hermitian structure h defined in (1.1), one can separate its real and imaginary part:

$$h(\psi, \phi) = \frac{1}{2}(\langle\psi|\phi\rangle + \langle\phi|\psi\rangle) + \frac{1}{2}(\langle\psi|\phi\rangle - \langle\phi|\psi\rangle) = g(\psi, \phi) + i\omega(\psi, \phi) \tag{2.5}$$

Notice that its real part is a symmetric, positive and nondegenerate form, while its imaginary part is an antisymmetric and nondegenerate form. So we can define the following objects:

$$g(\psi, \phi) = \frac{1}{2}(\langle\psi|\phi\rangle + \langle\phi|\psi\rangle) \tag{2.6}$$

$$\omega(\psi, \phi) = \frac{1}{2}(\langle\psi|\phi\rangle - \langle\phi|\psi\rangle) \tag{2.7}$$

The aim of this section will be to "promote" this objects to tensorial quantitities and use them as a metric tensor and a symplectic form on $\mathcal{H}_0$ [22].
Being $\mathcal{H}$ a vector space, we can do what follows: given a point p in $\mathcal{H}$ and a vector $\psi \in T_p\mathcal{H} \cong \mathcal{H}$ we can construct the constant vector field that associates ψ to every point of $\mathcal{H}$, let us call it X_ψ. Now we can redefine g and ω as tensors of type (0,2) in the following way:

$$g(\psi, \phi) = g(p)(X_\psi, X_\phi) \tag{2.8}$$

$$\omega(\psi, \phi) = \omega(p)(X_\psi, X_\phi) \tag{2.9}$$

The form g can now be considered a Riemannian metric and ω a symplectic form on an Hilbert manifold.
We will now switch to the Dirac notation and give explicit expressions of g, ω and J in an orthonormal basis $(|e_1\rangle, |e_2\rangle, \ldots, |e_n\rangle)$, in the following will be used the Einstein convention on summations over repeated indices:

$$|\psi\rangle = z^j_\psi |e_j\rangle = (x^j_\psi + iy^j_\psi) |e_j\rangle \tag{2.10}$$

$$|\phi\rangle = z^j_\phi |e_j\rangle = (x^j_\phi + iy^j_\phi) |e_j\rangle \tag{2.11}$$

$$h(\psi, \phi) = \langle\psi|\phi\rangle = x^j_\psi x_{\phi j} + y^j_\psi y_{\phi j} + i(x^j_\psi y_{\phi j} - y^j_\psi x_{\phi j}) \tag{2.12}$$

$$X_\psi = z^j_\psi \frac{\partial}{\partial z^j} = x^j_\psi \frac{\partial}{\partial x^j} + iy^j_\psi \frac{\partial}{\partial y^j} \tag{2.13}$$

$$X_\phi = z^j_\phi \frac{\partial}{\partial z^j} = x^j_\phi \frac{\partial}{\partial x^j} + iy^j_\phi \frac{\partial}{\partial y^j} \tag{2.14}$$

With these positions one can easily show that in order to satisfy (2.5), (2.8) and (2.9) g and ω have to assume the following form:

$$g = \sum_{j=1}^{n} (dx^j \otimes dx^j + dy^j \otimes dy^j) \tag{2.15}$$

$$\omega = \sum_{j=1}^{n} dx^j \wedge dy^j \tag{2.16}$$

We can also easily find the form of the tensor J:

$$J = dy^j \otimes \frac{\partial}{\partial x^j} - dx^j \otimes \frac{\partial}{\partial y^j} \tag{2.17}$$

The following equalities are satisfied:

$$J^2 = -\mathbb{I} \tag{2.18}$$

$$g(JX_\psi, JX_\phi) = g(X_\psi, X_\phi) \tag{2.19}$$

$$\omega(JX_\psi, JX_\phi) = \omega(X_\psi, X_\phi) \tag{2.20}$$

$$\omega(JX_\psi, X_\phi) = g(X_\psi, X_\phi) \tag{2.21}$$

Now we can see that $\mathcal{H}_0$ (or better its realified $\mathcal{H}_\mathbb{R}$) is a Kähler manifold and the triple (g, J, ω) is an *admissible triple.*

Let us write the complete expression of h:

$$h = \sum_{j=1}^{n}(dx^j \otimes dx^j + dy^j \otimes dy^j) + i \sum_{j=1}^{n} dx^j \wedge dy^j \qquad (2.22)$$

Notice that this can be simply written as:

$$h = \langle d\psi | d\psi \rangle \qquad (2.23)$$

Let us conclude this section by noticing that ω can be obtained in the following way:

$$\omega = \frac{1}{2} dJd(\langle \psi | \psi \rangle) \qquad (2.24)$$

i.e. that the function $\frac{1}{2}\langle \psi | \psi \rangle$ is a Kähler potential for the sympletic form ω

$$dJd(\frac{1}{2}\langle \psi | \psi \rangle) = dJ(x_j dx^j + y_j dy^j) = d(x_j dy^j - y_j dx^j) = \sum_{j=1}^{n} dx^j \wedge dy^j \qquad (2.25)$$

Using (2.17) and (2.21) one can obtain again (2.15).

2.2 A connection one-form on $\mathcal{H}_0$

On $\mathcal{H}_0$ one can construct the following vector fields:

$$\Delta = x^j \frac{\partial}{\partial x^j} + y^j \frac{\partial}{\partial y^j} \tag{2.26}$$

$$\Gamma = y^j \frac{\partial}{\partial x^j} - x^j \frac{\partial}{\partial y^j} = J(\Delta) \tag{2.27}$$

These definitions allow us to obtain the bundle structure constructed in chapter 1 in geometric terms: consider the distribution generated by Δ and Γ, being involutive this distribution will be associated to a foliation which is regular in $\mathcal{H}_0$, the quotient space with respect to this foliation is indeed $\mathcal{P}(\mathcal{H})$. The integral curves of Δ and Γ are respectively the orbits of the action of $\mathbb{R}_+$ and of $U(1)$ on $\mathcal{H}_0$, so they jointly generate the action of $\mathbb{C}_0$.
These vector fields will be tangent to the fibers of our principal bundle, so they will be vertical vector fields, from now on we will refer to these vector fields as *fundamental vector fields*, because they generate the action of $\mathbb{C}_0$.
From the previous remark it is clear that a generic vertical vector will be of the form:

$$V(p) = \alpha\, \Delta(p) + \beta\, \Gamma(p) \quad with \quad \alpha, \beta \in \mathbb{R} \tag{2.28}$$

Where p is a point in the total space. The fundamental vector fields can also be used to construct a connection on our principal bundle:

$$\mathcal{A} = \Delta \otimes \frac{1}{2} \frac{d\langle\psi|\psi\rangle}{\langle\psi|\psi\rangle} + \Gamma \otimes \frac{1}{2} \frac{d_J\langle\psi|\psi\rangle}{\langle\psi|\psi\rangle} \tag{2.29}$$

For future use, let us define the following form:

$$\theta = \frac{1}{2} \frac{d\langle\psi|\psi\rangle}{\langle\psi|\psi\rangle} \tag{2.30}$$

So we can rewrite our connection in the following way:

$$\mathcal{A} = \Delta \otimes \theta + \Gamma \otimes J(\theta) \tag{2.31}$$

Let us give the explict expression of θ and $J(\theta)$:

$$\theta = \frac{x_j dx^j + y_j dy^j}{x_j x^j + y_j y^j} \tag{2.32}$$

$$J(\theta) = \frac{x_j dy^j - y_j dx^j}{x_j x^j + y_j y^j} \tag{2.33}$$

From which one can easily show:

$$\theta(\Delta) = 1 \qquad (J(\theta))(\Delta) = 0 \tag{2.34}$$

$$\theta(\Gamma) = 0 \qquad (J(\theta))(\Gamma) = 1 \tag{2.35}$$

So $\mathcal{A}$ is the identity over vertical vectors:

$$\mathcal{A}(\Delta) = \Delta \tag{2.36}$$

$$\mathcal{A}(\Gamma) = \Gamma \tag{2.37}$$

Vectors that are in the kernel of this $(1, 1)$ type tensor field will be called horizontal vectors. Our expressions clearly show that these structures may be derived from a potential:

$$F = \frac{1}{2} \log \langle\psi|\psi\rangle \tag{2.38}$$

One has that

$$dF = \theta \tag{2.39}$$

So we can rewrite our connection in the following way:

$$\mathcal{A} = \Delta \otimes dF + J(\Delta) \otimes d_J F \tag{2.40}$$

Now we will prove that the function F defined in (2.38) is the only function of $\langle\psi|\psi\rangle$ that gives a connection in the form (2.40) that defines Δ and $\Gamma = J(\Delta)$ as vertical vectors.

This amounts to prove that the one form $\theta = dF$ satisfies properties (2.34) and (2.35). The first two of these relations in coordinates give the following relations:

$$\frac{\partial F}{\partial x^j} x^j + \frac{\partial F}{\partial y^j} y^j = 1 \tag{2.41}$$

$$\frac{\partial F}{\partial x^j} y^j - \frac{\partial F}{\partial y^j} x^j = 0 \tag{2.42}$$

The other two give the same relations.

Now assuming that

$$F = F(\langle\psi|\psi\rangle) \tag{2.43}$$

One gets that (2.42) is automatically verified, while (2.41) gives:

$$2\frac{\partial F}{\partial \langle\psi|\psi\rangle} x_j x^j + 2\frac{\partial F}{\partial \langle\psi|\psi\rangle} y_j y^j + = 1 \tag{2.44}$$

That means:

$$\frac{\partial F}{\partial \langle\psi|\psi\rangle} = \frac{1}{2\langle\psi|\psi\rangle} \tag{2.45}$$

And this finally gives:

$$F = \frac{1}{2} \log \langle\psi|\psi\rangle \tag{2.46}$$

Up to an additive constant.

2.3 Fubini-Study metric on $\mathcal{H}_0$

Now we want to obtain a metric tensor on $\mathcal{P}(\mathcal{H})$. In the first place we want to redefine the Hermitian tensor h in order to make it "gauge invariant" (that is invariant under the action of $\mathbb{C}_0$), we do this by normalizing it:

$$h = \frac{\langle d\psi | d\psi \rangle}{\langle \psi | \psi \rangle} \tag{2.47}$$

Notice that this tensor defined on $\mathcal{H}_0$, altough it is gauge invariant, cannot be seen as the pullback via the projection map (1.9) of a tensor defined on $\mathcal{P}(\mathcal{H})$, this can be done if:

$$\mathbf{i}_\Delta h = 0 \tag{2.48}$$

$$\mathbf{i}_\Gamma h = 0 \tag{2.49}$$

But this is not the case. What we can do is, in the spirit of Kaluza-Klein theory, add to this tensor some terms in order to make relations (2.48) and (2.49) satisfied. Now we are going to show that the quantity

$$\tilde{h} = h - \theta \otimes \theta - J(\theta) \otimes J(\theta) - i\, J(\theta) \wedge \theta \tag{2.50}$$

is a tensor on $\mathcal{H}$ that can be thought of as the pullback via the projection map defined in our bundle of a tensor defined on the base space.

$$\mathbf{i}_\Delta \tilde{h} = \mathbf{i}_\Delta h - \mathbf{i}_\Delta(\theta)\theta - \mathbf{i}_\Delta(J(\theta))J(\theta) - i\,\mathbf{i}_\Delta(J(\theta))\theta + i\,\mathbf{i}_\Delta(\theta)J(\theta) \tag{2.51}$$

So using (2.34) and (2.35):

$$\mathbf{i}_\Delta \tilde{h} = \mathbf{i}_\Delta h - \theta + iJ(\theta) \tag{2.52}$$

Using coordinate expression for h, θ and $J(\theta)$:

$$\mathbf{i}_\Delta \tilde{h} = \frac{x_j dx^j + y_j dx^j + i(x_j dy^j - y_j dx^j)}{x_j x^j + y_j y^j} - \frac{x_j dx^j + y_j dx^j}{x_j x^j + y_j y^j} + i\frac{y_j dx^j - x_j dy^j}{x_j x^j + y_j y^j} = 0 \quad (2.53)$$

Now turning to Γ:

$$\mathbf{i}_\Gamma \tilde{h} = \mathbf{i}_\Gamma h - \mathbf{i}_\Gamma(\theta)\theta - \mathbf{i}_\Gamma(J(\theta))J(\theta) - i\,\mathbf{i}_\Gamma(J(\theta))\theta + i\,\mathbf{i}_\Gamma(\theta)J(\theta) \quad (2.54)$$

And again using (2.34) and (2.35):

$$\mathbf{i}_\Gamma \tilde{h} = \mathbf{i}_\Gamma h + J(\theta) + i\theta \quad (2.55)$$

And again using coordinate expressions for this quantities:

$$\mathbf{i}_\Gamma \tilde{h} = \frac{-y_j dx^j + x_j dy^j - i(y_j dy^j + x_j dx^j)}{x_j x^j + y_j y^j} + \frac{y_j dx^j - x_j dy^j}{x_j x^j + y_j y^j} + i\frac{x_j dx^j + y_j dy^j}{x_j x^j + y_j y^j} = 0 \quad (2.56)$$

