Exponential families, Kähler geometry and quantum mechanics

Mathieu Molitor *
Department of Mathematics, Keio University
3-14-1, Hiyoshi, Kohoku-ku, 223-8522, Yokohama, Japan
e-mail: pergame.mathieu@gmail.com

Abstract
Exponential families are a particular class of statistical manifolds which are particularly important in statistical inference, and which appear very frequently in statistics. For example, the set of normal distributions, with mean \( \mu \) and deviation \( \sigma \), form a 2-dimensional exponential family.

In this paper, we show that the tangent bundle of an exponential family is naturally a Kähler manifold. This simple but crucial observation leads to the formalism of quantum mechanics in its geometrical form, i.e. based on the Kähler structure of the complex projective space, but generalizes also to more general Kähler manifolds, providing a natural geometric framework for the description of quantum systems.

Many questions related to this “statistical Kähler geometry” are discussed, and a close connection with representation theory is observed.

Examples of physical relevance are treated in details. For example, it is shown that the spin of a particle can be entirely understood by means of the usual binomial distribution.

This paper centers on the mathematical foundations of quantum mechanics, and on the question of its potential generalization through its geometrical formulation.

1 Introduction – summary

In the 70’s, it has been observed by Chernoff and Marsden [CM74] that the Schrödinger equation \( i \hbar \frac{d\psi}{dt} = H\psi \) is Hamiltonian with respect to the symplectic form coming from the imaginary part of the Hermitian scalar product of the Hilbert space \( \mathcal{H} \) of possible quantum states. Since then, this Hamiltonian view on quantum mechanics has been developed independently by several authors [CL84, CMP90, Hes84, Hes85, Kib79] and has led to a complete geometrization of the quantum formalism, entirely based on the Kähler properties of the complex projective space \( \mathbb{P}(\mathcal{H}) \). This reformulation, which is very elegant and complete, is now usually referred to as the geometrical formulation of quantum mechanics [AS99, CMP90].

The geometrical formulation was mainly motivated by the desire to generalize quantum mechanics, especially in view of quantum gravity. The basic idea is that, by geometrizing the quantum formalism, one frees it from its burdensome linearity and put it on a geometrical ground akin to Einstein’s theory of gravitation. Geometry is, in this regard, particularly “flexible”, and seems an appropriate setting for generalizations. For example, while it is not clear how to generalize Hilbert spaces, generalizations of the complex projective space is straightforward: instead of \( \mathbb{P}(\mathcal{H}) \), take an arbitrary Kähler manifold. Such possibilities have been discussed in [GH92, Hug95, Kaw, Kib79] and applications towards quantum gravity have been proposed in [AT03, MT04].

* Present address: Fakultät für Mathematik, Ruhr-Universität Bochum, Germany

1 Not to be confused with the geometric quantization of Kostant and Souriau [Kos70, Sou97]. In the geometrical formulation, the Hilbert space is considered as given, not as the result of a quantization scheme.
These proposals, however, are limited in their original scope by a severe limitation: the need to allow for a probabilistic interpretation. The latter, contrary to the quantum state space and dynamics, is extremely difficult to generalize in a purely geometrical context and is usually not even discussed. The reason is that in the geometrical formulation, all formulas related to probabilities rely on an expression of the form \( \cos^2 \left( d(\cdot, \cdot) \right) \), where \( d(\cdot, \cdot) \) is the geodesic distance on \( \mathbb{P}({\mathcal{H}}) \), expression which is clearly specific to \( \mathbb{P}({\mathcal{H}}) \) and which, consequently, cannot be generalized directly to arbitrary Kähler manifolds.

These difficulties –related to the probabilistic interpretation– address the following question: what is the link between the Kähler structure of \( \mathbb{P}(\mathbb{C}^n) \) and probabilities? This question, which is central in the present work, was already formulated in our previous paper \[\text{Molb}\] where an interesting, though puzzling connection with information geometry has been observed.

Let us recall, in this regard, that information geometry is a branch of statistics characterized by its use of differential geometrical techniques \[\text{AN00, MR93}\]. Its basic objects of study are statistical manifolds, i.e. manifolds whose points can be identified with probability density functions over some fixed measured space. For example, Gaussian distributions over \( \mathbb{R} \) form a 2-dimensional statistical manifold parameterized by the mean \( \mu \) and deviation \( \sigma \). In general –and this is what information geometry is about– a statistical manifold \( S \) possesses a rich geometry that encodes many of its statistical properties; it has a Riemannian metric \( h_F \), called Fisher metric, and a pair of dual affine connections \( \nabla^{(c)}, \nabla^{(m)} \), respectively called exponential connection and mixture connection, which can be used, for example, to give lower bounds in estimation problems (compare e.g. the Cramér-Rao inequality). Together, the triplet \( (h_F, \nabla^{(c)}, \nabla^{(m)}) \) forms what is called a dualistic structure, and it is probably the most important structure in information geometry.

Very little attention has been paid to dualistic structures outside the statistical community, but we can mention, in connection with \[\text{Molb}\], the work of Dombrowski\[\text{Dom62}\]. In a paper which already goes back to the 60’s \[\text{Dom62}\], Dombrowski shows that if a manifold \( M \) is endowed with a dualistic structure (\( M \) needs not be a statistical manifold here), then its tangent bundle \( TM \) becomes naturally, via a simple geometric construction, an almost Hermitian manifold\[\text{Dom62}\]. A direct consequence of Dombrowski’s construction, which seems to have been unnoticed in the existing literature, is that the tangent bundle of a statistical manifold is canonically an almost Hermitian manifold.

This observation, although mathematically very simple, is one of the most important of \[\text{Molb}\]. It tells us that statistics abounds with almost Hermitian manifolds. To illustrate this, let us consider what is probably the most simple example that one may think of. Take a finite set \( \Omega := \{x_1, \ldots, x_n\} \) and consider the space \( \mathcal{P}_n^\times \) of nowhere vanishing\[\text{N92}\] probabilities \( p : \Omega \to \mathbb{R}, \ p > 0, \ \sum_{k=1}^n p(x_k) = 1 \). This is a \((n-1)\)-dimensional statistical manifold, therefore its tangent bundle \( T\mathcal{P}_n^\times \) is an almost Hermitian manifold. Now the main observation in \[\text{Molb}\] may be formulated as follows: the canonical almost Hermitian structure of \( T\mathcal{P}_n^\times \) is locally isomorphic to the Kähler structure of \( \mathbb{P}(\mathbb{C}^n) \).

This result is intriguing. On one hand, it establishes a link between the geometrical formulation of quantum mechanics and information geometry, and suggests a possible information-theoretical origin of the quantum formalism. But on the other hand, the statistical relevance of Dombrowski’s construction is not at all clear, and since \( \mathcal{P}_n^\times \) is the only example in \[\text{Molb}\] for which explicit computations are performed, their are a priori no reasons for other statistical manifolds to yield interesting geometrical results of physical importance.

As such, the results in \[\text{Molb}\] are potentially fruitful, but they raise many questions that need further investigations.

In the present paper, we developed some of the ideas of \[\text{Molb}\] and present mathematical results –at

---

2See also \[\text{NS92}\].
3In \[\text{Dom62}\], Dombrowski is not explicitly using the language of dualistic structures, and his main concern is on the analytical properties of the almost complex structure that he constructs on the tangent bundle \( TM \) of a Riemannian manifold \((M, g)\) endowed with a connection \( \nabla \).
4The condition \( p > 0 \) (instead of \( p \geq 0 \)) is purely technical and ensures that \( \mathcal{P}_n^\times \) has no boundary nor corners.
the crossroad of information geometry, Kähler geometry, functional analysis and, to some extent, representation theory— which reinforce the idea that the quantum formalism has a statistical and information-theoretical origin. In particular, we want to describe a mechanism by which Kähler geometry emerges from information geometry, and to explain, by revisiting the geometrical formulation, how the latter, in its various aspects, can be understood within this larger information-theoretical setting.

A central role, in this development, is played by the so-called exponential families. Exponential families are a particular class of probability distributions which plays a key role in various ramifications of statistics, and especially in the context of statistical inference (see for example [AN00]). Their importance stems from the fact that among all possible parameterized statistical models, they are the only ones having efficient estimators, meaning roughly that it is possible, given sample data, to estimate the unknown parameters of the model in the “best possible way”. Examples of exponential families are found among the most common probability distributions: Bernoulli, beta, binomial, chi-square, Dirichlet, exponential, gamma, geometric, multinomial, normal, Poisson, to name but just a few.

For us, the important property of an exponential family $\mathcal{E}$ is that its canonical almost Hermitian structure (on $T\mathcal{E}$) is always a Kähler structure (Corollary 1.4), allowing us to define, via some refinements of Dombrowski’s arguments, what we shall call the Kählerification of an exponential family, denoted $\mathcal{E}^C$ (see 4, Definition 6.1). By construction, a Kählerification $\mathcal{E}^C$ is a Kähler manifold, and it always comes with a Riemannian submersion $\pi_E: \mathcal{E}^C \to \mathcal{E}$. As an important example, the Kählerification of $\mathcal{P}_n^\times$ yields $\mathbb{P}(\mathbb{C}^n)^\times$, i.e. $(\mathcal{P}_n^\times)^C \cong \mathbb{P}(\mathbb{C}^n)^\times$, where $\mathbb{P}(\mathbb{C}^n)^\times := \{[z_1, ..., z_n] \in \mathbb{P}(\mathbb{C}^n) | z_k \neq 0 \text{ for all } k = 1, ..., n\}$ (we use homogeneous coordinates, see 7 and Proposition 6.3). Other classical Kähler manifolds can be realized as the Kählerification of appropriate exponential families, like $\mathbb{C}^n$ and the Poincaré upper half plane $\mathbb{H}$.

Kählerifications (and their completions) generalize the usual quantum state space $\mathbb{P}(\mathbb{C}^n)$, but what about the observables? In the geometrical formulation, observables are Kähler functions, i.e. functions $f: \mathbb{P}(\mathbb{C}^n) \to \mathbb{R}$ whose associated Hamiltonian vector fields $X_f$ are Killing vector fields.

In 5, we investigate the properties of Kähler functions in the context of Kählerification and obtain a relation between the statistical structure of $\mathcal{E}$ and a class of Kähler functions on $\mathcal{E}^C$, as follows. If $(\Omega, dx)$ denotes the measured space on which $\mathcal{E}$ is defined and if $X: \Omega \to \mathbb{R}$ belongs to a certain class of random variables which depends on the exponential structure of $\mathcal{E}$ (see 6.4), then for any holomorphic isometry $\Phi: \mathcal{E}^C \to \mathcal{E}^C$, the function

$$\mathcal{E}^C \to \mathbb{R}, \quad z \mapsto \int_\Omega X(x) \left[ (\pi_\mathcal{E} \circ \Phi)(z) \right] (x) dx \quad (1)$$

is a Kähler function (Corollary 5.8).

This result is a “geometric analogue” of the usual spectral decomposition theorem for Hermitian matrices. For, when $\mathcal{E} = \mathcal{P}_n^\times$, then the space of Kähler functions $\mathcal{H}(\mathbb{P}(\mathbb{C}^n) \cap \mathcal{P}_n^\times)$ (the latter viewed as the natural “completion” of $(\mathcal{P}_n^\times)^C$), is isomorphic in the Lie algebra sense to the space of $n \times n$ skew Hermitian matrices $\mathfrak{u}(n)$, i.e., $\mathcal{H}(\mathbb{P}(\mathbb{C}^n) \cap \mathcal{P}_n^\times) \cong \mathfrak{u}(n)$, and the decomposition in 4 is in this case a rephrasing of the diagonalisability of a Hermitian matrix (see Lemma 7.6).