Let us write the terms that we added to h in another form:

$$\theta \otimes \theta + J(\theta) \otimes J(\theta) + iJ(\theta) \wedge \theta = \frac{(x_j x_k + y_j y_k)(dx^j \otimes dx^k + dy^j \otimes dy^k + i dx^j \wedge dy^k)}{(x_l x^l + y_l y^l)^2} = \frac{\langle d\psi|\psi\rangle \otimes \langle \psi|d\psi\rangle}{\langle \psi|\psi\rangle^2} \quad (2.57)$$

So we can rewrite $\tilde{h}$ in the following way:

$$\tilde{h} = \frac{\langle d\psi|d\psi\rangle}{\langle \psi|\psi\rangle} - \frac{\langle d\psi|\psi\rangle \otimes \langle \psi|d\psi\rangle}{\langle \psi|\psi\rangle^2} \quad (2.58)$$

And this is just what is usually called the Fubini-Study metric. We will prove that the imaginary part is given by:

$$dd_J\big(\frac{1}{4}\log\langle\psi|\psi\rangle\big) = dd_J\big(\frac{1}{2}F\big) \tag{2.59}$$

In fact we have:

$$\begin{aligned} dd_J\left(\frac{1}{4}\log\langle\psi|\psi\rangle\right) &= d\left(\frac{d_J\langle\psi|\psi\rangle}{4\langle\psi|\psi\rangle}\right) = \frac{1}{2}dJ(\theta) \\ &= \frac{dd_J(\langle\psi|\psi\rangle)\langle\psi|\psi\rangle}{\langle\psi|\psi\rangle^2} - J(\theta)\wedge d(\langle\psi|\psi\rangle) \\ &= \frac{dd_J\langle\psi|\psi\rangle}{2\langle\psi|\psi\rangle} - J(\theta)\wedge\theta \end{aligned} \tag{2.60}$$

And this is clearly equal to the imaginary part of (2.58), in fact we have, recalling (2.24), that:

$$\operatorname{Im} h = \frac{\omega}{\langle\psi|\psi\rangle} = \frac{dd_J(\langle\psi|\psi\rangle)}{2\langle\psi|\psi\rangle} \tag{2.61}$$

and this is just the imaginary part of the first term. The other term can be checked directly from (2.57).

Similarly can be proven that the symmetric part is:

$$Jdd_J(\frac{1}{4}\log\langle\psi|\psi\rangle) = Jdd_J\big(\frac{1}{2}F\big) \tag{2.62}$$

This makes clear that, in order to define ω and g the only addtional objects one needs are the complex structure J and the potential function $F = \frac{1}{2}\log\langle\psi|\psi\rangle$.

2.4 The pull-back to "trial submanifolds of states"

As a matter of facts, in some situation it may be useful to work actually into a subset of the Hilbert space rather then into the whole of it because of computational convenience or because of experimental constraints
This is the typical approach in the so-called *variational method*, the method can be described as follows: is given a certain problem, tipically finding the ground state of a certain Hamiltonian H, being unable to find the solution analitically, one restricts himself to a certain set of states M, called *trial states*, labelled with a certain set of parameters $\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$.
Then one finds the values $\bar{\boldsymbol{\lambda}}$ of said parameters that select the element $\bar{\phi}$ of this set that minimizes the expectation value of the Hamiltonian H. This will be the best approximation in M for the ground states of the Hamiltonian H.
This process can be translated in geometric terms in the following way: consider an immersion of the space of parameters Λ into the Hilbert space $\mathcal{H}$:

$$I : \Lambda \hookrightarrow \mathcal{H} \tag{2.63}$$

The image of this immersion will be the submanifold of trial states M, then we can bring objects of interest on Λ using the pullback of the immersion I and (hopefully) find the point $\bar{\lambda}$ that solves the given problem. Then the solution of our problem will be just the $I(\bar{\lambda})$.
Now we will give an example of restriction of a metric tensor to a trial submanifold of states, this will be done in the case that the full Hilbert space is $L^2(\mathbb{R})$ and we want to restrict Fubini-Study metric to Gaussian states.
This example points out a crucial advantage of this process, in fact being $L^2(\mathbb{R})$ an infinite dimensional Hilbert space, we should develop the calculus for such spaces. But as we will see we will never write the explicit expression for the metric on the full Hilbert space, avoiding technical difficulties.

Example 3. *In order to restrict to Gaussian states let us define an immersion i of the space of parameters that identify a Gaussian state in $L^2(\mathbb{R})$:*

$$i : \mathbb{R}^3 \ni (a, \mu, \phi) \longrightarrow \psi_{a,\mu,\phi} \in L^2(\mathbb{R}) \tag{2.64}$$

Where $\psi_{a,\mu,\phi}$ *is defined by:*

$$\psi_{a,\mu,\phi}(x) = \sqrt{\frac{\sqrt{2}}{a\sqrt{\pi}}} e^{i\phi(x-\mu) - \frac{(x-\mu)^2}{a^2}} \tag{2.65}$$

We want to restrict Fubini-Study metric:

$$h = \frac{\langle d\psi | d\psi \rangle}{\langle \psi | \psi \rangle} - \frac{\langle d\psi | \psi \rangle \langle \psi | d\psi \rangle}{\langle \psi | \psi \rangle^2} \tag{2.66}$$

to such states.

So to begin let us notice that:

$$i^* \left| d\psi \right\rangle = d\psi_{a,\mu,\phi} \tag{2.67}$$

because exterior derivative always commute with the pull-back of a smooth map, and the exterior derivative in the right side is simply done in the space of parameters.

Then we are able to calculate:

$$i^* \left| d\psi \right\rangle = \sqrt{\frac{\sqrt{2}}{a\sqrt{\pi}}} e^{i\phi(x-\mu) - \frac{(x-\mu)^2}{a^2}} \left[\left(\frac{2(x-\mu)^3}{a^3} - \frac{1}{2a} \right) da + \left(\frac{2(x-\mu)}{a^2} - i\phi \right) d\mu + i(x-\mu) d\phi \right] \tag{2.68}$$

Then:

$$i^* \langle d\psi | d\psi \rangle = \frac{\sqrt{2}}{a\sqrt{\pi}} \left[\left(\frac{4I_4}{a^6} - \frac{2I_2}{a^4} + \frac{I_0}{4a^2} \right) da \otimes da + \left(\frac{4I_2}{a^4} + I_0 \phi^2 \right) d\mu \otimes d\mu + I_2 d\phi \otimes d\phi + i\phi \left(\frac{2I_2}{a^3} - \frac{I_0}{2a} \right) da \wedge d\mu + i \frac{2I_2}{a^2} d\mu \wedge d\phi \right] \tag{2.69}$$

And:

$$i^*(\langle d\psi|\psi\rangle \otimes \langle\psi|d\psi\rangle = \frac{2}{a^2\pi}\left[\left(\frac{2I_2}{a^3} - \frac{I_0}{2a}\right)^2 da \otimes da + \phi^2 I_0^2 d\mu \otimes d\mu + i\phi\left(\frac{2I_2}{a^3} - \frac{I_0}{2a}\right) da \wedge d\mu\right] \quad (2.70)$$

Where:

$$I_0 = \int_{\mathbb{R}} e^{\frac{-2(x-\mu)^2}{a^2}} dx = \sqrt{\frac{\pi}{2}} a \quad (2.71)$$

$$I_2 = \int_{\mathbb{R}} (x-\mu)^2 e^{\frac{-2(x-\mu)^2}{a^2}} dx = \sqrt{\frac{\pi}{2}} \frac{a^3}{4} \quad (2.72)$$

$$I_4 = \int_{\mathbb{R}} (x-\mu)^4 e^{\frac{-2(x-\mu)^2}{a^2}} dx = \sqrt{\frac{\pi}{2}} \frac{3a^5}{16} \quad (2.73)$$

While the integrals:

$$I_1 = \int_{\mathbb{R}} (x-\mu) e^{\frac{-2(x-\mu)^2}{a^2}} dx \qquad I_3 = \int_{\mathbb{R}} (x-\mu)^3 e^{\frac{-2(x-\mu)^2}{a^2}} dx \quad (2.74)$$

are zero.

Notice also that:

$$\langle\psi|\psi\rangle = 1 \quad (2.75)$$

Then combining these results one obtains:

$$i^*h = \frac{1}{2a^2} da \otimes da + \frac{1}{a^2} d\mu \otimes d\mu + \frac{a^2}{4} d\phi \otimes d\phi + \frac{i}{2} d\mu \wedge d\phi \quad (2.76)$$

So we have that the pullback of the Hermitian tensor h can be split in its real part, that is a symmetric tensor, and its imaginary part, that is a *skew-symmetric tensor:*

$$i^*g = \frac{1}{2a^2} da \otimes da + \frac{1}{a^2} d\mu \otimes d\mu + \frac{a^2}{4} d\phi \otimes d\phi \tag{2.77}$$

$$i^*\omega = \frac{1}{2} d\mu \wedge d\phi \tag{2.78}$$

Notice that the imaginary part is a closed two form, but it is not symplectic, since it is degenerate, while the real part is genuinely a metric tensor.

Chapter 3

A glance at Information Geometry

Information Geometry is the study of statistical estimation from a geometric point of view, this means that we will give a geometric setting for the discussion and construct geometrical object that have significance from the point of view of Information Theory.

In this chapter we will give a brief review of this subject, a complete treatment can be found on [3]. This subject is widely studied and has a vast number of application, for example in Machine Learning and Signal Optimization [4].

3.1 Statistical manifolds and distinguishability

Let $\boldsymbol{\chi}$ be a sample space, let us assume for simplicity that it is finite (of cardinality n): $\boldsymbol{\chi} = \{x_1, x_2, \ldots, x_n\}$. We can construct the space $P(\boldsymbol{\chi})$ of probability distributions on $\boldsymbol{\chi}$, that is the space of maps of this kind:

$$p : \boldsymbol{\chi} \longrightarrow \mathbb{R} \quad s.t. \quad p(x_j) \geqslant 0; \ \sum_{j=0}^{n} p(x_j) = 1 \tag{3.1}$$

Then we could take some family S of elements of $P(\boldsymbol{\chi})$, for example one that can be parametrized by a set of parameters $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$. Then a statistical manifold is a manifold whose points are in a one to one correspondence with the elements of S. Let us give an example:

Example 4. *Let $\boldsymbol{\chi} = \{H, T\}$, this is a sample space of cardinality two, and can be seen as the sample space of a coin toss. The space of all probability distributions is just the mono-dimensional simplex, as we already pointed out in chapter 2. So for example a fair coin will give equal probabilities of getting a tail or a head, and will be associated to the barycenter of the simplex (see fig. 1.2), while a biased coin that always gives head will be associated to one of the extremal points of the simplex. Now let us consider the problem of distinguishing between these two coins. Distinguish between two probability distributions means that, having the possibility to make trials with a source that obeys some distribution p, we want to know what is the probability, after N trials, that we will observe a different distribution q.*
It is easy to convince ourselves that this task is more difficult (that is takes a larger number of trials to have the same confidence) if we can make trials on the biased coin, while if we can make trials on the fair coin, and we are lucky enough to get a tail, we could also conclude that the probability that the coin is biased is zero.
This simple example makes clear the fact that distinguishability between statistical distributions can not in general be seen as a distance on a statistical

manifold, since distance between two points is a symmetric quantity in the exchange of the two points, while we have shown that distinguishability is not.

So, in order to make a statistical discussion, we need a measure of distinguishability between distributions, this will be provided by the so-called *divergence function* (or *contrast function*). We will first give the definition in coordinates, then switch to a coordinate-free discussion of the argument in the next section.

Definition 1. *A divergence function F is a two point function on a statistical manifold M:*

$$F : M \times M \mapsto \mathbb{R} \tag{3.2}$$

that satisfies the following properties, introducing local coordinates (x, y) on $M \times M$:

$$F(x, y) \geqslant 0 \tag{3.3a}$$

$$F(x, y) = 0 \quad iff \quad x = y \tag{3.3b}$$

$$\left.\frac{\partial F}{\partial x^j}\right|_{x=y} = 0 \quad for \quad j = 1, 2, \ldots, n \tag{3.3c}$$

$$\left.\frac{\partial F}{\partial y^j}\right|_{x=y} = 0 \quad for \quad j = 1, 2, \ldots, n \tag{3.3d}$$

And such that the Hessian:

$$G_{jk} = \left.\frac{\partial^2 F}{\partial x^j \partial y^k}\right|_{x=y} \tag{3.4}$$

is a positive-definite matrix.

It is worth noting here that we did not demand that the function F is symmetric in the exchange of the two arguments. As we discussed in example 4, this property is necessary in order to makes the function F suitable as a distinguishability measure.

3.2 From divergence functions to metric tensors

Now we are going to restate the definition of divergence functions in an intrinsic fashion and then see how we can "promote" the Hessian in (3.4) to a metric tensor [32]. In order to do that we need to introduce some geometric tools on $M \times M$.

Let us define the diagonal immersion:

$$i_D : M \ni m \hookrightarrow (m, m) \in M \times M \tag{3.5}$$

In order to rewrite equations (3.3) in an intrinsic fashion we need to deal with bi-forms [35], a bi-form is an element of $\Omega^p(M) \otimes \Omega^q(M)$ so it can be regarded either as a q-form valued p-form on M or as a p-form valued q-form on M. We also need the following definitions:

$$d_1 \otimes \mathbf{I} : \Omega^p(M) \otimes \Omega^q(M) \longrightarrow \Omega^{p+1}(M) \otimes \Omega^q(M) \tag{3.6}$$

$$\mathbf{I} \otimes d_2 : \Omega^p(M) \otimes \Omega^q(M) \longrightarrow \Omega^p(M) \otimes \Omega^{q+1}(M) \tag{3.7}$$

Where d_1 and d_2 act as the canonical exterior derivative on $\Omega(M)$.

Now we are in the position to rewrite equations (3.3):

$$F(p) \geqslant 0 \qquad \forall p \in M \times M \tag{3.8a}$$

$$i_D^* F = 0 \tag{3.8b}$$

$$i_D^*(d_1 \otimes \mathbf{I})F = 0 \tag{3.8c}$$

$$i_D^*(\mathbf{I} \otimes d_2)F = 0 \tag{3.8d}$$

From now on we will use a simplified notation and simply write d_1 and d_2 when there is no ambiguity.

In order to rewrite the property of the Hessian of F to be positive-definite we

need to consider two immersions of vector fields of M into the vector fields on $M \times M$. Given a vector field $X \in \mathfrak{X}(M)$ we can construct the following vector fields in $\mathfrak{X}(M \times M)$: $X_l = X \oplus \{0\}$ or $X_r = \{0\} \oplus X$ where the direct sum is meant in the module space of vector fields.
Then we define the metric tensor g defined by F in the following way:

Definition 2. *We say that the metric tensor g is obtained from the divergence function F if:*

$$g(X, Y) := i_D^*\big((d_1 d_2 F)(X_l, Y_r)\big) = .i_D^*(L_{X_l} L_{Y_r} F) \tag{3.9}$$

for all X, Y in $\mathfrak{X}(M)$

Now we will retrace this whole procedure using as statistical manifold the $n-1$-dimensional simplex and as divergence function the *Shannon relative entropy.*

Example 5. *A point in the $n-1$-dimensional simplex $\mathcal{S}_{n-1}$ is individuated by an n-tuple $\boldsymbol{p} = (p_1, p_2, \ldots, p_n)$ s.t. $0 \leqslant p_j \leqslant 1 \quad \forall j = 1, 2, \ldots, n$ and they sum to one. Shannon relative entropy between two probability distribution $\boldsymbol{p}$ and $\boldsymbol{q}$ is defined as:*

$$S_{SH}(\boldsymbol{p}, \boldsymbol{q}) = \sum_{j=1}^{n} p_j \log \frac{p_j}{q_j} \tag{3.10}$$

This is a function from $\mathcal{S}_{n-1} \times \mathcal{S}_{n-1}$ to $\mathbb{R}$, this function satisfies the first two relations of (3.8) but not the other two. So in order to call it a divergence function we need to modify it in this way, for details and motivations see [3]:

$$\widetilde{S}_{SH}(\boldsymbol{p}, \boldsymbol{q}) = \sum_{j=1}^{n} p_j \log \frac{p_j}{q_j} - p_j + q_j \tag{3.11}$$

This new function satisfies the first two relations of (3.8) but now:

$$i_D^*(d_1\widetilde{S}_{SH}) = i_D^*\left(\sum_{j=1}^n dp_j \log p_j + p_j d(\log p_j) - dp_j \log q_j - dp_j\right) =$$
$$i_D^*\left(\sum_{j=1}^n dp_j \log p_j - dp_j \log q_j\right) = \sum_{j=1}^n dp_j \log p_j - dp_j \log p_j = 0 \quad (3.12)$$

And an analog calculation can be done for d_2*, so this is a divergence function. Anyway* S_{SH} *and* $\widetilde{S}_{SH}$ *define the same metric tensor, so we refer to* S_{SH} *as a divergence function and use it to construct the metric tensor:*

$$g_{FR} = i_D^*\big(d_1 d_2 S_{SH}\big) = \sum_{j=1}^n p_j d\log p_j \otimes d\log p_j \quad (3.13)$$

This is what is usually called the Fisher-Rao metric tensor. In section 3.4 we will see the importance of this metric tensor. This tensor can be rewritten in the following forms:

$$g_{FR} = \sum_{j=1}^n dp_j \otimes d\log p_j = \sum_{j=1}^n d\log p_j \otimes dp_j$$
$$= \sum_{j=1}^n p_j d\log p_j \otimes d\log p_j = \sum_{j=1}^n 4d\sqrt{p_j} \otimes d\sqrt{p_j} \quad (3.14)$$

the third form appears as an expectation-value 2-form and will be usually preferred.
If we introduce the parameters $x_j = 2\sqrt{p_j}$ *in the last form Fisher-Rao metric appears as an Euclidean metric:*

$$g_{FR} = \sum_{j=1}^n dx_j \otimes dx_j \quad (3.15)$$

Let us conclude this section with an interesting observation, we have seen how we can obtain a metric tensor on M from a divergence function as:

$$g(X,Y) = i_D^*\big((d_1 d_2 F)(X_l, Y_r)\big) \tag{3.16}$$

But in chapter 2 we also obtained a metric tensor (and a symplectic form) in a different fashion, using the complex structure defined on our complex manifold. Now, let us introduce on $M \times M$ the coordinates $\{x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n\}$, where the first n coordinates specify a point on the first copy of M and the last n do the same on the second copy of M.
We can define on $M \times M$ the following complex structure:

$$J = dx^j \otimes \frac{\partial}{\partial x^j} - dy^j \otimes \frac{\partial}{\partial y^j} \tag{3.17}$$

We proved that with this choice of the complex structure the following result holds:

$$i_D^*(dd_J F(JX,Y)) = -4 i_D^*(d_1 d_2 F(X_l, Y_r)) \tag{3.18}$$

Where d is the exterior derivative of $M \times M$, X and Y are vector fields defined on $M \times M$:

$$X = X_x^j \frac{\partial}{\partial x^j} + X_y^j \frac{\partial}{\partial y^j} \in \mathfrak{X}(M \times M) \tag{3.19}$$

$$Y = Y_x^j \frac{\partial}{\partial x^j} + Y_y^j \frac{\partial}{\partial y^j} \in \mathfrak{X}(M \times M) \tag{3.20}$$

While:

$$X_l = X_x^j \frac{\partial}{\partial x^j} \in \mathfrak{X}(M) \oplus \{0\} \tag{3.21}$$

$$Y_r = Y_y^j \frac{\partial}{\partial y^j} \in \{0\} \oplus \mathfrak{X}(M) \tag{3.22}$$

This makes a connection between the two procedures introduced in this and in the previous chapter to obtain a metric, but actually the procedure involving the complex structure is more general.

In fact we have shown that the two sides of (3.18) coincide only when we choose a J in the form in which we defined it in (3.17). This gives the possibility of obtaining a metric for every possible choice of the complex structure, and also has another advantage: the definitions of d_1, d_2, require that the manifold has the structure of a product $M \times M$, while defining d and d_J doesn't require such a structure.

Let us conclude this section with the proof of (3.18):

We will write the two sides of (3.18) and then compare them.

$$d_1 d_2 F(X_l, Y_r) = \frac{\partial^2 F}{\partial x^j \partial y^j}(dx^j(X_l)dy^k(Y_r) - dy^j(X_l)dx^k(Y_r)) = \frac{\partial^2 F}{\partial x^j \partial y^j} X_x^j Y_y^k \tag{3.23}$$

So:

$$i_D^*(d_1 d_2 F(X_l, Y_r)) = \left.\frac{\partial^2 F}{\partial x^j \partial y^j}\right|_{x=y} X_x^j Y_y^k \tag{3.24}$$

Now let us denote with z_j the two set of coordinates together:

$$z_j = x_j \quad \forall j = 1, \dots, n \tag{3.25}$$

$$z_j = y_j \quad \forall j = n+1, \dots, 2n \tag{3.26}$$

Then on the left hand side of (3.18) we have:

$$\begin{aligned} dd_J F(JX, Y) = {} & \frac{\partial^2 F}{\partial z^j \partial x^j}(dz^j(JX)dx^k(Y) - dx^j(JX)dz^k(Y)) \\ & - \frac{\partial^2 F}{\partial x^j \partial y^j}(dz^j(JX)dy^k(Y) - dx^j(JX)dz^k(Y)) \end{aligned} \tag{3.27}$$

So:

$$dd_J F(JX, Y) = -2\frac{\partial^2 F}{\partial x^j \partial y^j}(X_x^j Y_x^k + X_y^j Y_y^k) \tag{3.28}$$

Taking the pull-back of the diagonal immersion we get:

$$i_D^*(dd_J F(JX, Y)) = -4 \frac{\partial^2 F}{\partial x^j \partial y^j}\bigg|_{x=y} X_x^j Y_y^k \tag{3.29}$$

So (3.18) is proven.

3.3 Dual connections on statistical manifolds

In section 3.2 we have constructed a metric tensor from the second derivative of a divergence function F, now we will see what can be done with the third derivatives of a divergence function.
It can be easily shown that the following quantities:

$$\Gamma^l_{kh}(x) = g^{jl}\frac{\partial^3 F}{\partial x^l \partial x^k \partial y^h}\Big|_{x=y} \qquad \Gamma^{*l}_{jkh}(x) = g^{jl}\frac{\partial^3 F}{\partial x^l \partial y^k \partial y^h}\Big|_{x=y} \tag{3.30}$$

Transform under coordinates transformation as the coefficients of a connection, so these cannot be seen as the components of any tensor, but we can define the following quantities:

$$T^h_{jk}(x) = \Gamma^h_{jk}(x) - \Gamma^{*h}_{jk}(x) \tag{3.31}$$

And this quantities will transform as the components of a tensor. In order to understand the role of this two connections and of this tensor we need the definition of dual connections.
Let us denote with L_X the Lie derivative along the field X. Then we define dual connections (w.r.t. the metric g) implicitly with relations:

$$g(\boldsymbol{\nabla}_X Y, Z) := i^*(L_{X_l} L_{Y_l} L_{Z_r} F) \tag{3.32}$$

$$g(\boldsymbol{\nabla}^*_X Y, Z) := i^*(L_{X_l} L_{Y_r} L_{Z_r} F) \tag{3.33}$$

It is shown in [3] that (3.30) define a pair of dual connections with respect to the metric given by the same divergence function. Then we can define the skewness tensor as:

$$T(X, Y, Z) = g(\boldsymbol{\nabla}_X Y, Z) - g(\boldsymbol{\nabla}^*_X Y, Z) \tag{3.34}$$

And this clearly gives equation (3.31) when we write it in coordinates.

The dual connections defined in this section are torsionless, and their "average":

$$\Gamma_0 = \frac{1}{2}(\Gamma + \Gamma^*) \tag{3.35}$$

is also a metric connection for g, so it is the Levi-Civita connection associated to the metric g. The triple (M, g, T) is usually called *statistical model* or *Amari-Chentsov structure*.

One could wonder if there are other combinations of derivatives that give rise to other covariant tensors, in [15] there is an answer to this question.

In this work it is proved that from a divergence function one can extract only one metric, only two dually related connections and no tensors of rank higher than four can be extracted, if the divergence function is analytic.

3.4 Coarse graining and Chentsov theorem

We have seen that we can construct a metric tensor on a statistical manifold from a divergence function, but at this point the choice of the function, and therefore of the metric one is going to obtain seems quite arbitrary.
In this section we will give a criterion to select a certain class of divergence functions, and all these functions give rise to the same metric, that is the Fisher-Rao metric.
This work was made by Chentsov [10] working in the framework of category theory, we will give the result without proof, but we will give the intuition of the theorem working with more familiar objects [17].
Let $\boldsymbol{\chi} = \{x_1, x_2, \ldots, x_n\}$ be a finite sample space, let M_n be the statistical manifold of all probability distribution on $\boldsymbol{\chi}$, and assume we have defined a divergence function F on M. We can divide $\boldsymbol{\chi}$ in m subsets $\{X_1, X_2, \ldots, X_m\}$ with $m < n$ and such that:

$$\bigcup_{j=1}^{m} X_j = \boldsymbol{\chi} \quad X_j \cap X_k = \varnothing \quad \forall j \neq k \tag{3.36}$$

We can consider the case where we are not able to know the outcome of the trial, but only to determine at which subset X_j it belongs, this process goes under the name of *coarse graining.*
So, starting from a probability distribution $p(x) = (p(x_1), p(x_2), \ldots, p(x_n))$, we can consider a coarse grained probability distribution $\tilde{p}$ on a coarse grained sample space $\tilde{\boldsymbol{\chi}} = \{y_1, y_2, \ldots, y_m\}$ such that:

$$\tilde{p}(y_k) = \sum_{x_j \in X_k} p(x_j) \tag{3.37}$$

The space of these probability distribution will form another statistical manifold, let us call it M_m.
So coarse graining is realized by a map:

$$\phi : M_n \longrightarrow M_m \tag{3.38}$$

That maps probability distribution on $\boldsymbol{\chi}$ to probability distributions on $\tilde{\boldsymbol{\chi}}$. Coarse graining is a process with loss of information, so if we want our divergence function F to be a measure of distinguishability between probability distribution, it is reasonable to impose that this should not increase with coarse graining.
In order to do this let us introduce a family of divergence functions $\{F^k\}$ s.t. F^k is a two-point function on the statistical manifold M_k:

Definition 3. *The family $\{F_k\}$ satisfies the monotonicity property if:*

$$F^n(p,q) \geqslant F^m(\phi(p),\phi(q)) \tag{3.39}$$

For all $p,q \in M_n$ and for all maps ϕ of this kind.