In 6, while revisiting the geometrical formulation of quantum mechanics, we use this correspondence with spectral theory to propose a definition for the spectrum of a Kähler function $f: \mathcal{E}^C \to \mathbb{R}$ of the form given in 1. Our definition reads as follows: $\text{spec}(f) := \text{Im}(X)$, where $\text{Im}(X)$ denotes the image of the random variable $X: \Omega \to \mathbb{R}$. As the decomposition in 1 is usually not unique, our definition is only consistent when invariance properties are met. We also define, for a Kähler function $f$ as in 1 and a point $z \in \mathcal{E}^C$, what might be interpreted, in a physical jargon, as the probability that the observable $f$ yields, upon measurement, the eigenvalue $\lambda$ while the system is in the state $z: P_{f,z}(\lambda) := \int_{X^{-1}(\lambda)} \left[ (\pi_\mathcal{E} \circ \Phi)(z) \right] (x) dx$.

When $\mathcal{E} = \mathcal{P}_n^\times$, then $\text{spec}(f)$ and $P_{f,z}$ are well defined for all Kähler functions $f$ on $\mathbb{P}(\mathbb{C}^n)$ (we extend our definitions via density arguments) and together, they yield the usual probabilistic interpretation of the geometrical formulation. In particular, $P_{f,z}$ depends on the expression $\cos^2 \left( d(.,.) \right)$, where $d(.,.)$ is the geodesic distance on $\mathbb{P}(\mathbb{C}^n)$.
When $\mathcal{E} = B(n, q)$ is the space of binomial distributions with parameter $q \in [0, 1]$ defined over $\Omega := \{0, \ldots, n\}$, then we have the following results (see [3]). The Kählerification of $B(n, q)$ is, up to completion, the 2-dimensional sphere of radius $n$ (regarded as a submanifold of $\mathbb{R}^3$). The space of Kähler functions on the sphere is generated by the functions $1, x, y, z$ and is isomorphic, in the Lie algebra sense, to $u(2)$. For the spectral theory, if $(u,v,w)\in \Omega = B_m$ magnetic fields produced by the Stern-Gerlach devices.

As we observe, this remarkable result is that it only depends on the statistical structure of the binomial distribution $B(n, q)$, providing support to the idea that the quantum formalism owes part of its mathematical structure to statistical concepts. Also, it shows that the probabilistic interpretation of the geometrical formulation, through $\text{spec}(f)$ and $P_{f,z}$ can be extended to more general situations than the one originally considered with the complex projective space, situations which are physically relevant.

We have to emphasize, however, that not all the possibilities of the Stern-Gerlach experiment are exhausted with (3), which are interpreted, at first, as a limitation of the statistical approach. But actually it is not. The remaining probabilities, as it turns out, can be obtained fairly easily by means of the “universal” inclusion $\mathcal{E} \ni B(n, q) \subseteq P_\mathcal{N}^\times$, as follows. By “Kählerifying” this inclusion, one gets an embedding $S^2 \hookrightarrow P(\mathbb{C}^{n+1})$ which makes it possible to extend every Kähler function $f$ on $S^2$ to a unique Kähler function $\hat{f}$ on $P(\mathbb{C}^{n+1})$, the latter function having the advantage to carry more informations than the original one. In fact, we show that the map $f \mapsto \hat{f}$ is a homomorphism of Lie algebras which is, via the appropriate identifications, an irreducible unitary representation of $u(2)$ (see Proposition 9.7 and lemmas 0.8 and 0.9). This allows us to extract the remaining probabilities (recall that the spin is usually described by the unitary representations of $\mathfrak{su}(2)$). Mathematically, this brings an interesting link between a purely geometrical problem – extending Kähler functions– and representation theory.

Collecting our results, we conclude that the spin of a particle can be entirely understood by means of the binomial distribution $B(n, q)$.

In [10] we briefly consider the space $\mathcal{N}(\mu, 1)$ of Gaussian distributions of mean $\mu$ and fixed deviation $\sigma = 1$ over $\Omega = \mathbb{R}$, give its Kählerification and describe its associated “spectral theory”. As we observe, this exponential family is closely related to the quantum harmonic oscillator, a fact which can only be fully understood by the introduction of an infinite dimensional analogue of $P_\mathcal{N}^\times$. On this, however, we say very little due to space limitation and refer the reader to [Mola].

To summarize, we carried out, following ideas of [Molb], an analysis of the mathematical foundations of quantum mechanics, using a geometric and information-theoretical approach, which points towards the

---

5By “maximum spin” we mean that the eigenvalue of the usual spin operator of the particle along the direction of the magnetic field of the first Stern-Gerlach device, is, among the possible values $-j, -j+1, \ldots, j-1, j$, exactly $j$. Here $j \in \{0, 1/2, 1, 3/2, \ldots\}$ is the spin of the particle.

6By “universal”, we simply mean that for any statistical manifold $S$ defined over a finite set $\{x_1, \ldots, x_n\}$, there is a canonical inclusion $S \subseteq P_n^\times$. The space $P_n^\times$ thus appears as a “universal container”.

---
following conclusion: the quantum formalism is grounded on the Kähler geometry which naturally emerges from statistics. Examples like the spin support this claim, and the various mechanisms involved have been described; we defined the Kählerification of an exponential family and sketched the very basis of what may be considered as a “statistical spectral” theory for Kähler functions. In doing so, we observed an intriguing link between the problem of extending Kähler functions and representation theory which seems to connect our approach to the standard way physicists work.

The author is fully aware that the techniques and definitions introduced in this paper are still in an infant stage, and that they should probably be modified in the light of further progress. Nevertheless, it is likely that the relationship between statistics and Kähler geometry will grow in importance, and we hope that it may help to get a better understanding of the mathematical foundations of quantum mechanics. Deepening this comprehension might well lead to a viable generalization of quantum mechanics, or at least to a new comprehension of some of its conceptually puzzling aspects, especially those related to the measurement problem.

2 Information geometry

In this section, we review the basic concepts of information geometry needed throughout this paper. Our (very short) presentation follows the currently reference book [AN00] whose emphasis is on the Fisher metric and α connections of a given statistical model (see also [MR93]).

A statistical manifold (or statistical model), is a couple (S,j) where S is a manifold and where j is an injective map from S to the space of all probability density functions p defined on a fixed measured space (Ω, dx): \[ j : S \mapsto \left\{ p : \Omega \rightarrow \mathbb{R} \mid p \text{ is measurable, } p \geq 0 \text{ and } \int_{\Omega} p(x) \, dx = 1 \right\}. \] (3)

In the case of a discrete space S, it will be implicitly assumed that dx is the counting measure, i.e. \( dx(A) = \text{card}(A) \), where card(A) denotes the cardinality of a given subset \( A \subset \Omega \). In this situation, integration of a function \( X : \Omega \rightarrow \mathbb{R} \) with respect to the probability \( p \, dx \) (p being a probability density function), is simply given by:

\[ \int_{\Omega} X(x) p(x) \, dx = \sum_{x \in \Omega} X(x) p(x). \] (4)

As a matter of notation, if \( (\xi : U \subseteq S \rightarrow \mathbb{R}^n) \) is a chart of a statistical manifold S with local coordinates \( \xi = (\xi_1, ..., \xi_n) \), then we shall indistinctly write \( p(x; \xi) \) or \( p_\xi(x) \) for the probability density function determined by \( \xi \) and in the variable \( x \in \Omega \).

Now, given a “reasonable” statistical manifold S, it is possible to define a metric \( h_\xi \) and a family of connections \( \nabla^{(a)} \) on S (a ∈ R) in the following way: for a chart \( \xi = (\xi_1, ..., \xi_n) \) of S, define

- \( (h_\xi)_{\xi}(\partial_i, \partial_j) := E_{p_\xi}(\partial_i \ln (p_\xi) \cdot \partial_j \ln (p_\xi)) \),
- \( \Gamma_{ij,k}^{(a)}(\xi) := E_{p_\xi} \left[ \left( \partial_i \partial_j \ln (p_\xi) + \frac{1-a}{2} \partial_i \ln (p_\xi) \cdot \partial_j \ln (p_\xi) \right) \partial_k \ln (p_\xi) \right] \),

where \( E_{p_\xi} \) denotes the mean, or expectation, with respect to the probability \( p_\xi \, dx \), and where \( \partial_i \) is a shorthand for \( \partial / \partial \xi_i \).

7 Depending on the symbole we use for the variable living in \( \Omega \), for example “x”, “k”, etc., we shall use the notation “dx”, “dk”, etc., for the measure on \( \Omega \).
It can be shown that if the above expressions are defined and smooth for every chart of $S$ (this is not always the case), then $h_F$ is a well defined metric on $S$ called the Fisher metric, and that the $\Gamma^{(\alpha)}_{ij,k}$'s are the Christoffel symbols of a connection $\nabla^{(\alpha)}$ called the $\alpha$-connection. Among the $\alpha$-connections, the $(\pm 1)$-connections are particularly important; the 1-connection is usually referred to as the exponential connection, also denoted $\nabla^{(\alpha)}$, while the $(-1)$-connection is referred to as the mixture connection, denoted $\nabla^{(m)}$.

In this paper, we will only consider statistical manifolds $S$ for which the Fisher metric and $\alpha$-connections are well defined.

One particularity of the $(\pm \alpha)$-connections is that they are dual of each other with respect to the Fisher metric $h_F$, or equivalently, that they form a dualistic structure on $S$. The general definition of a dualistic structure on an arbitrary manifold $M$ is as follows: a dualistic structure on $M$ is a triple $(h, \nabla, \nabla^*)$ where $h$ is a Riemannian metric on $M$ and where $\nabla$ and $\nabla^*$ are connections satisfying

$$X(h(Y, Z)) = h(\nabla_X Y, Z) + h(Y, \nabla^*_X Z),$$

for all vector fields $X, Y, Z$ on $M$. The connection $\nabla^*$ is called the dual connection, or conjugate connection, of the connection $\nabla$ (and vice versa).

An example of dualistic structure is, as we already said, given by the triple $(h_F, \nabla^{(\alpha)}, \nabla^{(-\alpha)})$ that one can always consider for a fixed $\alpha \in \mathbb{R}$ on a statistical manifold $S$ (provided of course that the Fisher metric and $(\pm \alpha)$-connections exist).

An important class of dualistic structures is that of dually flat structures. A dually flat structure on a manifold $M$ is a dualistic structure $(h, \nabla, \nabla^*)$ for which both connections are flat, meaning that their torsions and curvature tensors vanish. As conventions are not uniform in the literature, let us agree that $\nabla^*$ is well defined. In particular, since $\alpha$-connections are always torsion-free, $(h_F, \nabla^{(\alpha)}, \nabla^{(-\alpha)})$ is dually flat if and only if $R^{(\alpha)}$ or $R^{(-\alpha)}$ vanishes identically (here $R^{(\alpha)}$ denotes the curvature tensor of the $\alpha$-connection).

3 Exponential families

**Definition 3.1.** An exponential family $\mathcal{E}$ on a measured space $(\Omega, dx)$ is a set of probability density functions $p(x; \theta)$ of the form

$$p(x; \theta) = \exp \left\{ C(x) + \sum_{i=1}^{n} \theta_i F_i(x) - \psi(\theta) \right\},$$

where $C(x)$ is the log-partition function, $\theta = (\theta_1, \ldots, \theta_n)$ is a set of parameters, $F_i(x)$ is the $i$-th sufficient statistic, and $\psi(\theta)$ is the natural parameter of the exponential family.

---

\footnote{Given a connection $\nabla$ on a Riemannian manifold $(M, h)$, there exists a unique connection $\nabla^*$ on $M$ such that $\nabla^* x = 0$ holds; it is thus justified to call $\nabla^*$ the dual connection of $\nabla$.}
where \( C, F_1, \ldots, F_n \) are measurable functions on \( \Omega \), \( \theta = (\theta_1, \ldots, \theta_n) \) is a vector varying in an open subset \( \Theta \) of \( \mathbb{R}^n \) and where \( \psi \) is a function defined on \( \Theta \).

In the above definition, it is understood that if \( \Omega \) is discrete, then \( dx \) should be the counting measure. It is also assumed that the family \( \{1, F_1, \ldots, F_n\} \) is linearly independent, so that the map \( p(x, \theta) \mapsto \theta \in \Theta \) becomes a bijection, hence defining a global chart of \( \mathcal{E} \). The parameters \( \theta_1, \ldots, \theta_n \) are called the natural or canonical parameters of the exponential family \( \mathcal{E} \).

Besides the natural parameters \( \theta_1, \ldots, \theta_n \), an exponential family \( \mathcal{E} \) possesses another particularly important parametrization which is given by the expectation or dual parameters \( \eta_1, \ldots, \eta_n \):

\[
\eta_i(p_\theta) := E_{p_\theta}(F_i) = \int_\Omega F_i(x) p_\theta(x) \, dx.
\]

It is not difficult, assuming \( \psi \) to be smooth, to show that \( \eta_i(p_\theta) = \partial_\theta_i \psi \). The map \( \eta = (\eta_1, \ldots, \eta_n) \) is thus a global chart of \( \mathcal{E} \) provided that \( (\partial_{\theta_1} \psi, \ldots, \partial_{\theta_n} \psi) : \Theta \to \mathbb{R}^n \) is a diffeomorphism onto its image, condition that we will always assume.

The natural and expectation parameters are important in that they form affine coordinate systems with respect to \( \nabla^{(e)} \) and \( \nabla^{(m)} \):

**Proposition 3.2.** [AN00] Let \( \mathcal{E} \) be an exponential family such as in [8]. Then \( (\mathcal{E}, h_F, \nabla^{(e)}, \nabla^{(m)}) \) is dually flat and \( \theta = (\theta_1, \ldots, \theta_n) \) is an affine coordinate system with respect to \( \nabla^{(e)} \) while \( \eta = (\eta_1, \ldots, \eta_n) \) is an affine coordinate system with respect to \( \nabla^{(m)} \). Moreover, the following relation holds:

\[
h_F(\partial_{\theta_i}, \partial_{\theta_j}) = \delta_{ij},
\]

where \( \delta_{ij} \) denotes the Kronecker symbol.

Let us now give some examples of exponential families, mostly taken from [AN00].

**Example 3.3** (Normal Distribution). Normal distributions

\[
p(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\} \quad (x \in \mathbb{R}),
\]

form a 2-dimensional statistical manifold parameterized by \( (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}^*_+ \) (here \( \mathbb{R}^*_+ := \{x \in \mathbb{R} \mid x > 0\} \)), subsequently denoted \( \mathcal{N}(\mu, \sigma^2) \). This family is easily seen to be an exponential one, for one may write

\[
p(x; \mu, \sigma) = \exp\left\{ \frac{\mu}{\sigma^2} x - \frac{1}{2\sigma^2} x^2 - \frac{\mu^2}{2\sigma^2} - \ln(\sqrt{2\pi\sigma}) \right\},
\]

and define

\[
\theta_1 = \frac{\mu}{\sigma^2}, \quad \theta_2 = -\frac{1}{2\sigma^2}, \quad C(x) = 0, \quad F_1(x) = x, \quad F_2(x) = x^2, \quad \psi(\theta) = -\left(\frac{\theta_1}{\theta_2}\right)^2 + \frac{1}{2} \ln\left(-\frac{\pi}{\theta_2}\right).
\]

**Example 3.4** (finite \( \Omega \)). For a finite set \( \Omega = \{x_1, \ldots, x_n\} \), define

\[
P^\times_n := \left\{ p : \Omega \to \mathbb{R} \mid p(x) > 0 \text{ for all } x \in \Omega \text{ and } \sum_{k=1}^n p(x_k) = 1 \right\}.
\]

\footnote{Let us recall that an affine coordinate system on a manifold \( M \) with a flat connection \( \nabla \), or simply a \( \nabla \)-affine chart, is a coordinate system in which all the Christoffel symbols associated to \( \nabla \) vanish.}
The space $\mathbb{P}_n^\chi$ is clearly a statistical manifold of dimension $n-1$, and it can be turned into an exponential family by means of the following parameterization:

$$p(x;\theta) = \exp\left\{ \sum_{i=1}^{n-1} \theta_i F_i(x) - \psi(\theta) \right\},$$  \hspace{1cm} (15)

where $x \in \Omega$, $\theta = (\theta_1, \ldots, \theta_{n-1}) \in \mathbb{R}^{n-1}$, $F_i(x) = \delta_{ij}$ and where $\psi(\theta) = -\ln(1 + \sum_{i=1}^{n-1} \exp(\theta_i))$.

### 4 Dombrowski’s construction

In this section, we explain, following Dombrowski’s paper [Dom62], how the tangent bundle of a given dually flat manifold can be turned into a Kähler manifold by a simple geometric construction. This implies in particular that the tangent bundle $\mathcal{T}E$ of an exponential family is naturally a Kähler manifold.

Most of the results of this section are due to Dombrowski, except for Lemma 4.2 and subsequent corollaries which are natural extensions of [Dom62].

Recall that if $M$ is a manifold endowed with an affine connection $\nabla$, then Dombrowski splitting Theorem holds (see [Dom62] and [Lan02]):

$$T(TM) \cong TM \oplus TM \oplus TM,$$

this splitting being viewed as an isomorphism of vector bundles over $M$, and the isomorphism, say $\Phi$, being

$$T_u TM \ni A_{u_x} \mapsto (u_x, \pi_u A_{u_x}, K A_{u_x}),$$  \hspace{1cm} (17)

where $\pi : TM \to M$ is the canonical projection and where $K : T(TM) \to TM$ is the canonical connector associated to the connection $\nabla$ (see [Lan02]).

Having $A_{u_x} = \Phi^{-1}((u_x, v_x, w_x)) \in T_u TM$, we shall write, for simplicity, $A_{u_x} = (u_x, v_x, w_x)$ instead of $\Phi^{-1}((u_x, v_x, w_x))$, i.e., we will drop $\Phi$. The second component $v_x$ is usually referred to as the horizontal component of $A_{u_x}$ (with respect to the connection $\nabla$) and $w_x$ the vertical component.

With the above notation, and provided that $M$ is endowed with a Riemannian metric $h$, it is a simple matter to define on $TM$ an almost Hermitian structure. Indeed, we define a metric $g$, a 2-form $\omega$ and an almost complex structure $J$ by setting

$$g_{u_x}((u_x, v_x, w_x), (u_x, v_x, w_x)) := h_x(v_x, \overline{v}_x) + h_x(w_x, \overline{w}_x),$$
$$\omega_{u_x}((u_x, v_x, w_x), (u_x, \overline{v}_x, \overline{w}_x)) := h_x(v_x, \overline{w}_x) - h_x(w_x, \overline{v}_x),$$
$$J_{u_x}((u_x, v_x, w_x)) := (u_x, -w_x, v_x),$$  \hspace{1cm} (18)

where $u_x, v_x, w_x, \overline{v}_x, \overline{w}_x \in T_x M$.

Clearly, $J^2 = -\text{Id}$ and $g(J., J.) = g(., .)$, which means that $(TM, g, J)$ is an almost Hermitian manifold, and one readily sees that $g$ and $\omega$ are compatible, i.e., that $\omega = g(J., .)$; the 2-form $\omega$ is thus the fundamental 2-form of the almost Hermitian manifold $(TM, g, J)$. This is Dombrowski’s construction.

Observe that the map $\pi : (TM, g) \to (M, h)$ is a Riemannian submersion.

In [Dom62], Dombrowski shows the following:

**Proposition 4.1** ([Dom62]). Let $\nabla$ be an affine connection defined on a manifold $M$, and let $J$ be the almost complex structure associated to $\nabla$ as in (18). Then,

$$J \text{ is integrable } \iff \nabla \text{ is flat}.$$  \hspace{1cm} (19)
The tangent bundle $TM$ of a manifold $M$ endowed with a flat connection $\nabla$ is thus naturally a complex manifold. If in addition $M$ is equipped with a Riemannian metric $h$, then $TM$ becomes a complex Hermitian manifold for the Hermitian structure $(g, J, \omega)$ considered above.

For the 2-form $\omega$ defined in (18), we have the following result:

**Lemma 4.2.** Let $(M, h)$ be a Riemannian manifold endowed with a flat connection $\nabla$, and let $\omega$ be the 2-form defined as in (18). Then,

$$d\omega = 0 \iff T^* = 0,$$

where $T^*$ denotes the torsion of the dual connection $\nabla^*$. 

**Proof.** Let us consider a $\nabla$-affine chart $(\varphi : U \subseteq M \rightarrow \mathbb{R}^n)$ with local coordinates $\varphi = (x_1, \ldots, x_n)$. In this chart, all Christoffel symbols $\Gamma^k_{ij}$ associated to $\nabla$ vanish. Let us also consider the chart $(\overline{\varphi} : \overline{U} \subseteq TM \rightarrow \mathbb{R}^n \times \mathbb{R}^n)$ with local coordinates $\overline{\varphi} = (x_1, \ldots, x_n, y_1, \ldots, y_n)$ canonically associated to $(U, \varphi)$, i.e., $\overline{U} := \pi^{-1}(U)$ ($\pi : TM \rightarrow M$ being the canonical projection) and where $\overline{\varphi}((\sum^n_{i=1} a_i \partial_{x_i}, x)) := (\varphi(x), a_1, \ldots, a_n)$.

In this chart, it is not hard to see that

$$\Omega(x, y) = \begin{pmatrix} 0 & (h(x))_{ij} \\ -(h(x))_{ij} & 0 \end{pmatrix},$$

where $(h(x))_{ij} := h(\partial_{x_i}, \partial_{x_j})$ and where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$.

Now, $\Omega$ is closed if and only if for all $i, j, k = 1, \ldots, n$,

$$(d\Omega)(\partial_{x_i}, \partial_{x_j}, \partial_{x_k}) = 0,$$

and it is easy, using (21), to see that the only possibly non-vanishing terms in the equations (22) and (23) are $(d\Omega)(\partial_{x_i}, \partial_{y_j}, \partial_{y_k}) = \partial_{x_i} h_{jk} - \partial_{x_j} h_{ik}$. Moreover, it is a simple calculation to show that

$$\partial_{x_i} h_{jk} - \partial_{x_j} h_{ik} = h(T^*(\partial_{x_i}, \partial_{x_j}), \partial_{x_k}).$$

Hence, $d\Omega = 0$ if and only if $T^* = 0$. The lemma follows. □

Recall that an almost Hermitian structure $(g, J, \omega)$ on a given manifold is Kähler when the following two analytical conditions are met: (1) $J$ is integrable; (2) $d\omega = 0$. Having this in mind, Proposition 3.2 and Lemma 4.2 readily imply the following two corollaries:

**Corollary 4.3.** Let $(h, \nabla, \nabla^*)$ be a dualistic structure on a manifold $M$ and let $(g, J, \omega)$ be the almost Hermitian structure on $TM$ associated to $(h, \nabla)$ via Dombrowski’s construction. Then,

$$(TM, g, J, \omega) \text{ is Kähler} \iff (M, h, \nabla, \nabla^*) \text{ is dually flat}.$$

**Corollary 4.4.** The tangent bundle $TE$ of an exponential family $E$ is a Kähler manifold for the Kähler structure $(g, J, \omega)$ associated to $(h_F, \nabla^c)$ via Dombrowski’s construction. 

In the sequel, by the Kähler structure of $TE$, we shall implicitly refer to the Kähler structure of $TE$ described in Corollary 4.3.
5 Kähler functions on an exponential family

Definition 5.1. Let \((N,g,J,\omega)\) be a Kähler manifold. We shall say that a function \(f : N \to \mathbb{R}\) is a Kähler function if

\[
\mathcal{L}_{X_f}g = 0, \tag{26}
\]

where \(X_f\) denotes the symplectic gradient of \(f\) with respect to the symplectic form \(\omega\), i.e., \(\omega(X_f, \cdot) = df(\cdot)\), and where \(\mathcal{L}_{X_f}\) denotes the Lie derivative in the direction \(X_f\).