A typical example is the family of Shannon entropies:

$$S^k_{SH}(p,q) = \sum_{j=1}^{k} p_k \log \frac{p_k}{q_k} \tag{3.40}$$

We can define the monotonicity property also for metric tensors, let $\{g^k\}$ be a family of metric tensors s.t. g^k is defined on M_k:

Definition 4. *The family $\{g_k\}$ satisfies the monotonicity property if:*

$$g^n(X,X) \geqslant (\phi^* g^m)(X,X) \tag{3.41}$$

For all X in $\mathfrak{X}(M_n)$ and for all maps ϕ of this kind.

One can prove [3] the following result:

Theorem 2. *If a divergence function F obeys the monotonicity property then the metric obtained from F with the procedure exposed in section 3.2 obeys the monotonicity property.*

Now let us state Chentsov theorem:

Theorem 3 (Chentsov)**.** *In the classical setting the only metric that satisfies the monotonicity property is Fisher-Rao metric tensor.*

Let us stress that in the work of Chentsov the fact that g can be obtained from a divergence function plays no role.
But then if one considers this possibility, Chentsov theorem allows to state the following remark:

Remark 7. *In the classical setting every divergence function that satisfies the monotonicity property gives rise to the only metric that satisfies the monotonicity property, that is Fisher-Rao metric.*

In the next chapter we will see that, switching to the quantum case, if one uses a category of maps relevant in this case, one can rewrite the monotonicity condition for two point functions and for metric tensors, and a theorem analogue to theorem 2 holds.
The point is that Chenstov theorem doesn't hold in this case, so there is no possibility of giving a unique metric in the quantum setting.

Chapter 4

From Classical to Quantum probabilities

As we already said, the usual setting for discussing classical probabilities for a finite sample space X of cardinality n is the $n-1$ dimensional simplex S_{n-1}. We can immerse $\mathbb{R}^n$ in the space of $n \times n$ matrices:

$$I : \boldsymbol{p} = (p_1, p_2, \dots, p_n) \longmapsto \rho_0 = \begin{pmatrix} p_1 & 0 & \dots & 0 \\ 0 & p_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p_n \end{pmatrix} \tag{4.1}$$

The image via I of the $n-1$ dimensional simplex S_{n-1} is the space of $n \times n$ diagonal, semi-positive definite and trace one matrices, let us call it $\mathbb{D}_n$. So this can be used to immerse the space of classical probability distributions over a finite sample space in the space of density states of a finite dimensional Hilbert space $\mathcal{H}$.

$$I : \mathbb{D}_n \hookrightarrow \mathcal{D}(\mathcal{H}) \tag{4.2}$$

And this gives the space of "classical states" of quantum mechanics, that is density states that commute with each other.

4.1 Quantum description of classical probabilities

Von Neumann-Umegaki [43] relative entropy is one of the entropies one can use in Quantum Information Theory, it has the form:

$$S_{VN}(\rho,\sigma) = \mathrm{Tr}\{\rho(\log\rho - \log\sigma)\} \qquad \rho,\sigma \in \mathcal{D} \tag{4.3}$$

Where ρ and σ are density states.
Now we can repeat the procedure illustrated in the previous chapter to obtain a metric tensor from this divergence function.
As we did for Shannon relative entropy, we modify it in this way:

$$\widetilde{S}_{VN}(\rho,\sigma) = \mathrm{Tr}\{\rho(\log\rho - \log\sigma) - \rho + \sigma\} \tag{4.4}$$

So that this satisfies conditions (3.8), let us verify condition (3.8c) as an example:

$$\begin{aligned} i_D^*(d_1\widetilde{S}_{VN}(\rho,\sigma)) &= i_D^*(\mathrm{Tr}\{d\rho\log\rho + \rho\, d\log\rho - d\rho\log\sigma - d\rho\} \\ &= \mathrm{Tr}\{d\rho\log\rho + \rho\, d\log\rho - d\rho\log\rho - d\rho\} = \mathrm{Tr}\{\rho\, d\log\rho - d\rho\} \end{aligned} \tag{4.5}$$

Now we are tempted to use $d\log\rho = d\rho/\rho$ but we can not, because $d\rho$ and $1/\rho$ do not commute. So we will use the fact that every Hermitian operator is diagonalized by a unitary transformation:

$$\rho = U\rho_0 U^\dagger \tag{4.6}$$

and the fact that, being the logarithm an analitic function, holds the following equality:

$$\log\left(BAB^{-1}\right) = B\log A\, B^{-1} \tag{4.7}$$

For any invertible matrix B and semi positive-definite matrix A.

$$
\begin{aligned}
i_D^*(d_1\widetilde{S}_{VN}(\rho,\sigma)) &= \mathrm{Tr}\big\{U\rho_0U^\dagger d\log\big(U\rho_0U^\dagger\big) - d(U\rho_0U^\dagger)\big\} \\
&= \mathrm{Tr}\big\{U\rho_0U^\dagger d(U\log\rho_0U^\dagger) - d(U\rho_0U^\dagger)\big\} \quad (4.8)
\end{aligned}
$$

And now, using the fact that $d(U^\dagger) = -U^\dagger dUU^\dagger$, the Leibniz rule and the cyclic property of the trace one gets:

$$
\begin{aligned}
i_D^*(d_1\widetilde{S}_{VN}(\rho,\sigma)) &= \mathrm{Tr}\{\rho_0U^\dagger dU\log\rho_0 + \rho_0 d\log\rho_0 - \rho_0\log\rho_0U^\dagger dU \\
&- U^\dagger dU\rho_0 - d\rho_0 + \rho_0U^\dagger dU\} = \mathrm{Tr}\{\rho_0 d\log\rho_0 - d\rho_0\} \quad (4.9)
\end{aligned}
$$

Now we can use $d\log\rho_0 = d\rho_0/\rho_0$, because they commute, and we find that:

$$
i_D^*(d_1\widetilde{S}_{VN}(\rho,\sigma)) = \mathrm{Tr}\{\rho_0 d\log\rho_0 - d\rho_0\} = \mathrm{Tr}\big\{\rho_0\rho_0^{-1}d\rho_0 - d\rho_0\big\} = 0 \quad (4.10)
$$

Now we can calculate the metric tensor induced by this divergence function with the same technique we used in the last calculation:

$$
\begin{aligned}
i_D^*(d_1d_2\widetilde{S}_{VN}(\rho,\sigma)) &= i_D^*\,\mathrm{Tr}\{d\rho\otimes d\log\sigma\} \\
&= i_D^*\,\mathrm{Tr}\big\{d(U\rho_0U^\dagger)\otimes d(V\log\sigma_0V^\dagger)\big\} = i_D^*\,\mathrm{Tr}\{(dU\rho_0U^\dagger + Ud\rho_0U^\dagger - \\
&U\rho_0U^\dagger dUU^\dagger)\otimes(dV\log\sigma_0V^\dagger + Vd\log\sigma_0V^\dagger - V\log\sigma_0V^\dagger dVV^\dagger)\} \\
&= \mathrm{Tr}\{U^\dagger dU\rho_0\otimes U^\dagger dU\log\rho_0 - U^\dagger dU\rho_0\otimes\log\rho_0U^\dagger dU \\
&+ \rho_0U^\dagger dU\otimes\log\rho_0U^\dagger dU - \rho_0U^\dagger dU\otimes U^\dagger dU\log\rho_0 + d\rho_0\otimes d\log\rho_0\} \\
&= \mathrm{Tr}\big\{\big[U^\dagger dU,\rho_0\big]\otimes\big[U^\dagger dU,\log\rho_0\big] + d\rho_0\otimes d\log\rho_0\big\} \quad (4.11)
\end{aligned}
$$

This metric tensor can be written in this way:

$$
g_{VN} = g_U + g_{FR} \quad (4.12)
$$

Where:

$$g_{FR} = \mathrm{Tr}\{d\rho_0 \otimes d\log\rho_0\} \tag{4.13}$$

$$g_U = \mathrm{Tr}\{[U^\dagger dU, \rho_0] \otimes [U^\dagger dU, \log\rho_0]\} \tag{4.14}$$

So what we found is that the metric splits in two parts, one is just Fisher-Rao metric on the diagonal part, and the other is a metric on $U(n)$ that is written in terms of the Maurer-Cartan one form of $U(n)$. Now we will write explicitly the metric obtained from Von Neumann-Umegaki relative entropy in the case of a qubit.

Example 6. *A generic diagonal density state is written in the following form:*

$$\rho_0 = \begin{pmatrix} \frac{1+w}{2} & 0 \\ 0 & \frac{1-w}{2} \end{pmatrix} = \frac{1}{2}(\mathbb{I} + w\sigma_3) \qquad w \in [-1;1] \tag{4.15}$$

So we have:

$$\begin{aligned}\mathrm{Tr}\{d\rho_0 \otimes d\log\rho_0\} &= \mathrm{Tr}\left\{d\left(\frac{1}{2}(\mathbb{I}+w\sigma_3)\right) \otimes d\left(\log\frac{1}{2}(\mathbb{I}+w\sigma_3)\right)\right\} \\ &= \frac{1}{2}\mathrm{Tr}\{\sigma_3 dw \otimes (\mathbb{I}+w\sigma_3)^{-1}\sigma_3 dw)\} = \frac{1}{2}\mathrm{Tr}\{(\mathbb{I}+w\sigma_3)^{-1}\}dw \otimes dw \\ &= \frac{1}{1-w^2}dw \otimes dw\end{aligned} \tag{4.16}$$

It is a well known result [34] that Maurer-Cartan one form can be written in the following form:

$$U^\dagger dU = \frac{i}{2}\sigma_j\theta^j \tag{4.17}$$

where the θ^j are the elements of a basis of left invariant one forms on $U(2)$. This is the Fisher-Rao part of the metric. For the other term we have:

$$\left[U^{\dagger}dU,\rho_0\right]\otimes\left[U^{\dagger}dU,\log\rho_0\right]=-\frac{1}{4}[\sigma_j,\rho_0][\sigma_k,\log\rho_0]\theta^j\otimes\theta^k \tag{4.18}$$

Clearly:

$$[\sigma_3,\rho_0]=[\sigma_3,\log\rho_0]=0 \tag{4.19}$$

After a tedious but straightforward calculation one gets:

$$\mathrm{Tr}\{[\sigma_1,\rho_0][\sigma_1,\log\rho_0]\}=2w\log\frac{1-w}{1+w} \tag{4.20}$$

$$\mathrm{Tr}\{[\sigma_2,\rho_0][\sigma_2,\log\rho_0]\}=2w\log\frac{1-w}{1+w} \tag{4.21}$$

$$\mathrm{Tr}\{[\sigma_2,\rho_0][\sigma_1,\log\rho_0]\}=0 \tag{4.22}$$

$$\mathrm{Tr}\{[\sigma_1,\rho_0][\sigma_2,\log\rho_0]\}=0 \tag{4.23}$$

So one gets:

$$g=\frac{w}{2}\log\frac{1+w}{1-w}(\theta^1\otimes\theta^1+\theta^2\otimes\theta^2)+\frac{1}{1-w^2}dw\otimes dw \tag{4.24}$$

This tensor is usually called Bogoliubov-Kubo-Mori metric.
We can notice here that this metric is singular when $w\to 1$ *and when* $w\to -1$*, that is when one goes to pure states. Notice that this is a metric on states, but the first term is not a metric on* $\mathcal{U}(2)$*, since it is degenerate, we can easily see this by noticing that it lacks any term involving* θ^3*.*

Now we can obtain again Shannon relative entropy if we take the pullback via the immersion I of the Von Neumann-Umegaki relative entropy.
Let

$$I(\boldsymbol{p})=\rho_0 \qquad I(\boldsymbol{q})=\sigma_0 \tag{4.25}$$

Then we have that

$$I^* S_{VN} = I^*(\mathrm{Tr}\{\rho(\log\rho - \log\sigma)\}) = \sum_{j=1}^{n} p_j(\log p_j - \log q_j) \tag{4.26}$$

Then we can take the pullback via I also of the metric (4.12), and clearly in this process the part along the orbit of the unitary group vanishes, and we have:

$$\begin{split} I^* g_{VN} = I^*\big(\mathrm{Tr}\big\{\big[U^\dagger dU, \rho_0\big] \otimes \big[U^\dagger dU, \log\rho_0\big] \\ + d\rho_0 \otimes d\log\rho_0\big\}\big) = \sum_{j=1}^{n} dp_j \otimes d\log p_j \end{split} \tag{4.27}$$

And this is just Fisher-Rao metric.

So we have the following commutative diagram:

$$\begin{array}{ccc} S_{VN} & \xrightarrow{I^*} & S_{SH} \\ \big\downarrow {\scriptstyle i_D^* d_1 d_2} & & \big\downarrow {\scriptstyle i_D^* d_1 d_2} \\ g_{VN} & \xrightarrow{I^*} & g_{FR} \end{array}$$

Notice that this is not trivial, because doesn't fall under the case of the commutation of the exterior derivative with the pullback of a map, since $d_1 d_2$ is not the exterior derivative of $\mathcal{D}(\mathcal{H}) \times \mathcal{D}(\mathcal{H})$.