Clearly, a Kähler function \(f : N \to \mathbb{R}\) preserves the Kähler structure of \(N\) in the sense that \(\mathcal{L}_{X_f}g = 0\) and \(\mathcal{L}_{X_f}\omega = 0\), hence the terminology. Following [CMP90], we shall also denote by \(\mathcal{H}(N)\) the space of all Kähler functions defined on a Kähler manifold \(N\). When \(N\) has a finite number of connected components, then the space of Kähler functions on \(N\) is a finite dimensional Lie algebra for the natural Poisson bracket \(\{f,g\} := \omega(X_f,X_g)\).

In the case of a Kähler structure associated to a dually flat manifold via Dombrowski’s construction, we shall use the following terminology:

Definition 5.2. Let \((h,\nabla,\nabla^*)\) be a dually flat structure on a given manifold \(M\). We shall say that a function \(f : M \to \mathbb{R}\) is a Kähler function if \(f \circ \pi : TM \to \mathbb{R}\) is a Kähler function with respect to the Kähler structure of \(TM\) associated to \((h,\nabla)\) via Dombrowski’s construction (here \(\pi : TM \to M\) is the canonical projection).

We shall denote by \(\mathcal{H}(M)\) the space of Kähler functions on a dually flat manifold \(M\). Clearly, \(\mathcal{H}(M) \subseteq \mathcal{H}(TM)\) via the map \(f \mapsto f \circ \pi\).

We now want to characterize the space of Kähler functions on a dually flat manifold. To this end, recall that a vector field \(X\) on a manifold \(M\) is said to be \(\nabla\)-parallel with respect to a given connection \(\nabla\) if \(\nabla_Y X = 0\) for all vector fields \(Y\) on \(M\).

Proposition 5.3. Let \((h,\nabla,\nabla^*)\) be a dually flat structure on a manifold \(M\), and let \((g,J,\omega)\) be the Kähler structure on \(TM\) associated to \((h,\nabla)\) via Dombrowski’s construction. For a given function \(f : M \to \mathbb{R}\), we have:

\[
\text{\(f\) is a Kähler function} \iff \text{\(\text{grad}^h(f)\) is \(\nabla\)-parallel}, \tag{27}
\]

where \(\text{grad}^h(f)\) is the Riemannian gradient of \(f\) with respect to \(h\), i.e. \(h(\text{grad}^h(f), \cdot) = df(\cdot)\).

In order to show Proposition 5.3 we need a lemma.

Lemma 5.4. Under the hypothesis of Proposition 5.3 and, using the identification given in [17], we have for \(u_x \in T_xM\),

\[
(X_{f \circ \pi})_u = (u_x,0,-\text{grad}^h(f))_x \quad \text{and} \quad \varphi_t^{X_{f \circ \pi}}(u_x) = u_x - t \text{grad}^h(f)_x, \tag{28}
\]

where \(X_{f \circ \pi}\) is the symplectic gradient of \(f \circ \pi : TM \to \mathbb{R}\) with respect to \(\omega\) and where \(\varphi_t^{X_{f \circ \pi}}\) denotes the flow of \(X_{f \circ \pi}\).

\[10\] The fact that \(\mathcal{H}(N)\) is finite dimensional comes from the following result: if \((M,h)\) is a connected Riemannian manifold, then its space of Killing vector fields \(\mathfrak{K}_{\text{Killing}}(M) := \{X \in \mathfrak{X}(M) \mid \mathcal{L}_X h = 0\}\) is finite dimensional (see for example [Bon02]).
Proof. For \( A_{u_*} \in T_{u_*} TM \), we have by definition of \( \omega \) (see (13)) and \( \text{grad}^h(f) \):

- \((f \circ \pi)_{u_*} A_{u_*} = \omega((X_{f \circ \pi})_{u_*}, A_{u_*}) = h(\pi_{u_*}(X_{f \circ \pi})_{u_*}, K A_{u_*}) - h(\pi_{u_*} A_{u_*}, K(X_{f \circ \pi})_{u_*}) \), \hspace{1cm} (29)

- \((f \circ \pi)_{u_*} A_{u_*} = f_{u_* \pi_{u_*}} A_{u_*} = h(\text{grad}^h(f)_x, \pi_{u_*} A_{u_*}) \). \hspace{1cm} (30)

Comparing (29) and (30), we get that (see Lemma 5.4), \( \pi_{u_*}(X_{f \circ \pi})_{u_*} = 0 \) and \( K(X_{f \circ \pi})_{u_*} = -\text{grad}^h(f)_x \), which, in view of the identification given in (17), implies the first equation in (28).

The second equation in (28) is an easy consequence of the first. The lemma follows.

\[
\begin{align*}
\text{Remark 5.5.} \quad \text{A simple consequence of Lemma 5.4 is that } \mathcal{K}(M), \text{ viewed as a Lie subalgebra of } \mathcal{K}(TM), \text{ is commutative.}
\end{align*}
\]

\textbf{Proof of Proposition 5.3.} In this proof, we use the identification given in (17) as well as the notation introduced in Lemma 5.4.

Let \( A_{u_*} = (u_x, v_x, w_x) \) and \( \mathcal{A}_{u_*} = (u_x, \tau_x, \omega_x) \) be two tangent vectors in \( T_{u_*} TM \). Since \((\pi \circ \varphi_{t}^X_{f \circ \pi})(u_x) = \pi(u_x - t \text{ grad}^h(f)_x) = \pi(u_x) \), \( \pi_{u_*}(\varphi_{t}^X_{f \circ \pi})_{u_*} A_{u_*} = \pi_{u_*}(u_x, v_x, w_x) = v_x \), and thus, recalling the definition of \( g \) given in (13),

\[
\begin{align*}
\left( (\varphi_{t}^X_{f \circ \pi})^* g \right)_{u_x} (A_{u_*}, \mathcal{A}_{u_*}) &= g ((\varphi_{t}^X_{f \circ \pi})_{u_*} A_{u_*}, (\varphi_{t}^X_{f \circ \pi})_{u_*} \mathcal{A}_{u_*}) \\
&= h(v_x, \tau_x) + h(K(\varphi_{t}^X_{f \circ \pi})_{u_*} A_{u_*}, K(\varphi_{t}^X_{f \circ \pi})_{u_*} \mathcal{A}_{u_*}).
\end{align*}
\]

We have to compute \( K(\varphi_{t}^X_{f \circ \pi})_{u_*} A_{u_*} \). For this, observe that

\[
\begin{align*}
K(\varphi_{t}^X_{f \circ \pi})_{u_*} A_{u_*} &= K(\varphi_{t}^X_{f \circ \pi})_{u_*} (u_x, v_x, w_x) \\
&= K(\varphi_{t}^X_{f \circ \pi})_{u_*} (u_x, v_x, 0) + K(\varphi_{t}^X_{f \circ \pi})_{u_*} (u_x, 0, w_x). 
\end{align*}
\]

(32)

The second term in (32) is easily computed:

\[
K(\varphi_{t}^X_{f \circ \pi})_{u_*} (u_x, 0, w_x) = K \frac{d}{ds} \bigg|_0 \varphi_{t}^X_{f \circ \pi} (u_x + sw_x) = K \frac{d}{ds} \bigg|_0 (u_x + sw_x - t \text{ grad}^h(f)_x) = w_x. \quad (33)
\]

For the first term in (32), we introduce a curve \( V(s) \) in \( TM \) such that \( dV(s)/ds|_0 = (u_x, v_x, 0) \) and such that \( V(s) \) is horizontal for all \( s \). Observe that \( d(\pi \circ V(s))/ds|_0 = \pi_{u_*}(u_x, v_x, w_x) = v_x \). Using this curve, we see that

\[
\begin{align*}
K(\varphi_{t}^X_{f \circ \pi})_{u_*} (u_x, v_x, 0) &= K \frac{d}{ds} \bigg|_0 \varphi_{t}^X_{f \circ \pi} (V(s)) = K \frac{d}{ds} \bigg|_0 (V(s) - t \text{ grad}^h(f)_{\pi(V(s))}) \\
&= \nabla_{u_*} (V(s) - t \text{ grad}^h(f)_{\pi(V(s))}) = \nabla_{u_*} V(s) - t \nabla_{u_*} \text{ grad}^h(f)_{\pi(V(s))} \\
&= -t \nabla_{v_x} \text{ grad}^h(f)_{\pi(V(s))}. 
\end{align*}
\]

(34)

Now, (31), (32), (33) and (34) yield

\[
\begin{align*}
\left( (\varphi_{t}^X_{f \circ \pi})^* g \right)_{u_x} (A_{u_*}, \mathcal{A}_{u_*}) &= g(A_{u_*}, \mathcal{A}_{u_*}) - t \left( h(\nabla_{v_x} \text{ grad}^h(f), \omega_x) + h(\nabla_{v_x} \text{ grad}^h(f), w_x) \right) \\
&+ t^2 h(\nabla_{v_x} \text{ grad}^h(f), \nabla_{v_x} \text{ grad}^h(f))
\end{align*}
\]

from which we clearly see that \( \varphi_{t}^X_{f \circ \pi} \) is an isometry for all \( t \) if and only if \( \text{grad}^h(f) \) is \( \nabla \)-parallel. The proposition follows. 

\[
\Box
\]
Let us now specialize to the case of an exponential family. So, let $\mathcal{E}$ be an exponential family defined on a measured space $(\Omega, dx)$ with elements of the form $p(x; \theta) = \exp\{C(x) + \sum_{i=1}^{n} \theta_i F_i(x) - \psi(\theta)\}$ as in (35), and let us consider the following space of functions:

$$\mathcal{A}_\mathcal{E} := \text{Vect}_{\mathbb{R}}\{1, F_1, ..., F_n\},$$

(36)

i.e., $\mathcal{A}_\mathcal{E}$ is the real vector space generated by the constant function 1 and the functions $F_1, ..., F_n : \Omega \to \mathbb{R}$.

**Proposition 5.6.** For a function $f : \mathcal{E} \to \mathbb{R}$, we have:

$$f \text{ is a Kähler function} \iff \left( \exists X \in \mathcal{A}_\mathcal{E} : f(p) = \int_\Omega X(x) p(x) \, dx \ \forall p \in \mathcal{E} \right).$$

(37)

In order to show Proposition 5.6, we need the following lemma.

**Lemma 5.7.** The expectation parameters $\eta_i : \mathcal{E} \to \mathbb{R}, p \mapsto \int_\Omega F_i(x)p(x)dx$ satisfy the following relation:

$$\text{grad}^{kr}(\eta_i) = \partial_{\eta_i}.$$  

(38)

In particular, expectation parameters are Kähler functions.

**Proof.** Using the duality between the natural and expectation parameters (see (111)), one easily sees that

$$\partial_{\eta_i} = \sum_{k=1}^{n} (h_F)^{ik}\partial_{\eta_k},$$

where $(h_F)^{ik}$'s are the coefficients of the inverse of the matrix $(h_F)_{ij} := h_F(\eta_i, \eta_j)$, and also that $h_F(\partial_{\eta_i}, \partial_{\eta_j}) = (h_F)^{ij}$. It follows that

$$\text{grad}^{kr}(\eta_i) = \sum_{a,b=1}^{n} (h_F)_{ab}\frac{\partial \eta_i}{\partial \eta_a} = \sum_{a,b=1}^{n} (h_F)_{ab}\delta_{ia}\partial_{\eta_a} = \sum_{b=1}^{n} (h_F)^{ib}\partial_{\eta_b}$$

$$= \sum_{b,k=1}^{n} (h_F)^{ib}(h_F)^{bk}\partial_{\theta_b} = \sum_{k=1}^{n} \delta_{ik}\partial_{\theta_k} = \partial_{\eta_i},$$

(39)

which is the desired relation. 

**Proof of Proposition 5.6.** Let $f : \mathcal{E} \to \mathbb{R}$ be a function. If $f$ is a Kähler function, then according to Proposition 5.3, $\text{grad}^{kr}(f)$ is $\nabla (\cdot)$-parallel, which means that when expressed in the $\nabla (\cdot)$-affine chart $\theta = (\theta_1, ..., \theta_n)$, $\text{grad}^{kr}(f)$ is a constant vector field. We can thus write $\text{grad}^{kr}(f) = \sum_{i=1}^{n} a_i \partial_{\theta_i}$, where $a_1, ..., a_n$ are some real constants. But then, according to Lemma 5.7 and the definition of $\eta_i$, 

$$\text{grad}^{kr}(f) = \sum_{i=1}^{n} a_i \partial_{\theta_i} = \sum_{i=1}^{n} a_i \text{grad}^{kr}(\eta_i) = \text{grad}^{kr}\left( \sum_{i=1}^{n} a_i \eta_i \right)$$

$$= \text{grad}^{kr}\left( \sum_{i=1}^{n} a_i \int_\Omega F_i(x)p(x)dx \right) = \text{grad}^{kr}\left( \int_\Omega \sum_{i=1}^{n} a_i F_i(x)p(x)dx \right).$$

(40)

Hence, and up to an additive constant, $f(p) = \int_\Omega \sum_{i=1}^{n} a_i F_i(x)p(x)dx$ which shows one direction of the proposition. The other direction being trivial, the proposition follows. 

A direct consequence of Proposition 5.6 is the following result:

**Corollary 5.8.** Functions of the form

$$T\mathcal{E} \to \mathbb{R}, \quad z \mapsto \int_\Omega X(x)[\{\pi \circ \Phi\}(z)](x)dx,$$

(41)

where $X \in \mathcal{A}_\mathcal{E}$ and where $\Phi : T\mathcal{E} \to T\mathcal{E}$ is a holomorphic isometry, are Kähler functions on $T\mathcal{E}$.
6 Kählerification of an exponential family

Let $\mathcal{E}$ be an exponential family and let $(g, J, \omega)$ be the Kähler structure of $T\mathcal{E}$. We define a subgroup $\Gamma(\mathcal{E})$ of the group of all diffeomorphisms $\text{Diff}(T\mathcal{E})$ of $T\mathcal{E}$ by letting

$$\Gamma(\mathcal{E}) := \{ \phi \in \text{Diff}(T\mathcal{E}) \mid \phi^* g = g, \ \phi^* J = J, \ \phi^* \omega = \omega, \ \text{and} \ f \circ \phi = f \ \text{for all} \ f \in \mathcal{K}(T\mathcal{E}) \}.$$  \hspace{1cm} (42)

**Definition 6.1.** Let $\mathcal{E}$ be an exponential family having a discrete $\Gamma(\mathcal{E})$ and whose natural action\footnote{The natural action of $\Gamma(\mathcal{E})$ on $T\mathcal{E}$ is simply given by $\gamma \cdot u_x := \gamma(u_x)$, where $\gamma \in \Gamma(\mathcal{E})$ and $u_x \in T\mathcal{E}$.} on $T\mathcal{E}$ is free and proper. The quotient space $T\mathcal{E}/\Gamma(\mathcal{E})$ is thus naturally a Kähler manifold for which the quotient map $T\mathcal{E} \to T\mathcal{E}/\Gamma(\mathcal{E})$ becomes a holomorphic Riemannian submersion. We shall call this quotient the Kählerification of $\mathcal{E}$, and use the following notation:

$$E^C := T\mathcal{E}/\Gamma(\mathcal{E}).$$ \hspace{1cm} (43)

At this point, it is worth mentioning that in all the examples considered in this paper, the group $\Gamma(\mathcal{E})$ is discrete and that its natural action on $T\mathcal{E}$ is free and proper. In the sequel, we will always assume that the exponential families under consideration fulfill the conditions of Definition 6.1.

Let us investigate the geometrical structure of $E^C$.

**Lemma 6.2.** For every $\gamma \in \Gamma(\mathcal{E})$, we have

$$\pi \circ \gamma = \pi,$$ \hspace{1cm} (44)

where $\pi : T\mathcal{E} \to \mathcal{E}$ is the canonical projection.

**Proof.** Since expectation parameters $\eta_i : \mathcal{E} \to \mathbb{R}$ are Kähler functions (see Lemma 5.7), and since $\eta = (\eta_1, ..., \eta_n)$ is a chart of $\mathcal{E}$, we have by definition of $\Gamma(\mathcal{E})$,

$$\eta_i \circ \pi \circ \gamma = \eta_i \circ \pi \ \text{for all} \ i \ \Rightarrow \ \eta(\pi(\gamma(x))) = \eta(\pi(x)) \ \text{for all} \ x \in T\mathcal{E}$$

$$\Rightarrow \ \pi \circ \gamma = \pi.$$ \hspace{1cm} (45)

This is the desired relation. \hfill \Box

**Remark:** Lemma 6.2 readily implies that the projection $\pi : T\mathcal{E} \to \mathcal{E}$ factorizes through $E^C$, yielding a submersion $E^C \to \mathcal{E}$ that we shall denote by $\pi_E$, or simply $\pi$. A Kählerification has thus the structure of a fiber bundle induced by the submersion

$$\pi_E : E^C \to \mathcal{E};$$ \hspace{1cm} (46)

whose fiber over $p \in \mathcal{E}$ is diffeomorphic to $T_p\mathcal{E}/\Gamma(\mathcal{E})$.

Clearly, $\pi_E : E^C \to \mathcal{E}$ is a Riemannian submersion. Also, $\mathcal{K}(T\mathcal{E}) \cong \mathcal{K}(E^C)$ (Lie algebra isomorphism), and consequently there are analogues of Proposition 5.6 and Corollary 5.8 for the space of Kähler functions on $E^C$. Music
7 Kählerification of $\mathcal{P}_n^\times$ and complex projective spaces

Recall from Example 3.4 that $\mathcal{P}_n^\times$ is the space of non-vanishing probability density functions $p$ defined on a finite set $\Omega = \{x_1, \ldots, x_n\}$, i.e.,

$$\mathcal{P}_n^\times := \left\{ p : \Omega \to \mathbb{R} \mid p(x_i) > 0 \text{ for all } x_i \in \Omega \text{ and } \sum_{i=1}^n p(x_i) = 1 \right\}.$$  \hfill (47)

This space is clearly a connected manifold of dimension $n - 1$, and as we already saw, $\mathcal{P}_n^\times$ is an exponential family.

In the context of information geometry, it is customary to describe the tangent bundle of $\mathcal{P}_n^\times$ using the exponential representation:

$$T_p\mathcal{P}_n^\times \cong \left\{ u = (u_1, \ldots, u_n) \in \mathbb{R}^n \mid u_1 p_1 + \cdots + u_n p_n = 0 \right\},$$  \hfill (48)

where $p \in \mathcal{P}_n^\times$, and where by definition, $p_i := p(x_i)$ for all $x_i \in \Omega$.

If $u \in \mathbb{R}^n$ is a vector satisfying $u_1 p_1 + \cdots + u_n p_n = 0$ for a given probability density function $p : \Omega \to \mathbb{R}$, then we shall denote by $[u]_p$ the unique tangent vector of $\mathcal{P}_n^\times$ at the point $p$ determined by the exponential representation. One easily sees that if $p(t)$ is a smooth curve in $\mathcal{P}_n^\times$, then

$$\frac{d}{dt} \bigg|_0 p(t) = [u]_p(0) \iff \frac{d}{dt} \bigg|_0 p_i(t) = p_i(0) u_i \text{ for all } i = 1, \ldots, n,$$

(49)

where $p_i(t) := \langle p(t) \rangle(x_i)$.

Equation (49) is actually one way to define the exponential representation.

In term of the exponential representation, the Fisher metric $h_F$ has the following expression:

$$(h_F)_p([u]_p, [v]_p) = \sum_{i=1}^n p_i u_i v_i,$$  \hfill (50)

while the covariant derivative $D^{(c)}[V]_{p(t)}/dt$ of a vector field $[V]_{p(t)}$ along a curve $p : I \subseteq \mathbb{R} \to \mathcal{P}_n^\times$ with respect to the exponential connection $\nabla^{(c)}$ is given by

$$\frac{D^{(c)}}{dt} [V(t)]_{p(t)} = [\dot{V}(t) - E_{p(t)}(\dot{V}(t))]_{p(t)},$$  \hfill (51)

where $E_{p(t)}(\dot{V}(t)) := p_1(t) \dot{V}_1(t) + \cdots + p_n(t) \dot{V}_n(t)$ is the mean of the vector $\dot{V}(t) = (\dot{V}_1(t), \ldots, \dot{V}_n(t))$ with respect to the probability $p(t)$ and where $E_{p(t)}(\dot{V}(t))$ in (51) as to be understood as the vector $(E_{p(t)}(\dot{V}(t))_1, \ldots, E_{p(t)}(\dot{V}(t))_n) \in \mathbb{R}^n$.

For later purposes, let us also give the following result which gives an explicit description of the inverse of the map $\Phi : T(T\mathcal{P}_n^\times) \to \mathcal{P}_n^\times \oplus \mathcal{P}_n^\times \oplus \mathcal{P}_n^\times$ introduced in (17). The proof may be found in [Molv].

**Lemma 7.1.** For $[u]_p, [v]_p, [w]_p \in T_p\mathcal{P}_n^\times$, we have:

$$\Phi^{-1}([u]_p, [v]_p, [w]_p) = \frac{d}{dt} \bigg|_0 [u + tw - E_{p(t)}(u + tw)]_{p(t)},$$  \hfill (52)

where $p(t)$ is a smooth curve in $\mathcal{P}_n^\times$ satisfying $p(0) = p$ and $dp(t)/dt|_0 = [v]_p$. 

14
We now want to relate the natural Kähler structure of $TP_n^\times$ to the Kähler structure of the complex projective space $\mathbb{P}(\mathbb{C}^n)$. To this end, and for the reader’s convenience, let us digress a little on $\mathbb{P}(\mathbb{C}^n)$.

Recall that the complex projective space $\mathbb{P}(\mathbb{C}^n)$ is the quotient $(\mathbb{C}^n - \{0\})/\sim$, where the equivalence relation “$\sim$” is defined by

$$(z_1, \ldots, z_n) \sim (w_1, \ldots, w_n) \iff \exists \lambda \in \mathbb{C} - \{0\}: (z_1, \ldots, z_n) = \lambda(w_1, \ldots, w_n).$$

For $z = (z_1, \ldots, z_n) \in \mathbb{C}^n - \{0\}$, we shall denote by $[z] = [z_1, \ldots, z_n]$ the corresponding element of $\mathbb{P}(\mathbb{C}^n)$.

One may identify Proposition 7.2 in [Molb], the following is shown:

The map $\mathbb{C}^n \times \{1\} \rightarrow \mathbb{P}(\mathbb{C}^n)$ where

$$z \mapsto [z],$$

is a universal covering map whose deck transformation group is isomorphic to $\mathbb{Z}^{n-1}$.

In [Molb], the following is shown:

**Proposition 7.2 (Molb).**

(i) The map $\tau : TP_n^\times \rightarrow \mathbb{P}(\mathbb{C}^n)^\times$, $[u]_p \mapsto \left[\sqrt{\lambda_1} e^{iu_1/2}, \ldots, \sqrt{\lambda_n} e^{iu_n/2}\right]$, where $\mathbb{P}(\mathbb{C}^n)^\times$ is the open subset of $\mathbb{P}(\mathbb{C}^n)$ defined by

$$\mathbb{P}(\mathbb{C}^n)^\times := \{[z_1, \ldots, z_n] \in \mathbb{P}(\mathbb{C}^n) \mid z_i \neq 0 \text{ for all } i = 1, \ldots, n\}.$$
Remark 7.3. In \[\text{Molb}\], we were defining the Fisher metric \( h_F \) as being the one considered in this paper, but multiplied by a factor \( 1/4 \). Because of that, the first two formulas in (59) differ from the corresponding formulas in \[\text{Molb}\] by a factor \( 1/4 \).

Remark 7.4. Observe that every deck transformation of \( TP_n^\times \) has to be a holomorphic isometry.

We can now state the main result of this section.