4.2 From classical to quantum metric tensors

In section 1.2 we described the state of density states $\mathcal{D}(\mathcal{H})$ on the Hilbert space $\mathcal{H}$, and saw that its bulk coincides with the space $\mathcal{D}_n(\mathcal{H})$ of invertible density states (where n is the complex dimension of $\mathcal{H}$).
In section 3.4 we have used coarse graining maps to introduce the monotonicity property of divergence functions and metric tensors. Now we want to do the same in the quantum setting, so we need to identify the class of *quantum stochatic maps*. From now on and until next chapter we will denote with $\mathcal{D}_n$ the state of invertible density states over an Hilbert space of complex dimension n. Let $\mathcal{D}_l$ and $\mathcal{D}_m$ be two such spaces and let us give the definition of quantum stochatic maps:

Definition 5. *Let $\phi : \mathcal{D}_l \longrightarrow \mathcal{D}_m$ be a completely positive trace preserving[1] map, if:*

$$\phi(\mathcal{D}_l) \subseteq \mathcal{D}_m \tag{4.28}$$

then ϕ is a quantum stochastic map.

Now we will give definitions of the monotonicity property for metric tensors in the quantum setting.
Let us consider the family $\{\mathcal{D}_k\}$ of spaces of invertible density states and a family $\{g_k\}$ of metric tensors, where k can be any natural number except 0 and 1 and g_k is defined on $\mathcal{D}_k$. Then:

Definition 6. *The family $\{g_k\}$ of metric tensors satisfies the monotonicity property if:*

$$g_l(X, X) \geqslant (\phi^* g_m)(X, X) \tag{4.29}$$

for all X in $\mathfrak{X}(\mathcal{D}_l)$ and for all quantum stochastic maps ϕ from $\mathcal{D}_l$ to $\mathcal{D}_m$.

Now we will give the definition of the monotonicity property also for divergence functions, in the quantum case this property is usually called **data processing inequality** (DPI).

[1] For a definition of completely positive maps see [21] pg. 66,67. for motivations of this choice see [30]

Let us consider a family of divergence functions $\{F_k\}$ where as before k is a natural number stricty larger than one and each function F_k is defined on $\mathcal{D}_k \times \mathcal{D}_k$. Then:

Definition 7. *If the family $\{F_k\}$ satisfies:*

$$F_l(\rho, \sigma) \geqslant F_m(\phi(\rho), \phi(\sigma)) \tag{4.30}$$

For all ρ, σ in $\mathcal{D}_l$ and for all quantum stochastic maps ϕ from $\mathcal{D}_l$ to $\mathcal{D}_m$, then we say that it satisfies the DPI.

The content of this definition is analogous the content of its classical counterpart (definition 3 in section 3.4), that is we demand that the information encoded in the divergence function does not increase under stochastic maps.
Now what one can prove [16] is the following proposition:

Proposition 1. *If the family of divergence functions $\{F_k\}$ satisfies the DPI, than the family of metric tensors $\{g_k\}$ obtained from these divergence functions satisfies the monotonicity property.*

In this proposition when we say that $\{g_k\}$ is obtained from $\{F_k\}$ we mean that the procedure described in section 3.2 is applied (in the appropriate space) to every element of the family $\{F_k\}$.
At this point the picture seems quite similar to the classical case, but there is a crucial difference, there is no analogue for Chentsov theorem in the quantum case, that is we don't have a unique metric that satisfies the monotonicity property in the quantum case.
So we have different function satisfying the monotonicity property that give rise to metrics that give rise to monotonicity property but we don't have a criterion to choose one funtion amongst the others, or one metric.
In fact in Quantum Information Theory there exist entire families of quantum divergence function that generate families of quantum metrics, for example one can consider relative Tsallis entropy, as done in [32], or the $(q-z)$-Rényi relative entropies, as in [11].

4.3 Metric tensors on orbits of Lie groups

In this section we will see how we can construct metric tensors on submanifolds of the full space of quantum states, in the particular case that this submanifold is an orbit of the action of a Lie group on the space of quantum states.
Imagine that we define the action of a Lie group G on our space of quantum states $\mathcal{S}$:

$$A : G \times \mathcal{S} \ni (g, \rho) \longmapsto A_g(\rho) \in \mathcal{S} \tag{4.31}$$

in the following we will denote this action just with $g\rho$.
Let ρ_0 be a state of $\mathcal{S}$, the orbit of the action of G on ρ_0 will be a submanifold $\mathcal{O}_{\rho_0}$. Let us define the following immersion:

$$i_G : G \ni g \longmapsto g\rho_0 \in \mathcal{S} \tag{4.32}$$

clearly if e is the neutral element of G we have that $\rho_e = \rho_0$, the image of this immersion will be just $\mathcal{O}_{\rho_0}$.
So now we can work on the Lie group G in an analogous way as we have done in section 2.4. Given a function $F \in \mathcal{F}(\mathcal{S} \times \mathcal{S})$ we can obtain a function on $G \otimes G$ by taking the pullback of F via i_G.
Then the construction of the metric tensor follows as described in section 3.2:

$$g(X, Y) = i_D^*(d_1 d_2 F)(X_l, Y_r) \tag{4.33}$$

and the exterior derivatives can be performed on the Lie group, while pullback of the diagonal immersion basically brings us from $G \times G$ to G. Let us make an example of this procedure.

Example 7. *On $\mathcal{H}$ we can obtain Fubini-Study metric from the following divergence function:*

$$F = 1 - \frac{|\langle\psi|\phi\rangle|^2}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle} \tag{4.34}$$

Where $|\psi\rangle$ and $|\phi\rangle$ are two vectors of $\mathcal{H}$. Let us compute its restriction to the orbit of a Lie group G.
This means that we consider $|\psi\rangle$ and $|\phi\rangle$ in (4.34) belonging to the orbit of G through the "fiducial" vector $|\psi_0\rangle$, namely, we are assuming that they can be both obtained by acting on $|\psi_0\rangle$ with two elements, say g and h, in G:

$$|\psi\rangle = g\,|\psi_0\rangle \qquad |\phi\rangle = h\,|\psi_0\rangle \qquad g,h \in G \tag{4.35}$$

This allows us to write F as a function on $G \times G$:

$$F = 1 - \frac{\langle\psi_0|\, g^\dagger h\,|\psi_0\rangle\langle\psi_0|\, h^\dagger g\,|\psi_0\rangle}{\langle\psi_0|\, h^\dagger h\,|\psi_0\rangle\langle\psi_0|\, g^\dagger g\,|\psi_0\rangle} \tag{4.36}$$

Then with the following positions:

$$N = \langle\psi_0|\, g^\dagger h\,|\psi_0\rangle\langle\psi_0|\, h^\dagger g\,|\psi_0\rangle \tag{4.37}$$

$$D = \langle\psi_0|\, h^\dagger h\,|\psi_0\rangle\langle\psi_0|\, g^\dagger g\,|\psi_0\rangle \tag{4.38}$$

We have:

$$d_1 d_2 F = -\frac{(D d_1 d_2 N + d_1 D d_2 D - d_1 N d_2 D - N d_1 d_2 D)D}{D^3} - \frac{d_1(D^2)(D d_2 N - N d_2 D)}{D^3} \tag{4.39}$$

Let us compute some terms:

$$d_1 N = \langle\psi_0| dg^\dagger h |\psi_0\rangle\langle\psi_0| h^\dagger g |\psi_0\rangle + \langle\psi_0| g^\dagger h |\psi_0\rangle\langle\psi_0| h^\dagger dg |\psi_0\rangle \tag{4.40}$$

$$d_2 N = \langle\psi_0| g^\dagger dh |\psi_0\rangle\langle\psi_0| h^\dagger g |\psi_0\rangle + \langle\psi_0| g^\dagger h |\psi_0\rangle\langle\psi_0| dh^\dagger g |\psi_0\rangle \tag{4.41}$$

$$d_1 D = \langle\psi_0| h^\dagger h |\psi_0\rangle \left(\langle\psi_0| dg^\dagger g |\psi_0\rangle + \langle\psi_0| g^\dagger dg |\psi_0\rangle\right) \tag{4.42}$$

$$d_2 D = \langle\psi_0| g^\dagger g |\psi_0\rangle \left(\langle\psi_0| dh^\dagger h |\psi_0\rangle + \langle\psi_0| h^\dagger dh |\psi_0\rangle\right) \tag{4.43}$$

$$d_1 d_2 D = \frac{d_1 D \otimes d_2 D}{D} \tag{4.44}$$

$$\begin{aligned} d_1 d_2 N = {} & \langle\psi_0| dg^\dagger dh |\psi_0\rangle\langle\psi_0| h^\dagger g |\psi_0\rangle + \langle\psi_0| g^\dagger dh |\psi_0\rangle\langle\psi_0| h^\dagger dg |\psi_0\rangle \\ & + \langle\psi_0| dh^\dagger dg |\psi_0\rangle\langle\psi_0| g^\dagger h |\psi_0\rangle + \langle\psi_0| dg^\dagger h |\psi_0\rangle\langle\psi_0| dh^\dagger g |\psi_0\rangle \end{aligned} \tag{4.45}$$

Then combining all of this result and taking the pullback of the resulting tensor via the diagonal immersion we get:

$$i_D^*(h) = -2\left(\frac{\langle\psi_0| dg^\dagger \otimes dg |\psi_0\rangle}{\langle\psi_0| g^\dagger g |\psi_0\rangle} - \frac{\langle\psi_0| dg^\dagger g |\psi_0\rangle \otimes \langle\psi_0| g^\dagger dg |\psi_0\rangle}{\langle\psi_0| g^\dagger g |\psi_0\rangle^2}\right) \tag{4.46}$$

That is just (-2 times) the Fubini-Study metric restricted to the orbit of the group G [5].
Notice that in the case that G is just the unitary group $\mathcal{U}(\mathcal{H})$ we have:

$$i_D^*(h) = -2\left(\frac{\langle\psi_0| dU^\dagger \otimes dU |\psi_0\rangle}{\langle\psi_0| U^\dagger U |\psi_0\rangle} - \frac{\langle\psi_0| dU^\dagger U |\psi_0\rangle \otimes \langle\psi_0| U^\dagger dU |\psi_0\rangle}{\langle\psi_0| U^\dagger U |\psi_0\rangle^2}\right) \tag{4.47}$$

Noting that $U^\dagger = U^{-1}$ this expression can be rewritten as:

$$i_D^*(h) = -2\Bigg(\frac{\langle\psi_0|\, dU^\dagger U \otimes U^\dagger dU\, |\psi_0\rangle}{\langle\psi_0|\psi_0\rangle} - \frac{\langle\psi_0|\, dU^\dagger U\, |\psi_0\rangle \otimes \langle\psi_0|\, U^\dagger dU\, |\psi_0\rangle}{\langle\psi_0|\psi_0\rangle^2}\Bigg) \quad (4.48)$$

So it is written in terms of the Maurer-Cartan form of the unitary group.

At the end of section 3.2 we also pointed out that we can obtain the same metric tensor (with a factor -4) if we use the derivatives d and d_J instead of d_1 and d_2:

$$i_D^* d_1 d_2 F(X_l, Y_r) = -4 i_D^* d d_J F(JX, Y) \quad (4.49)$$

In order to do this in this setting we need to define a complex structure on $G \times G$. A "privileged" choice for the complex structure on $G \times G$ could be the following:

$$J = \tilde{X}^j \otimes \theta_j - X^j \otimes \tilde{\theta}_j \quad (4.50)$$

Where X^j and $\tilde{X}^j$ are the elements of a basis of left-invatiant vector fields respectively on $\mathfrak{X}(G) \oplus \{0\}$ and on $\{0\} \oplus \mathfrak{X}(M)$, while θ_j and $\tilde{\theta}_j$ are the elements of a basis of left-invariant one forms respectively on $\Omega^1(G) \oplus \{0\}$. and on $\{0\} \oplus \Omega^1(G)$. Now we will give an example of this kind of procedure to obtain a metric tensor.