Proposition 7.5. The group \( \Gamma(P_n^\times) \) coincides with the deck transformation group of the universal covering map \( \tau : TP_n^\times \to P(C^n)^\times \). In particular, if we multiply both the Fubini-Study metric \( g_{FS} \) and the Fubini-Study symplectic form \( \omega_{FS} \) by a factor 4, then we get a natural identification of Kähler manifolds:

\[
(P_n^\times)^C \cong P(C^n)^\times.
\] (60)

Moreover, in term of the above identification, the canonical projection \( \pi_P^\times : (P_n^\times)^C \to P_n^\times \) becomes

\[
\pi_P^\times : \mathbb{P}(\mathbb{C}^n)^\times \to P_n^\times, \quad \pi_P^\times([z])(x_k) := \frac{\bar{z}_k}{\langle z, z \rangle}.
\] (61)

We will show Proposition 7.5 with a series of lemmas.

Lemma 7.6. Let \( \mathcal{K}(P(C^n)) \) be the space of Kähler functions on \( P(C^n) \) and let \( u(n) \) be the space of complex \( n \times n \) skew Hermitian matrices. If \( \mathcal{K}(P(C^n)) \) is endowed with its natural Poisson bracket \( \{ f, g \} := \omega_{FS}(X_f, X_g) \), then the map \( u(n) \to \mathcal{K}(P(C^n)), A \mapsto \xi^A \), where

\[
\xi^A([z]) := \frac{i}{2} \frac{\langle z, A \cdot z \rangle}{\langle z, z \rangle}, \quad (z \in \mathbb{C}^n - \{0\})
\] (62)

is a Lie algebra isomorphism.

Proof. See for example [CMP90]. \qed

Lemma 7.7. Let \([z], [w] \) be two points in \( P(C^n) \). We have:

\[
[z] = [w] \iff f([z]) = f([w]) \text{ for all } f \in \mathcal{K}(P(C^n)).
\] (63)

Proof. Let \( z, w \in \mathbb{C}^n - \{0\} \) be two vectors. Using Proposition 7.6 and especially (62), it is easy to see that if \( f([z]) = f([w]) \) for all Kähler functions \( f \) on \( P(C^n) \), then

\[
2 \Re \left( \sum_{a<b} A_{ab} (\bar{z}_a z_b - \bar{w}_a w_b) \right) + \sum_a A_{aa} (|z_a|^2 - |w_a|^2) = 0,
\] (64)

where \( A_{ab} (a, b = 1, \ldots, n, a < b) \) are arbitrary complex numbers, and where \( A_{aa} (a = 1, \ldots, n) \) are arbitrary real numbers (one may think of \( A_{ab} \) as the coefficients of a Hermitian matrix). From (64), we deduce that

\[
\bar{z}_a z_b = \bar{w}_a w_b
\] (65)

for all \( a, b = 1, \ldots, n \). By introducing polar decompositions and with some algebraic manipulations, it is then easy to see that \( z \) and \( w \) are collinear. The lemma follows. \qed

Lemma 7.8. A function \( f : TP_n^\times \to \mathbb{R} \) is a Kähler function if and only if there exists a Kähler function \( \overline{f} : P(C^n) \to \mathbb{R} \) such that

\[
f = \overline{f} \circ \tau.
\] (66)
Proof. Let $f : TP_n^\times \to \mathbb{R}$ be a Kähler function. Since $\tau : TP_n^\times \to \mathbb{P}(\mathbb{C}^n)^\times$ is a covering map, for every $z \in TP_n^\times$, there exists an open and connected set $U_z \subseteq TP_n^\times$ containing $z$ and such that the restriction of $\tau$ to $U_z$ becomes a diffeomorphism between $U_z$ and $\tau(U_z)$. Let us denote this restriction by $\tau|_{U_z}$. According to (ii) in Proposition 7.2, $\tau|_{U_z}$ is a holomorphic isometry; this implies that the map $f|_{(\tau|_{U_z})^{-1}} : \tau(U_z) \to \mathbb{R}$ is a Kähler function, which means in particular that $X_{f \circ (\tau|_{U_z})^{-1}}$ is a Killing vector field on $\tau(U_z)$. But now, since $\mathbb{P}(\mathbb{C}^n)$ is a connected, simply connected and complete (in the Riemannian sense) Kähler manifold, there exists an open and connected set $U \subseteq \mathbb{P}(\mathbb{C}^n)$ such that the restriction of $\tau$ to $U \cap \tau(U_z)$ coincides. But now, since Kähler functions on $\mathbb{P}(\mathbb{C}^n)$ extend uniquely and continuously, the Kähler function $\tau \circ f$ coincides. This implies $f|_{U_z} = f|_{U \cap \tau(U_z)}$, which means that the map $f|_{U_z}$ is a Killing vector field on $\tau(U_z)$. But now, since $\tau$ is a holomorphic isometry, it follows from (62) that if two Kähler functions on the complex projective space coincide on $\tau(U_z)$, then they are equal. This implies $f|_{U_z} = f|_{U \cap \tau(U_z)}$, and therefore $f|_{U_z} = f|_{U \cap \tau(U_z)}$. Let us now show that it is also globally true. So let $U, V \subseteq TP_n^\times$ be two open connected sets whose intersection is not empty, and such that there exist two Kähler functions $f_U, f_V : \mathbb{P}(\mathbb{C}^n) \to \mathbb{R}$ verifying $f_U = f_V$ on $U \cap V$. Since $f_U \circ \tau$ and $f_V \circ \tau$ coincide on the intersection of $U$ and $V$, there exists a connected open subset of $\mathbb{P}(\mathbb{C}^n)$ on which $f_U$ and $f_V$ coincide. But now, since Kähler functions on $\mathbb{P}(\mathbb{C}^n)$ are of the form $f^A$ (see Lemma 7.6), it is clear from (62) that if two Kähler functions on the complex projective space coincide on an open subset, then they are equal. This implies $f_U = f_V$ from which the lemma follows.

The above formula shows that the statement in the lemma is locally true. Let us now show that it is also globally true. So let $U, V \subseteq TP_n^\times$ be two open connected sets whose intersection is not empty, and such that there exist two Kähler functions $f_U, f_V : \mathbb{P}(\mathbb{C}^n) \to \mathbb{R}$ verifying $f_U = f_V$ on $U \cap V$. Since $f_U \circ \tau$ and $f_V \circ \tau$ coincide on the intersection of $U$ and $V$, there exists a connected open subset of $\mathbb{P}(\mathbb{C}^n)$ on which $f_U$ and $f_V$ coincide. But now, since Kähler functions on $\mathbb{P}(\mathbb{C}^n)$ are of the form $f^A$ (see Lemma 7.6), it is clear from (62) that if two Kähler functions on the complex projective space coincide on an open subset, then they are equal. This implies $f_U = f_V$ from which the lemma follows.

Since every deck transformation of $TP_n^\times$ is a holomorphic isometry, it follows from Lemma 7.9 that the deck transformation group of the universal covering map $\tau : TP_n^\times \to \mathbb{P}(\mathbb{C}^n)^\times$ is a subgroup of $\Gamma(\mathbb{P}_n^\times)$. The converse is given by the following lemma.

Lemma 7.9. An element $\gamma \in \Gamma(\mathbb{P}_n^\times)$ is necessarily a deck transformation.

Proof. Let $\gamma$ be an element of $\Gamma(\mathbb{P}_n^\times)$. By definition of $\Gamma(\mathbb{P}_n^\times)$, and taking into account Lemma 7.7 and Lemma 7.8, we see that

\[
\left( f \circ \tau \circ \gamma = f \circ \tau \quad \forall \ f \in \mathcal{K}(\mathbb{P}(\mathbb{C}^n)) \right) \Rightarrow \tau \circ \gamma = \tau,
\]

i.e., $\gamma$ is a deck transformation. The lemma follows.

Equation (61) being straightforward, Proposition 7.5 is now a direct consequence of Lemma 7.8 and Lemma 7.9.

Let us end this section with a few important remarks on the map $\pi_{\mathbb{P}_n^\times} : \mathbb{P}(\mathbb{C}^n)^\times \to \mathbb{P}_n^\times$. Clearly, $\pi_{\mathbb{P}_n^\times}$ extends uniquely as a continuous map $\pi_{\mathbb{P}_n} : \mathbb{P}(\mathbb{C}^n) \to \mathbb{P}_n$, where

\[
\mathbb{P}_n := \left\{ p : \Omega \to \mathbb{R} \mid p(x_i) \geq 0 \ \text{for all} \ x_i \in \Omega \ \text{and} \ \sum_{i=1}^n p(x_i) = 1 \right\}.
\]

Notice that $\mathbb{P}_n^\times \subseteq \mathbb{P}_n$, and that these two spaces are distinguished only by the conditions $p > 0$ for $\mathbb{P}_n^\times$ and $p \geq 0$ for $\mathbb{P}_n$. Notice also that $\mathbb{P}_n$ is not a smooth manifold, for it has a boundary and corners.

\[12\text{It is well known that every Killing vector field on } \mathbb{P}(\mathbb{C}^n) \text{ can be realized as the Hamiltonian vector field of an appropriate}\]

\[13\text{We endow } \mathbb{P}_n \text{ with the topology induced by } \mathbb{R}^n \text{ via the injection } \mathbb{P}_n \hookrightarrow \mathbb{R}^n, \ p \mapsto (p(x_1), ..., p(x_n)).\]
The map $\pi_{P_n} : P(C^n) \to P_n$ has the property that it makes the following diagram commutative

$$P(C^n)^\times \xrightarrow{\pi_{P_n}^\times} P_n^\times$$
$$\downarrow \quad \downarrow$$
$$P(C^n) \xrightarrow{\pi_{P_n}} P_n$$

($i, j$ are inclusions), which allows to carry over many structural properties of the space of Kähler functions on $P(C^n)^\times$ to the space of Kähler functions on $P(C^n)$. Indeed, let $\mathcal{K}(P_n)$ denotes the following space of functions

$$\left\{ f : P_n \to \mathbb{R} \mid f(p) = \sum_{k=1}^{n} X_k p(x_k), X = (X_1, ..., X_n) \in \mathbb{R}^n \right\}. \quad (71)$$

By using Corollary 5.8, Lemma 7.8, some obvious continuity arguments and the fact that $A_{P_n^\times} = \{X : \Omega \to \mathbb{R} \cong \mathbb{R}^n \}$ (see the general definition of $A_E$ given in [30] and Example 3.4), one easily shows the following:

**Proposition 7.10.** We have:

(i) $\mathcal{K}(P(C^n)^\times) \cong \mathcal{K}(P(C^n))$, (Lie algebra isomorphism),

(ii) $\mathcal{K}(P_n^\times) \cong \mathcal{K}(P_n)$,

(iii) functions on $P(C^n)$ of the form $f \circ \pi_{P_n} \circ \Phi$, where $\Phi : P(C^n) \to P(C^n)$ is a holomorphic isometry and where $f \in \mathcal{K}(P_n)$, are Kähler functions.

We see from Proposition 7.10 that the map $\pi_{P_n} : P(C^n) \to P_n$ behaves like the canonical projection of a Kählerification. It is thus a natural “completion” of the map $\pi_{P_n^\times} : P(C)^\times \to P_n^\times$, and formally we have $(P_n)^C = P(C^n)$.

# 8 Kählerification and the geometrical formulation of quantum mechanics

The goal of this section is to rederive the geometrical formulation of quantum mechanics in finite dimension (based on the Kähler properties of $P(C^n)$ as in [AS99]), using a statistically oriented approach through the equation $(P_n^\times)^C = P(C^n)^\times$ and its formal version $(P_n^\times)^C = P(C^n)$.

This exercise is necessary, for we want to express all the relevant quantities of the geometrical formulation in terms of statistical concepts, aiming to generalize them to situations where $P_n^\times$ is replaced by a more general exponential family $E$.

Let us start with the following “statistical” characterization of Kähler functions on the complex projective space (see also Corollary 5.8).

**Proposition 8.1.** Let $f : P(C^n) \to \mathbb{R}$ be a smooth function. Then, $f$ is a Kähler function if and only if there exist a random variable $X : \Omega = \{x_1, ..., x_n\} \to \mathbb{R}$ and an unitary matrix $U \in U(n)$ such that

$$f([z]) = \int_{\Omega} X(x) \left( (\pi_{P_n} \circ \Phi_U)([z]) \right)(x)dx,$$

where $\pi_{P_n} : P(C^n) \to P_n$ is the map considered at the end of [7] and where $\Phi_U$ is the holomorphic isometry of $P(C^n)$ defined by $\Phi_U([z]) = [U \cdot z]$. 