Example 8. *In the first place let us show that Gaussian states can be obtained acting with the affine group $Aff(\mathbb{R})$ on the state:*

$$\psi_0(x) = e^{-x^2} \quad (4.51)$$

In fact let us consider an affine transformation on the coordinates:

$$\chi(x) = ax + b \tag{4.52}$$

we have that:

$$\psi_0(\chi(x)) = e^{-a^2(x+a^{-1}b)^2} \tag{4.53}$$

That is just a Gaussian state of mean value $-a^{-1}b$ *and variance* $1/a$*. It is clear that every Gaussian state can be written as a certain affine transformation acting on the state* ψ_0*. But to make better contact with the usual way gaussian states are written we will use not the transformation* χ *but its inverse:*

$$\chi^{-1}(x) = a^{-1}x - a^{-1}b \tag{4.54}$$

So that we have:

$$\psi_0(\chi^{-1}(x)) = e^{-\frac{(x-b)^2}{a^2}} \tag{4.55}$$

In this way the mean value is b *and the variance is* a*.*
This provides us with the following immersion:

$$i_{Aff(\mathbb{R})} : \mathbb{R}_0 \otimes \mathbb{R} \ni (a, b) \longrightarrow \psi_{a,b} \in L^2(\mathbb{R}) \tag{4.56}$$

Now let us construct the complex structure on $Aff(\mathbb{R}) \times Aff(\mathbb{R})$*: it is easy to check that a basis for the left invariant vector fields on* $Aff(\mathbb{R})$ *is given by:*

$$X^1 = a\frac{\partial}{\partial a} \tag{4.57}$$

$$X^2 = a\frac{\partial}{\partial b} \tag{4.58}$$

while a basis for left invariant one forms on $Aff(\mathbb{R})$ *is given by:*

$$\theta_1 = a^{-1} da \tag{4.59}$$

$$\theta_2 = a^{-1} db \tag{4.60}$$

So the complex structure on $Aff(\mathbb{R}) \times Aff(\mathbb{R})$ is:

$$J = \tilde{a}a^{-1}(\frac{\partial}{\partial \tilde{a}} \otimes da + \frac{\partial}{\partial \tilde{b}} \otimes db) - a\tilde{a}^{-1}(\frac{\partial}{\partial a} \otimes d\tilde{a} + \frac{\partial}{\partial b} \otimes d\tilde{b}) \tag{4.61}$$

Now we want to compute a metric tensor from the divergence function:

$$F = 1 - \frac{|\langle\psi|\phi\rangle|^2}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle} \tag{4.62}$$

We will proceed taking in the first place the pullback of F via $i_{Aff(\mathbb{R})}$ and then acting on this function with dd_J:

$$i^*_{Aff(\mathbb{R})}F = 1 - \frac{2e^{-2\frac{(b-\tilde{b})^2}{a^2+\tilde{a}^2}}}{a\tilde{a}(1/a^2 + 1/\tilde{a}^2)} \tag{4.63}$$

Then we have:

$$d_J(i^*_{Aff(\mathbb{R})}F) = \tilde{a}a^{-1}\left(\frac{\partial F}{\partial \tilde{a}}da + \frac{\partial F}{\partial \tilde{b}}db\right) - a\tilde{a}^{-1}\left(\frac{\partial F}{\partial a}d\tilde{a} + \frac{\partial F}{\partial b}d\tilde{b}\right) \tag{4.64}$$

Now we should take the exterior derivative of this one form, then make it act on JX and Y:

$$X = X_a\frac{\partial}{\partial a} + X_b\frac{\partial}{\partial b} + X_{\tilde{a}}\frac{\partial}{\partial \tilde{a}} + X_{\tilde{b}}\frac{\partial}{\partial \tilde{b}} \tag{4.65}$$

$$X = Y_a\frac{\partial}{\partial a} + Y_b\frac{\partial}{\partial b} + Y_{\tilde{a}}\frac{\partial}{\partial \tilde{a}} + Y_{\tilde{b}}\frac{\partial}{\partial \tilde{b}} \tag{4.66}$$

$$JX = \tilde{a}a^{-1}\left(X_a\frac{\partial}{\partial \tilde{a}} + X_b\frac{\partial}{\partial \tilde{b}}\right) - a\tilde{a}^{-1}\left(X_{\tilde{a}}\frac{\partial}{\partial a} + X_{\tilde{b}}\frac{\partial}{\partial b}\right) \tag{4.67}$$

and then take the diagonal immersion The calculations are really long but straightforward, so let us jump to the results:

$$i_D^*(dd_J i_{Aff(\mathbb{R})}^* F(JX,Y)) = \frac{4}{a^2}X_a Y_a + \frac{8}{a^2}X_b Y_b \tag{4.68}$$

Let us notice that this tensor is 8 times (2.77)*, in example 3, if we put* $\phi = 0$*. This is in complete agreement with the results of section 3.2 (in particular equation* (3.18)*) and of the previous example.*
The fact that the imaginary part doesn't appear here is clearly because of the fact that we have chosen to work with real functions from the start, so it is no surprise.

Chapter 5

The groupoid viewpoint

In this chapter we will discuss a recent proposal [13] [14] of a new picture of Quantum Mechanics based on an approach due to Schwinger [38]. Schwinger's approach was based on the concept of selective measurement, let us describe it briefly.

Given a physical system, we can consider an *ensemble* $\mathcal{E}$[1] associated with it, then we have a family of observables $\mathbf{A}$ that represent physical quantities, so every $A \in \mathbf{A}$ has some possible outcomes $a \in \mathbb{R}$ when A is measured on a physical system $S \in \mathcal{E}$. Two observables A_1 and A_2 are said to be compatible if the outcomes of their respective measurements are not affected by the outcomes of the other, a set $\mathbf{A}$ of observables is said to be compatible if every couple of observables in $\mathbf{A}$ is compatible.

Consider a family of compatible observables $\mathbf{A} = \{A_1, A_2, \ldots, A_n\}$ and a collection $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ of outcomes respectively of $A_1, A_2, \ldots A_n$. A selective measurement $M(\mathbf{a})$ is a process that rejects all elements S of the ensemble $\mathcal{E}$ whose outcomes are different from $\mathbf{a}$. So we will write $M(\mathbf{a})S = S$ if the result of the measurement of $\mathbf{A}$ on S gives the outcome $\mathbf{a}$, and $M(\mathbf{a})S = \emptyset$ in the case that this measurement gives another outcome.

For the sake of clarity. let us stress here that we are referring to non-destructive measurements so, if we take the example of Stern-Gerlach measurements, we consider only the interaction with the magnetic field without the act of registering

[1] See [36] pg. 45,46 for a definition of ensemble, that we omit here for sake of brevity

the outcome by letting the particle hit a screen.
Then we can define another class of measurement $M(\mathbf{a}', \mathbf{a})$, this is a measurement that first rejects all systems whose outcomes are not $\mathbf{a}$ and then transforms the accepted systems in such a way that their outcomes are $\mathbf{a}'$.
Now consider the case that we want to take two consecutive measurements of this kind, to describe this situation we can define a composition law between measurements that has the meaning of performing first the measurement on the right and then the measurement on the left, let $M(\mathbf{a}', \mathbf{a})$ and $M(\mathbf{a}'', \mathbf{a}')$ be two selective measurements, then:

$$M(\mathbf{a}'', \mathbf{a}') \circ M(\mathbf{a}', \mathbf{a}) = M(\mathbf{a}'', \mathbf{a}) \tag{5.1}$$

clearly only some couples of measurements can be composed, specifically $M(\mathbf{a}''', \mathbf{a}'')$ can be composed with $M(\mathbf{a}', \mathbf{a})$ iff $\mathbf{a}'' = \mathbf{a}'$.
We can also compose together the two kind of measurements introduced:

$$M(\mathbf{a}') \circ M(\mathbf{a}', \mathbf{a}) = M(\mathbf{a}', \mathbf{a}) \qquad M(\mathbf{a}', \mathbf{a}) \circ M(\mathbf{a}) = M(\mathbf{a}', \mathbf{a}) \tag{5.2}$$

The composition law beetwen selective transformation, when it can be performed, is clearly associative:

$$(M(\mathbf{a}''', \mathbf{a}'') \circ M(\mathbf{a}'', \mathbf{a}')) \circ M(\mathbf{a}', \mathbf{a}) = M(\mathbf{a}''', \mathbf{a}'') \circ (M(\mathbf{a}'', \mathbf{a}') \circ M(\mathbf{a}', \mathbf{a})) \tag{5.3}$$

Moreover we will assume that if exists the selective measurement $M(\mathbf{a}', \mathbf{a})$, then exists another measurement $M(\mathbf{a}, \mathbf{a}')$ such that:

$$M(\mathbf{a}', \mathbf{a}) \circ M(\mathbf{a}, \mathbf{a}') = M(\mathbf{a}') \qquad M(\mathbf{a}, \mathbf{a}') \circ M(\mathbf{a}', \mathbf{a}) = M(\mathbf{a}) \tag{5.4}$$

Now that we have briefly descripted the premises of Schwinger's approach we can start our discussion.
In the first section of this chapter we will develop the theory of groupoids we need in order to understand the approach depicted in [12], [13] and [14]. In the second

section we will see how this links to Schwinger's idea we briefly descripted and finally in a third section we will see its link with our work.

5.1 Groupoids

Roughly speaking a groupoid is a set with the properties of a group, except for the fact that the binary operation doesn't have to be defined for every couple of elements of the group. Now we will give a rigorous definition, but in order to do that we have to make a little deviation and give some definitions of Category Theory.

A category **C** consist of a family of objects that we will denote $Ob(\mathbf{C})$ and a family of morphisms, that we will denote $Mor(\mathbf{C})$:

$$Mor(\mathbf{C}) \ni \gamma : x \mapsto y \quad x, y \in Ob(\mathbf{C}) \tag{5.5}$$

We can define two maps from $Mor(\mathbf{C})$ two $Ob(\mathbf{C})$ called *source* and *target* that gives respectively the starting object and the arrival object of the morphism:

$$\gamma : x \mapsto y \tag{5.6}$$

$$s(\gamma) = x \quad t(\gamma) = y \tag{5.7}$$

The category is equipped with a composition law $\circ$ between morphisms:

$$\alpha : x \mapsto y \quad \beta : y \mapsto z \quad x, y, z \in Ob(\mathbf{C}) \tag{5.8}$$

$$\beta \circ \alpha : x \mapsto z \tag{5.9}$$

This composition law is defined on every pair of composable morphisms, that is when $t(\alpha) = s(\beta)$, this composition law is also associative.

Another requirement for **C** to be a category is that there exist a family of morphisms $\mathbb{I}_x$ such that $\alpha \circ \mathbb{I}_x = \alpha$ and $\mathbb{I}_y \circ \alpha = \alpha$, where α is defined as before.

A category can be denoted as:

$$\begin{array}{c} Mor(\mathbf{C}) \\ s\Big\downarrow \Big\downarrow t \\ Ob(\mathbf{C}) \end{array}$$

Where the double arrows denote the assignments to each morphism $\alpha : x \mapsto y$ of the source object x and the target object y respectively.
A morphism $\alpha : x \mapsto y$ is said to be invertible if exists another morphism $\beta : y \mapsto x$, this will be denoted as α^{-1}.
A category is said to be small when his objects, family of objects and family of morphisms are sets, and their morphisms are maps between sets.
A groupoid $\mathbf{G}$ is a small category whose morphisms are all invertible.
Now before we introduce other elements in the discussion let us present a simple example:

Example 9. *Given a set Ω we can consider the set of pairs of elements of Ω, and this set has the structure of a groupoid. In fact, given a pair (x, y) with $x, y \in \Omega$, we can associate to it a morphism γ such that:*

$$s(\gamma) = x \qquad t(\gamma) = y \tag{5.10}$$

The composition rule will give:

$$(x, y) \circ (y, z) = (x, z) \tag{5.11}$$

This composition is clearly associative.
The unit morphisms are given by $\mathbb{I}_x = (x, x)$ and the inverse of γ will be given by $\gamma^{-1} = (y, x)$.

Now consider a finite groupoid $\mathbf{G}$, this means that $\mathbf{G}$ is made of a finite number of morphisms:

$$\mathbf{G} = \{\gamma_j | j = 1, \dots, N\} \tag{5.12}$$

Clearly also its set of objects Ω will be finite

$$\Omega = \{x_j | j = 1, \ldots, n\} \tag{5.13}$$

Now we can consider formal complex linear combinations of elements of **G**:

$$\mathbf{a} = \sum_{j=1}^{N} a_j \gamma_j \tag{5.14}$$

$$\mathbf{b} = \sum_{j=1}^{N} b_j \gamma_j \tag{5.15}$$

with $a_j, b_j \in \mathbb{C} \quad \forall j = 1, \ldots, N$.

Then we can define a product between these objects in the following way:

$$\mathbf{a} \cdot \mathbf{b} = \sum_{j,k=1}^{N} a_j b_k \delta(\gamma_j, \gamma_k) \gamma_j \circ \gamma_k \tag{5.16}$$

Where $\delta(\gamma_j, \gamma_k)$ is 1 whenever γ_j and γ_k are composable and 0 otherwise. This product gives to the set of objects like **a** and **b** the structure of an associative algebra (the associativity of the product defined by (5.16) comes from the associativity of the composition law between morphisms).This structure will be called *groupoid algebra* and will be indicated with the symbol $\mathbb{C}(\mathbf{G})$.Clearly a basis of $\mathbb{C}(\mathbf{G})$ will be given by the set of all transitions in **G**.

On $\mathbb{C}(\mathbf{G})$ one can also introduce an involution operator $* : \mathbb{C}(\mathbf{G}) \ni \mathbf{a} \mapsto \mathbf{a}^* \in \mathbb{C}(\mathbf{G})$ defined in the following way:

$$\mathbf{a}^* = \sum_{\gamma} \bar{a}_\gamma \gamma^{-1} \tag{5.17}$$

Moreover given a finite groupoid we can define the Hilbert space $\mathcal{H}_\Omega$ as the complex linear space with elements:

$$|\psi\rangle = \sum_{j=1}^{n} \psi_j |x_j\rangle \quad v_j \in \mathbb{C} \tag{5.18}$$

What we have done is associate every object $x_j \in \Omega$ with a linear space :

$$\Pi : \Omega \ni x_j \mapsto \mathbb{C} |x_j\rangle \tag{5.19}$$

Then we define an inner product in our Hilbert space:

$$\langle \psi, \phi \rangle = \sum_{j=1}^{n} \bar{\psi}_j \phi_j \quad \forall \ |\psi\rangle , |\phi\rangle \in \mathcal{H}_\Omega \tag{5.20}$$

With this definition then the vectors $|x_1\rangle , |x_2\rangle , \ldots , |x_n\rangle$ form an orthonormal basis of $\mathcal{H}_\Omega = \oplus_{j=1}^{n} \mathbb{C} |x_j\rangle$.
Then what about the morphisms of the groupoid? They can now be seen as maps between linear spaces: this means that we associate to the morphism $\gamma : x_j \mapsto x_k$ the linear map:

$$\Pi(\gamma) : \Pi(x_j) \mapsto \Pi(x_k) \tag{5.21}$$

$$\Pi(\gamma) |x_j\rangle = |x_k\rangle \tag{5.22}$$

This is what is called the fundamental representation of the groupoid **G**.
As one could expect, the support $\mathcal{H}_\Omega$ of this representation will be used to make contact with the usual Dirac-Schrödinger picture of Quantum Mechanics.
Notice that the fundamental representation can be applied also to elements of the groupoid algebra $\mathbb{C}(\mathbf{G})$ assuming complex linearity:

$$\Pi(\mathbf{a}) |\psi\rangle = \sum_{j,k} a_j \psi_k \Pi(\gamma_j) |x_k\rangle \tag{5.23}$$

If we introduce the symbol $\delta(\gamma_j, x_k)$ defined as 1 if $s(\gamma_j) = x_k$ and 0 in any other case we can recast the last definition in the following form:

$$\Pi(\mathbf{a}) |\psi\rangle = \sum_{j,k} a_j \psi_k \delta(\gamma_j, x_k) |t(\gamma_j)\rangle \tag{5.24}$$

Now we can take functions $f : \mathbf{G} \longrightarrow \mathbb{C}$ that associate a complex number to a transition. In the set of all such functions $\mathcal{F}(\mathbf{G})$ we can define a convolution product:

$$(f * g)(\gamma) = \sum_{\alpha \circ \beta = \gamma} f(\alpha) g(\beta) \tag{5.25}$$

Equipped with this product the set $\mathcal{F}(\mathbf{G})$ becomes an algebra (in general non-commutative). One can also define an involution:

$$f^*(\gamma) = \overline{f(\gamma^{-1})} \tag{5.26}$$

That makes it into a $*$-algebra.
A basis of this algebra will be given by the function δ_γ defined in the following way:

$$\delta_\gamma(\alpha) = \begin{cases} 1 & \text{if} \quad \alpha = \gamma \\ 0 & \text{if} \quad \alpha \neq \gamma \end{cases} \tag{5.27}$$

So that every element of $\mathcal{F}(\mathbf{G})$ can be written in the following way:

$$f = \sum_\gamma f(\gamma) \delta_\gamma \tag{5.28}$$

Where the summation is made on all transitions.
This definition shows clearly that the algebras $\mathbb{C}(\mathbf{G})$ and $\mathcal{F}(\mathbf{G})$ are dual.
We can also define a pairing $\langle ., . \rangle : \mathcal{F}(\mathbf{G}) \times \mathbb{C}(\mathbf{G}) \longrightarrow \mathbb{C}$ between this two sets, let:

$$\mathbf{a} = \sum_\gamma a_\gamma \gamma \quad \in \mathbb{C}(\mathbf{G}) \tag{5.29}$$

$$f = \sum_\gamma f(\gamma) \delta_\gamma \quad \in \mathcal{F}(\mathbf{G}) \tag{5.30}$$

Then:

$$\langle f, \mathbf{a} \rangle = \sum_\gamma a_\gamma f(\gamma) \tag{5.31}$$

Then we can associate to every element of $\mathbb{C}(\mathbf{G})$ an element of $\mathcal{F}(\mathbf{G})$ and vice-versa:

$$\mathbf{a}_f = \sum_\gamma f(\gamma)\gamma \tag{5.32}$$

$$f_{\mathbf{a}} = \sum_\gamma a_\gamma \delta_\gamma \tag{5.33}$$

With this identification one can easily prove the following identities:

$$\mathbf{a}_f \cdot \mathbf{a}_g = \mathbf{a}_{f*g} \tag{5.34}$$

$$f_{\mathbf{a}} * f_{\mathbf{b}} = f_{\mathbf{a}\cdot\mathbf{b}} \tag{5.35}$$

$$\mathbf{a}_f^* = \mathbf{a}_{f*} \tag{5.36}$$

$$f_{\mathbf{a}}^* = f_{\mathbf{a}}^* \tag{5.37}$$

We will prove the first and the third of these relations as an example, the other two proofs being analogous:

$$\mathbf{a}_f \cdot \mathbf{a}_g = \sum_{\gamma_1,\gamma_2} f(\gamma_1)g(\gamma_2)\delta(\gamma_1,\gamma_2)\gamma_1 \circ \gamma_2 = \sum_{\gamma_1\circ\gamma_2=\gamma} f(\gamma_1)g(\gamma_2)\gamma = \mathbf{a}_{f*g} \tag{5.38}$$

and this proves the first relation.
Now the third:

$$\mathbf{a}_f^* = (\sum_\gamma f(\gamma)\gamma)^* = \sum_\gamma \overline{f(\gamma)}\gamma^{-1} \tag{5.39}$$

$$\mathbf{a}_{f^*} = \sum_\gamma f^*(\gamma)\gamma = \sum_\gamma \overline{f(\gamma^{-1})}\gamma \tag{5.40}$$

And clearly the two expression are the same.
Now, using the correspondence we estabilished between elements in the groupoid algebra and functions defined on the groupoid let us define the action of the

fundamental representation also on elements of $\mathcal{F}(\mathbf{G})$:

$$\Pi(f)\,|x_j\rangle \equiv \Pi(\mathbf{a}_f)\,|x_k\rangle = \sum_{\gamma} f(\gamma)\delta(\gamma, x_k)\,|t(\gamma)\rangle \tag{5.41}$$

So the fundamental representation allows us to associate an endomorphism of $\mathcal{H}_\Omega$ to a function in $\mathcal{F}(\mathbf{G})$. For this reason we will also write $A_f = \Pi(f)$, moreover one gets:

$$A_{f*} = A_f^\dagger \tag{5.42}$$

Where the adjoint of an operator A is defined w.r.t. the canonical inner product (5.20). So the fundamental representation is a $*$-representation w.r.t. this inner product.

5.2 Groupoids and Quantum Mechanics

In the introduction of this chapter we briefly described Schwinger's approach, in this section we will see how the mathematical structures introduced in the previous section can be used to reformulate such approach.

In the first place let us consider the set of all possible outcomes of the family of compatible observables $\mathbf{A}$ and let us call it $\Omega_{\mathbf{A}}$, a selective measurements $M(\mathbf{a}', \mathbf{a})$ can be considered as a morphism between two elements $\mathbf{a}$ and $\mathbf{a}'$ of this set, we will call the set of all transitions $\mathbf{G_A}$. Equation (5.1) then defines a composition rule between morphisms that is not globally defined on $\mathbf{G_A}$, this composition rule, when it can be performed, is clearly associative by (5.3).

For every element $\mathbf{a} \in \Omega_{\mathbf{A}}$ we have the selective measurement $M(\mathbf{a})$ that is represented by the morphism that maps $\mathbf{a}$ into itself. As can be seen from (5.2) these morphisms represent left and right identities for the composition rule between transitions.

Moreover by (5.4) for every transition in $\mathbf{G_A}$ there exists an inverse transition, these considerations let us conclude that $\mathbf{G_A}$ can be seen as a groupoid with object space $\Omega_{\mathbf{A}}$.

We are actually neglecting an important aspect here, we could choose to use another set of compatible observables, let us call it $\mathbf{B}$, to describe the same quantum system, then transformations between these two should be taken care of. This can be done adding another layer to the groupoid structure, we introduce transformations as morphisms between transitions, so the resulting structure is a 2-groupoid, the outcomes are the objects of the groupoid $\mathbf{G_A}$, transitions are morphisms between outcomes but can also be considered as the objects of the groupoid $\Gamma_{\mathbf{A}}$ whose morphisms are transformations.

For details on this aspect see [13], we will skip these details and bring on our discussion using only one set of compatible observables.

Now we can use the fundamental representation of the groupoid $\mathbf{G_A}$ to construct the Hilbert space $\mathcal{H}_{\Omega_{\mathbf{A}}}$ and to associate to every function $f \in \mathcal{F}(\mathbf{G_A})$ an operator acting on it. A basis of $\mathcal{H}_{\Omega_{\mathbf{A}}}$ is given by the vectors $|\mathbf{a}\rangle$ associated to the elements of $\Omega_{\mathbf{A}}$.

Notice that, considering (5.42), real elements in $\mathcal{F}(\mathbf{G_A})$ (that is functions f s.t. $f^* = f$) are associated to self-adjoint operators acting on $\mathcal{H}_{\Omega_{\mathbf{A}}}$, from now on we will refer both to the real elements of $\mathcal{F}(\mathbf{G_A})$ and to the operators associated with them with the name of *observables*.

Let us evaluate the following quantity:

$$\langle \mathbf{a}', A_f \mathbf{a} \rangle = \langle \mathbf{a}' | \left(A_f | \mathbf{a} \rangle \right) = \sum_{\gamma} \langle \mathbf{a}' | f(\gamma) \delta(\gamma, \mathbf{a}) | t(\gamma) \rangle$$
$$= \sum_{\gamma} f(\gamma) \delta(\gamma, \mathbf{a}) \delta(\mathbf{a}', t(\gamma) = \sum_{\gamma: \mathbf{a} \mapsto \mathbf{a}'} f(\gamma) \quad (5.43)$$

So it is basically the sum of the values of f on all transitions that connect $\mathbf{a}$ to $\mathbf{a}'$.

In particular when f is an observable and for $\mathbf{a} = \mathbf{a}'$ we will get a real number that can be interpreted as the expected value of f on the state $|\mathbf{a}\rangle$.

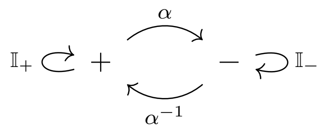
What we are doing here is using $\mathcal{H}_{\Omega_{\mathbf{A}}}$ as the Hilbert space of the Dirac-Schrödinger picture, alternatively one can define states as normalized positive functionals on $\mathcal{F}(\mathbf{G_A})$ and then construct the Hilbert space with the GNS construction [14].

Let us conclude this section with an example in order to better understand this construction:

Example 10. *Imagine that the family of observables* $\mathbf{A}$ *gives only two possible outcomes, that we will call* $+$ *and* $-$*, so we have* $\Omega = \{+, -\}$*. We have the transition* $\alpha : + \mapsto -$ *and its inverse* α^{-1} *together with the identities of the two objects:*

$$\mathbf{G_A} = \{\mathbb{I}_+, \mathbb{I}_-, \alpha, \alpha^{-1}\} \quad (5.44)$$

This groupoid can be represented by the following graph:



We can use the fundamental representation of this groupoid: $|+\rangle = \Pi(+)$ *and* $|-\rangle = \Pi(-)$*. Then the Hilbert space* $\mathcal{H}_\Omega$ *that supports this representation will be isomorphic to* $\mathbb{C}^2$*.*

The morphisms are represented by the following linear operator acting on $\mathcal{H}_\Omega$*:*

$$\Pi(\mathbb{I}_+)\,|+\rangle = |+\rangle \quad \Pi(\mathbb{I}_-)\,|-\rangle = |-\rangle \tag{5.45}$$

$$\Pi(\alpha)\,|+\rangle = |-\rangle \quad \Pi(\alpha^{-1})\,|-\rangle = |+\rangle \tag{5.46}$$

So we have that they are represented by 2×2 *matrices:*

$$\Pi(\mathbb{I}_+) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \Pi(\mathbb{I}_-) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{5.47}$$

$$\Pi(\alpha) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \Pi(\alpha^{-1}) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{5.48}$$

(5.49)

The groupoid algebra will be the (non-commutative) associative algebra generated by $\Pi(\mathbb{I}_+)$*,* $\Pi(\mathbb{I}_-)$*,* $\Pi(\alpha)$ *and* $\Pi(\alpha^{-1})$*, that we will denote with* e_1, e_2, e_3, e_4*.*

We have the following relations:

$$e_1^2 = e_1 \qquad e_1e_2 = 0 \qquad e_1e_3 = 0 \qquad e_1e_4 = e_4 \tag{5.50}$$

$$e_2e_1 = 0 \qquad e_2^2 = e^2 \qquad e_2e_3 = e_3 \qquad e_2e_4 = 0 \tag{5.51}$$

$$e_3e_1 = e_3 \qquad e_3e_2 = 0 \qquad e_3^2 = 0 \qquad e_3e_4 = e_2 \tag{5.52}$$

$$e_4e_1 = 0 \qquad e_4e_2 = e_4 \qquad e_4e_3 = e_1 \qquad e_4^2 = 0 \tag{5.53}$$

So the groupoid algebra is given by $M_2(\mathbb{C})$ *with structure constants given by the previous relations.*
Notice that if we remove the transitions α *and* α^{-1} *we get the groupoid* $\mathbf{G}_C = \{\mathbb{I}_+, \mathbb{I}_-\}$*, it can be represented by the following graph:*

$$\mathbb{I}_+ \circlearrowleft + \qquad\qquad - \circlearrowright \mathbb{I}_-$$

The support of the fundamental representation is still $\mathcal{H}_\Omega \cong \mathbb{C}^2$ *but the groupoid algebra will be the two-dimensional Abelian algebra defined by the relations:*

$$e_1^2 = e^1 \qquad e_1e_2 = 0 \tag{5.54}$$

$$e_2e_1 = 0 \qquad e_2^2 = e_2 \tag{5.55}$$

Notice that ignoring these two transitions we are saying that experimental devices do not modify the system, so we are describing the classical bit.