18
Proof. Simply use Lemma 7.6 together with the usual spectral decomposition theorem.

Proposition 8.1 implies that every Kähler function on \( \mathbb{P}(\mathbb{C}^n) \) can be realized as an expectation of the form \( f([z]) = E_{\pi \Phi_U([z])}(X) \), where \( X : \Omega \to \mathbb{R} \) is a random variable. The image \( \text{Im}(X) \) of \( X \) has thus an important statistical meaning that we would like to relate with the usual spectrum of \( f \) as defined in the geometrical formulation of quantum mechanics.

To this end, we could use Lemma 7.6. Proposition 8.1 and relate \( \text{Im}(X) \) with the usual spectrum of an appropriate Hermitian matrix. We prefer, however, to use a generalization of a statistical result that we now present.

Recall that if \( X : \Omega \to \mathbb{R} \) is a random variable, and if \( p \in \mathcal{P}_n \) is a given probability, then the variance of \( X \) with respect to \( p \) is given by \( V_p(X) = E_p((X - E_p(X))^2) \).

**Proposition 8.2 (Cramér-Rao equality).** Let \( f([z]) = \int_{\Omega} X(x)\left(\pi_{\mathbb{C}^n} \circ \Phi_U([z])\right)(x)dx \) be a Kähler function on \( \mathbb{P}(\mathbb{C}^n) \). For all \([z] \in \mathbb{P}(\mathbb{C}^n)\), we have:

\[
V_p\left(\pi_{\mathbb{C}^n} \circ \Phi_U([z])\right)(X) = \frac{1}{4} \| \text{grad}(f)\|_{[z]}^2,
\]

where \( \text{grad}(f) \) denotes the Riemannian gradient of \( f \) with respect to the Fubini-Study metric \( g_{FS} \).

**Proof.** If \( f : \mathbb{P}(\mathbb{C}^n) \to \mathbb{R} \) is a (non-necessarily Kähler) function, and if \( \phi : \mathbb{P}(\mathbb{C}^n) \to \mathbb{P}(\mathbb{C}^n) \) is an isometry, then, for \([z] \in \mathbb{P}(\mathbb{C}^n)\), we have:

\[
\text{grad}(f \circ \phi)\|_{[z]} = (\phi^{-1})_*\left(\text{grad}(f)\right)\|_{[\phi([z])]}.
\]

From this formula, and the fact that \( \Phi_U \) is an isometry, we see that it is sufficient to show the proposition for \( U = I_n \). So let \( X : \Omega \to \mathbb{R} \) be a random variable and assume that \( f([z]) = \sum_{k=1}^n X(x_k)\pi_{\mathbb{C}^n}([z])(x_k) \).

Taking into account Proposition 7.2, we see that the function \( \tau : f \circ \tau : TP_n^x \to \mathbb{R} \) satisfies

\[
\tau_{\tau([u])}\| \text{grad}(\tau)\|_{[u]} = \frac{1}{4} \| \text{grad}(f)\|_{[\tau([u])]}^2,
\]

where \( \text{grad}(\tau) \) denotes the Riemannian gradient of \( \tau \) with respect to the Riemannian metric \( g \) (see the definition of \( g \) in Proposition 7.2). Moreover, for \( A_{[u]} = ([u], [v], [w]) = d/dt|_0[u + tw - E_{p(t)}(u + tw)]p(t) \) as in Lemma 7.1 and taking into account the relation \( \left(\pi_{\mathbb{C}^n} \circ \tau\right)([u]) = p \), we see that

\[
g_{[u]}(\text{grad}(\tau)_{[u]}p, A_{[u]}) = \frac{d}{dt}|_0[tu + tw - E_{p(t)}(u + tw)]p(t) = \sum_{k=1}^n p(t)(x_k)X(x_k) = \sum_{k=1}^n \tau_{\tau([u])}\| \text{grad}(\tau)\|_{[u]}^2,
\]

and thus,

\[
\tau_{\tau([u])}\| \text{grad}(\tau)\|_{[u]} = \left([u], [X - E_p(X)]p, [0]p\right).
\]

From this equation, and taking into account \( 7.1 \), we get

\[
\| \text{grad}(\tau)_{[u]}^2 = E_p((X - E_p(X))^2) = V_p(X).
\]

The proposition is now a consequence of this last equation together with \( 7.1 \) and the fact that \( \tau(TP_n^x) \) is dense in \( \mathbb{P}(\mathbb{C}^n) \).
Remark 8.3. Proposition \ref{prop:cramer-rao} is a direct generalization of a formula which is well known in the context of information geometry, namely:

\begin{equation}
\| \text{grad}(E_p(X)) \|^2 = V_p(X),
\end{equation}

where \( E_p(X) \) denotes the function \( P_n^X \rightarrow \mathbb{R}, p \mapsto E_p(X) \) \((X : \Omega \rightarrow \mathbb{R} \text{ being a given random variable})\), and where the norm and the Riemannian gradient are taken with respect to the Fisher metric \( h_F \). The above formula is sometimes called Cramér-Rao equality for, it allows to recover the usual Cramér-Rao inequality, the latter being, roughly, an inequality which gives a “lower bound” for the variance-covariance matrix of an unbiased estimator on a given statistical model \( S \) (see \cite{Arnold} for details).

Remark 8.4. More generally, if \( \mathcal{E} \) is an exponential family whose elements are of the form \( p(x; \theta) = \exp \{ C(x) + \sum_{i=1}^{n} \theta_i F_i(x) - \psi(\theta) \} \) on a fixed measured space \((\Omega, dx)\) (see Definition \ref{def:exponential-family}), and if \( X : \Omega \rightarrow \mathbb{R} \) is a linear combination of \( 1, F_1, ..., F_n \), then one can easily show the following identity

\begin{equation}
\| \text{grad}(E_p(X)) \|^2 = V_p(X),
\end{equation}

where \( E_p(X) \) is viewed as the function \( \mathcal{E} \rightarrow \mathbb{R}, p \mapsto E_p(X) \) and where the norm and the gradient are taken with respect to the Fisher metric \( h_F \).

By inspection of the Cramér-Rao equality (as formulated in Proposition \ref{prop:cramer-rao}), one deduces easily the following corollary:

Corollary 8.5. Let \( f([z]) = \int_{\Omega} X(x) \| [\pi_{P_n} \circ \Phi_U([z])] \| (x) dx \) be a Kähler function on \( \mathbb{P}(\mathbb{C}^n) \). Then a real number \( \lambda \) belongs to \( \text{Im}(X) \) if and only if \( \lambda \) is a critical value of \( f \), i.e. if and only if

\begin{equation}
\exists [z] \in \mathbb{P}(\mathbb{C}^n) \text{ such that } f_{|[z]} = 0 \text{ and } f([z]) = \lambda.
\end{equation}

The above corollary implies that the set \( \text{Im}(X) \) doesn’t depend on the particular decomposition of \( f \) given in Proposition \ref{prop:critical-values}. We can thus give the following definition:

Definition 8.6. The spectrum of a Kähler function \( f([z]) = \int_{\Omega} X(x) \| [\pi_{P_n} \circ \Phi_U([z])] \| (x) dx \) on \( \mathbb{P}(\mathbb{C}^n) \) is the subset of \( \mathbb{R} \) given by

\begin{equation}
\text{spec}(f) := \text{Im}(X).
\end{equation}

This is the set of all critical values of \( f \).

Following \cite{Arnold}, we shall call elements of \( \text{spec}(f) \) eigenvalues and the corresponding critical points eigenpoints. These are the geometrical analogues, in the geometrical formulation, of the usual eigenvalues and eigenvectors of Hermitian matrices used in the standard formulation of quantum mechanics.

Given a Kähler function \( f([z]) = \int_{\Omega} X(x) \| [\pi_{P_n} \circ \Phi_U([z])] \| (x) dx \) and a point \([z] \in \mathbb{P}(\mathbb{C}^n)\), there is an obvious associated probability \( P_{f,[z]} \) on \( \text{spec}(f) \) :

\begin{equation}
P_{f,[z]}(A) := \int_{X^{-1}(A)} \| [\pi_{P_n} \circ \Phi_U([z])] \| (x) dx,
\end{equation}

where \( A \subseteq \text{spec}(f) \) is a given subset. This is the pushforward of \( \| [\pi_{P_n} \circ \Phi_U([z])] \| (x) dx \) via the random variable \( X : \Omega \rightarrow \mathbb{R} \).

As for \( \text{spec}(f) \), we would like to show that this probability doesn’t depend on the particular decomposition of \( f \) given in \ref{prop:critical-values}. To this end, we introduce, for a given real number \( \lambda \), the following space:

\begin{equation}
M_{f,\lambda} := \{ [z] \in \mathbb{P}(\mathbb{C}^n) \mid f_{|[z]} = 0 \text{ and } f([z]) = \lambda \}.
\end{equation}

Observe that if \( \lambda \notin \text{spec}(f) \), then \( M_{f,\lambda} = \emptyset \).

\footnote{The absence of the factor 1/4 in \ref{eq:grad-squared} compared to \ref{eq:grad-squared} is due to the normalizing factor of the Fubini-Study metric \( g_{FS} \) used throughout this section (see also Proposition \ref{prop:fisher-metric}).}
Lemma 8.7. Let $\lambda \in \text{spec}(f)$ be an eigenvalue of a given Kähler function $f([z]) = \int X(x)(\pi\nu_n \circ \Phi_U)([z]) dx$, and let us fix, for notational convenience, some indices $k_1, \ldots, k_{m(\lambda)} \in \{1, \ldots, n\}$ so that we can write $X^{-1}(\lambda) = \{x_{k_1}, \ldots, x_{k_{m(\lambda)}}\}$. Let us also denote by $\{e_1, \ldots, e_n\}$ the canonical basis for $\mathbb{C}^n$. Then,

$$M_{f,\lambda} = \left\{ U^* \cdot [c_1 \cdot e_{k_1} + \ldots + c_{m(\lambda)} \cdot e_{k_{m(\lambda)}}] \in \mathbb{P}(\mathbb{C}^n) \mid c_1, ..., c_{m(\lambda)} \in \mathbb{C} \right\}. \quad (85)$$

In particular, $M_{f,\lambda} \cong \mathbb{P}(\mathbb{C}^{m(\lambda)})$.

Proof. Let $[z]$ be an element of $\mathbb{P}(\mathbb{C}^n)$. We have:

$$[z] \in M_{f,\lambda} \iff \left( f_{\ast}[z] = 0 \text{ and } f([z]) = \lambda \right) \iff V_{(\pi\nu_n \circ \Phi_U)([z])}(X) = 0 \text{ and } E_{(\pi\nu_n \circ \Phi_U)([z])}(X) = \lambda \iff \sum_{k=1}^{n} (X(x_k) - \lambda)^2(\pi\nu_n \circ \Phi_U)([z])(x_k) = 0 \text{ and } E_{(\pi\nu_n \circ \Phi_U)([z])}(X) = \lambda \iff \left\{ (X(x_k) - \lambda)^2(\pi\nu_n \circ \Phi_U)([z])(x_k) = 0 \text{ for all } k \in \{1, \ldots, n\} \right. \text{ and } E_{(\pi\nu_n \circ \Phi_U)([z])}(X) = \lambda \iff \left\{ (\pi\nu_n \circ \Phi_U)([z])(x_k) = 0 \text{ for all } k \in \{1, \ldots, n\} - \{k_1, \ldots, k_{m(\lambda)}\} \right. \text{ and } \left. E_{(\pi\nu_n \circ \Phi_U)([z])}(X) = \lambda \right. \iff (\pi\nu_n \circ \Phi_U)([z])(x_k) = 0 \text{ for all } k \in \{1, \ldots, n\} - \{k_1, \ldots, k_{m(\lambda)}\}. \quad (86)$$

Now observe that for $k \in \{1, \ldots, n\}$,

$$\langle \pi\nu_n \circ \Phi_U([z])(x_k) \rangle = \frac{\langle U \cdot z, e_k \rangle^2}{\langle z, z \rangle} = \frac{|\langle z, U^* e_k \rangle|^2}{\langle z, z \rangle}, \quad (87)$$

and thus,

$$[z] \in M_{f,\lambda} \iff \langle z, U^* e_k \rangle = 0 \text{ for all } k \in \{1, \ldots, n\} - \{k_1, \ldots, k_{m(\lambda)}\} \quad (88)$$

from which the lemma follows. \qed

From Lemma 8.7 we see that the cardinal of $X^{-1}(\lambda)$ doesn’t depend on the decomposition of $f$; we shall call this number the multiplicity of the eigenvalue $\lambda$ and denote it by $m(\lambda)$. As we already saw, $M_{f,\lambda} \cong \mathbb{P}(\mathbb{C}^{m(\lambda)})$. In [AS99], $M_{f,\lambda}$ is called the eigenmanifold of $f$ associated to $\lambda$.