5.3 Why groupoids?

In this section we will connect this formulation in terms of groupoids of Schwinger's Picture of Quantum Mechanics with the main theme of the work.
Let us take another look at the purification of mixed states in the case of the qubit, in section 1.3 we set the problem in a bundle picture. Now we want to see how this problem can be formulated in terms of groupoids.
To see this let us limit ourselves to pure states for the moment, that is, instead of taking $\mathcal{D}_n(\mathcal{H})$ as the base space, take $\mathcal{D}_1(\mathcal{H})$. Then in our base space there will be rank-one projectors and in the total space rank-one linear operators.
Let us recall here the projection maps π_1 and π_2 introduced in section 1.2, so for example we have:

$$\pi_1(|-\rangle\langle+|) = |+\rangle\langle-|-\rangle\langle+| = |+\rangle\langle+| = \rho_+ \tag{5.56}$$

$$\pi_2(|-\rangle\langle+|) = |-\rangle\langle+|+\rangle\langle-| = |-\rangle\langle-| = \rho_- \tag{5.57}$$

Where we are considering normalized vectors for simplicity.
Recalling example 10, one can notice that $|-\rangle\langle+|$ is just the image via the the fundamental representation of the transition α, what we have called $\Pi(\alpha)$. The projection map π_1 applied on it gives $\mathbb{I}_+$, the unit in the groupoid associated to the element $+$ of the space of objectsl, that is the unit of the groupoid associated to the source of α.
On the other hand we have that π_2 applied on $\Pi(\alpha)$ gives $\mathbb{I}_-$, the unit of the groupoid associated to the target of α.
Units of the groupoids can be clearly put in a one-to-one correspondence with the set of objects of the groupoid, and can be used to immerse the set of objects into the morphisms of the groupoid.
So the idea behind this simple example is the following: the purification procedure can be seen from the point of view of the Schwinger picture, and in this setting it is a lift from the units (or from the objects) of a groupoid to its morphisms.
Moreover the two projection map are respectively on the source and on the target

of the transition. This can give a criterion on the choice between the two, that was something lacking in the way we proposed them in section 1.2.

This represent a further development of the idea of going from probabilities to amplitudes, in fact we have seen in chapter 1 that this transition can be seen as a lift from the projective space $\mathcal{P}(\mathcal{H})$ to the Hilbert space $\mathcal{H}_0$, and now we have reformulated this transition in terms of a lift in the context of groupoids.

Clearly at this level we are "purifying" pure states, but this discussion can be repeated for mixed states, and the investigation of this approach can be fruitful and could be object of future work.

There is also another interesting aspect, given a statistical manifold M, we can construct the groupoids of pairs of points of M, like we have done in example 9. So a two-point function on the statistical manifold M can be seen as a one-point function of the groupoid of pairs of set of M. This possibility is discussed in the contest of Lie groupoids on [28].

Conclusions

In this work we studied the connection between Information Geometry and Geometric Quantum Mechanics. The *leitmotiv* is the transition from probabilities to amplitudes [23], starting from the motivation given in the introduction, namely: Fisher-Rao metric can be seen as a term of Fubini-Study metric.

In fact Fubini-Study metric can be split in two parts: one involving phases and one involving probabilities, and Fisher-Rao metric is just the term relative to probabilities.

In chapter 1 we saw how this transition from probabilities to amplitudes can be seen from the geometrical point of view as a lift in a bundle picture for quantum states.

The same idea was applied in chapter 4 to Information Geometry, and led us to the transition from Classical to Quantum Information Geometry. Clearly in order to do that we needed some background in Information Geometry, and this was given in chapter 3,

In chapter 2 we introduced some geometric structure in the bundle picture presented in 1, in fact we studied the Riemannian and symplectic structure of $\mathcal{H}_0$ also in relation with the bundle structure described in the first chapter.

Finally in the last chapter we introduced a formulation of Schwinger's approach to Quantum Mechanics in term of groupoids, this formulation is interesting by itself and offers several advantages [12] [13] [14], but for our purpose the more relevant part is the fact that it gives a natural setting for studying two-points functions, that is the pair groupoid [28].

Investigating these subjects we were able to formulate a certain number of interesting observations, like the ones contained in section 4.1, where is given a geometric description of the transition from Quantum to Classical Information Theory.

However, during the development of this work many more questions have been asked and are unanswered at the moment, let us enumerate some of them:

- At the end of chapter 1 we introduced the purification process and gave a geometric picture to describe it as a lift in an appropriate bundle, then what one could ask is: if a purification procedure assigns a certain path on the total space to an assigned path on the base space, can we find a connection that defines the tangent vectors of the path in the total space as horizontal? When this connection will be metric? If it is metric, how does this metric look like?

- In chapter 2 we saw the relevant role of the complex structure in defining the metric and the symplectic structure, we assumed implicitly that it satisfied Nijenhuis condition, but what if it doesn't? Or more generally, what is the role of Nijenhuis tensor in this setting?

- In section 2.3 we found with a straightforward procedure the Fubini-Study metric and underlined what are the essential objects needed to construct it. Is it possible to reproduce the same procedure to obtain a metric in the mixed state case? Will this procedure give rise to Bures metric [9] [42]?

- In section 3.2 we estabilished a contact point between two procedures, one typical of Information Geometry and one typical of Geometric Quantum Mechanics, to obtain metric tensors. What one could ask is: is it possible to estabilish such connection also for the third order tensors usually defined in Information Geometry?

The presence of so many interesting question that arise in this context signals that this work has hit an interesting spot, from which a large amount of work can be carried out in order to answer these and other questions.

Bibliography

[1] Ralph Abraham, Jerrold E Marsden, and Tudor Ratiu. *Manifolds, tensor analysis, and applications*, volume 75. Springer Science & Business Media, New York, 2012.

[2] AC Aitken and H Silverstone. Xv.—on the estimation of statistical parameters. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 61(2):186–194, 1942.

[3] Shun-ichi Amari. *Information geometry and its applications*, volume 194. Springer, New York, 2016.

[4] Shun-ichi Amari and Hiroshi Nagaoka. *Methods of information geometry*, volume 191. American Mathematical Society, Providence, RI, 2007.

[5] P Aniello, G Marmo, and GF Volkert. Classical tensors from quantum states. *International Journal of Geometric Methods in Modern Physics*, 6(03):369–383, 2009.

[6] Ingemar Bengtsson and Karol Życzkowski. *Geometry of quantum states: an introduction to quantum entanglement.* Cambridge university press, Cambridge, 2017.

[7] Michael Victor Berry. Quantal phase factors accompanying adiabatic changes. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 392(1802):45–57, 1984.

[8] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Springer Science & Business Media, New York, 2010.

[9] Donald 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.

[10] Nikolaj Nikolaevic Chentsov. *Statistical decision rules and optimal conclusions*. American Mathematical Society, Providence, RI, 1976.

[11] Florio M Ciaglia, Fabio Di Cosmo, Marco Laudato, Giuseppe Marmo, Fabio M Mele, Franco Ventriglia, and Patrizia 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.

[12] Florio M Ciaglia, Alberto Ibort, and Giuseppe Marmo. A gentle introduction to schwinger's formulation of quantum mechanics: The groupoid picture. *Modern Physics Letters A*, 33(20):1850122, 2018.

[13] Florio M. Ciaglia, Alberto Ibort, and Giuseppe Marmo. Schwinger's Picture of Quantum Mechanics I: Groupoids. *arXiv e-prints*, page arXiv:1905.12274, May 2019.

[14] Florio M Ciaglia, Alberto Ibort, and Giuseppe Marmo. Schwinger's picture of quantum mechanics ii: Algebras and observables. *arXiv preprint arXiv:1907.03883*, 2019.

[15] Florio M Ciaglia, Giuseppe Marmo, and Juan Manuel Pérez-Pardo. Generalised potential functions in differential geometry and information geometry. *arXiv preprint arXiv:1804.10414*, 2018.

[16] Florio Maria Ciaglia. *The space of Quantum States, a Differential Geometric Setting*. PhD thesis, Università degli Studi di Napoli Federico II, 2017.

[17] Imre Csiszár. Information-type measures of difference of probability distributions and indirect observation. *studia scientiarum Mathematicarum Hungarica*, 2:229–318, 1967.

[18] Paul Adrien Maurice Dirac. *The principles of quantum mechanics*. Oxford university press, Oxford, UK, 1981.

[19] J Dittmann and A Uhlmann. Connections and metrics respecting purification of quantum states. *Journal of Mathematical Physics*, 40(7):3246–3267, 1999.

[20] Jochen Dittmann and Gerd Rudolph. On a connection governing parallel transport along $2\times$ 2 density matrices. *Journal of Geometry and Physics*, 10(1):93–106, 1992.

[21] J. Dixmier. *C*-algebras*. North-Holland mathematical library. North-Holland, Amsterdam, 1982.

[22] Elisa Ercolessi, Giuseppe Marmo, and Giuseppe Morandi. From the Equations of Motion to the Canonical Commutation Relations. *Rivista del Nuovo Cimento*, 33(8-9), 2010.

[23] Paolo Facchi, Ravi Kulkarni, VI Man'ko, Giuseppe Marmo, ECG Sudarshan, and Franco Ventriglia. Classical and quantum fisher information in the geometrical formulation of quantum mechanics. *Physics Letters A*, 374(48):4801–4803, 2010.

[24] Guido Fubini. Sulle metriche definite da una forma hermitiana. *Istituto Veneto*, 63(2):502–513, 1904.

[25] Paolo Gibilisco, Fumio Hiai, and Dénes Petz. Quantum covariance, quantum fisher information, and the uncertainty relations. *IEEE Transactions on Information Theory*, 55(1):439–443, 2008.

[26] Paolo Gibilisco, Daniele Imparato, and Tommaso Isola. Uncertainty principle and quantum fisher information. ii. *Journal of mathematical physics*, 48(7):072109, 2007.

[27] Paolo Gibilisco and Tommaso Isola. Uncertainty principle and quantum fisher information. *Annals of the Institute of Statistical Mathematics*, 59(1):147–159, 2007.

[28] Katarzyna Grabowska, Janusz Grabowski, Marek Kuś, and Giuseppe Marmo. Lie groupoids in information geometry. *arXiv preprint arXiv:1904.00709*, 2019.

[29] Janusz Grabowski, Marek Kuś, and Giuseppe Marmo. Geometry of quantum systems: density states and entanglement. *Journal of Physics A: Mathematical and General*, 38(47):10217, 2005.

[30] Alexander S Holevo. *Statistical structure of quantum theory*, volume 67. Springer Science & Business Media, New York, 2003.

[31] Luigi Malagò and Giovanni Pistone. Combinatorial optimization with information geometry: The newton method. *Entropy*, 16(8):4260–4289, 2014.

[32] Vladimir I Man'ko, Giuseppe Marmo, Franco Ventriglia, and Patrizia Vitale. Metric on the space of quantum states from relative entropy. tomographic reconstruction. *Journal of Physics A: Mathematical and Theoretical*, 50(33):335302, 2017.

[33] Giuseppe Marmo and Alessandro Zampini. Kähler geometry on complex projective spaces via reduction and unfolding. *arXiv e-prints*, page arXiv:1809.09993, Sep 2018.

[34] Mikio Nakahara. *Geometry, topology and physics.* Institute of Physics, Bristol, 2003.

[35] Helen Kelsall Nickerson, Donald Clayton Spencer, and Norman Earl Steenrod. *Advanced calculus.* Van Nostrand, Princeton, New Jersey, 2011.

[36] Asher Peres. *Quantum theory: concepts and methods*, volume 57. Springer Science & Business Media, New York, 2006.

[37] C Radhakrishna Rao. Information and the accuracy attainable in the estimation of statistical parameters. In *Breakthroughs in statistics*, pages 235–247. Springer, New York, 1992.

[38] Julian Schwinger. The algebra of microscopic measurement. *Proceedings of the National Academy of Sciences of the United States of America*, 45(10):1542, 1959.

[39] Claude Elwood Shannon. A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423, 1948.

[40] Barry Simon. Holonomy, the quantum adiabatic theorem, and berry's phase. *Physical Review Letters*, 51(24):2167, 1983.

[41] Armin Uhlmann. A gauge field governing parallel transport along mixed states. *Letters in Mathematical Physics*, 21(3):229–236, 1991.

[42] Armin Uhlmann. The metric of bures and the geometric phase. In *Groups and related Topics*, pages 267–274. Springer, 1992.

[43] Hisaharu Umegaki. Conditional expectation in an operator algebra, iv (entropy and information). In *Kodai Mathematical Seminar Reports*, volume 14, pages 59–85. Department of Mathematics, Tokyo Institute of Technology, 1962.

[44] V. Vedral and M. B. Plenio. Entanglement measures and purification procedures. *Phys. Rev. A*, 57:1619–1633, Mar 1998.