Now recall that the geodesic distance $d(\cdot, \cdot)$ on $\mathbb{P}(\mathbb{C}^n)$ induced by the Fubini-Study metric $g_{FS}$ is given, for $[z], [w] \in \mathbb{P}(\mathbb{C}^n)$, by:

$$d([z], [w]) = \cos^{-1}\left( \frac{|\langle z, w \rangle|}{\|z\| \cdot \|w\|} \right). \quad (89)$$

The above formula together with the Cauchy-Schwarz inequality readily implies the following lemma:

Lemma 8.8. Let $\lambda \in \text{spec}(f)$ be an eigenvalue and let $[z]$ be a point in $\mathbb{P}(\mathbb{C}^n)$. Then there exists a unique point $p_{M_{f,\lambda}}([z])$ in $M_{f,\lambda}$ verifying

$$d\left(p_{M_{f,\lambda}}([z]), [z] \right) < d([w], [z]) \quad (90)$$

for all $[w] \in M_{f,\lambda}$ such that $[w] \neq p_{M_{f,\lambda}}([z])$. Moreover, if $X^{-1}(\lambda) = \{x_{k_1}, \ldots, x_{k_{m(\lambda)}}\}$, then

$$p_{M_{f,\lambda}}([z]) = \left[ \sum_{i=1}^{m(\lambda)} \langle U^* e_{k_i}, z \rangle \cdot U^* e_{k_i} \right]. \quad (91)$$
Clearly, if \([z] \in \mathbb{P}(\mathbb{C}^n)\), then \(d([[z]], pM_{f,\lambda}([[z]])\) is the geodesic distance between \([z]\) and the subset \(M_{f,\lambda};\) we shall write \(d([[z]], M_{f,\lambda}) = d([[z]], pM_{f,\lambda}([[z]])\).

**Proposition 8.9.** Let \(f([[z]]) = \int_{\Omega} X(x)[(\pi_{P_n} \circ \Phi_U)([[z]])](x)dx\) be a Kähler function, and let \(\lambda \in \text{spec}(f)\) be an eigenvalue. For \([[z]] \in \mathbb{P}(\mathbb{C}^n)\), we have:

\[
\int_{X^{-1}(\lambda)} \left[(\pi_{P_n} \circ \Phi_U)([[z]])\right](x)dx = \cos^2 \left(d([[z]], M_{f,\lambda})\right) .
\]

**Proof.** Let \(z \in \mathbb{C}^n\) be a normalized vector, and assume that \(X^{-1}(\lambda) = \{x_1, ..., x_{m(\lambda)}\}\). If \(z_\lambda := \sum_{i=1}^{m(\lambda)} (U^*_i e_{k_i}, z) \cdot U^*_i e_{k_i}\), then clearly \(<z_\lambda, z> = <z_\lambda, z_\lambda>\), and according to Lemma \(8.8\) \([z_\lambda] = pM_{f,\lambda}([[z]])\) and \(d([[z]], [z_\lambda]) = d([[z]], M_{f,\lambda})\). Hence,

\[
\int_{X^{-1}(\lambda)} \left[(\pi_{P_n} \circ \Phi_U)([[z]])\right](x)dx = \sum_{i=1}^{m(\lambda)} \left|\langle (\pi_{P_n} \circ \Phi_U)([[z]])(x_{k_i}) \right|^2 = \sum_{i=1}^{m(\lambda)} \left|\langle U \cdot z, e_{k_i}\right|^2 = \sum_{i=1}^{m(\lambda)} \left|\langle z_\lambda, z\right|^2 = \left(\frac{d([[z]], [z_\lambda])}{\|z\| \cdot \|z_\lambda\|}\right)^2 = \left(\cos \left(d([[z]], pM_{f,\lambda}([[z]])\right)\right)^2 .
\]

which is exactly \(92\). The proposition follows. \(\square\)

A direct consequence of Proposition \(8.9\) is that the measure \(P_{f,[z]}\) on spec\(f\) defined in \(8.3\) doesn’t depend on a particular decomposition of \(f\) such as in \(8.1\).

With the above proposition, we have completed our “statistical” study of the geometrical formulation of quantum mechanics in finite dimension.

The important points are: the configuration space \(\mathbb{P}(\mathbb{C}^n)\) is the (formal) Kählerification of \(\mathcal{P}_n;\) observables are Kähler functions \(f : \mathbb{P}(\mathbb{C}^n) \to \mathbb{R}\) that can be decomposed as \(f([[z]]) = \int_{\Omega} X(x)[(\pi_{P_n} \circ \Phi_U)([[z]])](x)dx\), where \(X \in \mathcal{A}_{P_n}^*\) and where \(\Phi_U\) is a holomorphic isometry of \(\mathbb{P}(\mathbb{C})\); the spectrum of a Kähler function is \(\text{Im}(X)\) and its associated probability is \(P_{f,[z]}(\lambda) = \int_{X^{-1}(\lambda)} X(x)[(\pi_{P_n} \circ \Phi_U)([[z]])](x)dx\).

As we see, these quantities depends only on the exponential structure of \(\mathcal{P}_n^\mathcal{E}\). Thus, we can try to generalize them to a given exponential family, as follows.

Let \(\mathcal{E}\) be an exponential family defined on a measured space \((\Omega, dx)\) with elements of the form \(p(x; \theta) = \exp\{C(x) + \sum_{i=1}^{n} \theta_i F_i(x) - \psi(\theta)\}\), \(\mathcal{A}_\mathcal{E} := \text{Vect}_\mathbb{R}\{1, F_1, ..., F_n\}\) and \(\pi_\mathcal{E} : \mathcal{E}^\mathcal{E} \to \mathcal{E}\) its associated Kählerification.

Regarding \(\mathcal{E}\) as the underlying statistical model of a “generalized quantum system”, we are led to the following definitions:

- **Configuration space**: \(\mathcal{E}^\mathcal{E}\), viewed as a Kähler manifold,
- **Observables**: this is the set of functions \(f : \mathcal{E}^\mathcal{E} \to \mathbb{R}\) of the form

\[
\mathcal{E}^\mathcal{E} \to \mathbb{R} , \quad z \mapsto \int_{\Omega} X(x)[(\pi_\mathcal{E} \circ \Phi)(z)](x)dx ,
\]

where \(X \in \mathcal{A}_\mathcal{E}\) and where \(\Phi\) is a holomorphic isometry of \(\mathcal{E}^\mathcal{E}\). Such a function is necessarily a Kähler function according to Proposition \(8.8\).

\(\text{Disclaimer:}\) Although being physically clear, this statement has still, on mathematical grounds, to be clarified since one has to make precise the passage from \(\mathbb{P}(\mathbb{C}^n)\) to its natural “completion” \(\mathbb{P}(\mathbb{C}^n)\). In this paper, we shall not treat this technical question, preferring to focus on the general procedure and the physical applications.
\textbf{Dynamics}: it is given by the flow of the Hamiltonian vector field $X_H$ associated to a given observable $H : \mathcal{E}^C \to \mathbb{R}$ with respect to the natural symplectic form of the Kähler manifold $\mathcal{E}^C$.

\textbf{Spectrum of an observable}: the spectrum of an observable $f$ as in (94) is given by the image of the random variable $X$,

$$\text{spec}(f) := \text{Im}(X),$$

(95)

\textbf{Probabilities associated to an observable}: the probability that an observable $f$ as in (94) yields upon measurement an eigenvalue belonging to a subset $A \subseteq \text{spec}(f)$ while the system is in the state $z \in \mathcal{E}^C$ is:

$$P_{f,z}(A) := \int_{\pi^{-1}(A)} \left[ (\pi_2 \circ \phi)(z) \right] dx.$$ (96)

\textbf{Remark 8.10.} Usually, the decomposition of a Kähler function $f$ as in (94) is not unique, and thus $\text{spec}(f)$ and $P_{f,z}$ are only well defined when invariance properties are met.

Of course, and from a physical point of view, the above definitions cannot be taken too literally. For example when $\text{spec}(f)$ is not unique, and thus $\text{spec}(f)$ and $P_{f,z}$ are only well defined when invariance properties are met.

Despite these technical difficulties and ambiguities, we shall use the above definitions as a basis for our physical investigations, and adapt them in an obvious way when a natural “completion” exists. As we will see, this already leads to interesting physical results.

9 \hspace{1cm} \textbf{Binomial distribution and the spin of a particle}

Let $B(n, q)$ be the space of binomial distributions defined over $\Omega := \{0, ..., n\}$. By definition, an element $p \in B(n, q)$ is characterized by a real parameter $q \in [0, 1]$ verifying, for $k \in \Omega$, $p(k) = \binom{n}{k} q^k (1 - q)^{n-k}$, where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. The set of binomial distributions forms a 1-dimensional statistical manifold parameterized by $q$ and is easily seen to be an exponential family, for one may write

$$p(k) = \binom{n}{k} q^k (1 - q)^{n-k} = \exp \left\{ \ln \binom{n}{k} + k \theta - n \ln (1 + \exp \theta) \right\},$$

(97)

where $\theta := \ln(\frac{q}{1-q})$. In particular, setting $C(k) := \ln \binom{n}{k}$, $F(k) := k$ and $\psi(\theta) := n \ln (1 + \exp \theta)$, one has

$$p(k) = \exp\{C(k) + \theta \cdot F(k) - \psi(\theta)\}.$$  

Let $S^2 := \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$ be the unit sphere endowed with its natural Kähler structure $(g_{S^2}, J_{S^2}, \omega_{S^2})$ and let us write $(S^2)^\times := S^2 - \{(1, 0, 0), (-1, 0, 0)\}$.

\textbf{Proposition 9.1.} If $S^2$ is endowed with the Kähler structure $(n \cdot g_{S^2}, J_{S^2}, n \cdot \omega_{S^2})$ (i.e. its natural Kähler structure is multiplied by $n$), then

$$B(n, q)^C \cong (S^2)^\times,$$

(98)

and in term of this identification, the map $\pi_{B(n, q)} : B(n, q)^C \to B(n, q)$ becomes

$$\pi_{B(n, q)} : (S^2)^\times \to B(n, q), \quad \pi_{B(n, q)}(x, y, z)(k) = \frac{1}{2^n} \binom{n}{k} (1 + x)^k (1 - x)^{n-k}.$$ (99)
Proposition 9.1 follows from direct computations. Indeed, in term of the natural parameter \( \theta \in \mathbb{R} \), the Fisher metric \( h_F \) on \( \mathcal{B}(n,q) \) is (see \cite{2})

\[
h_F(\theta) = \frac{n \exp \theta}{(1 + \exp \theta)^2}
\]

from which one easily sees that the canonical structure \( (g, J, \omega) \) of \( T\mathcal{B}(n,q) \) is, using the identification \( T\mathcal{B}(n,q) \cong \mathbb{R}^2, \partial \partial_0 \mapsto (\theta, \dot{\theta}) \) as well as \cite{13},

\[
g(\dot{\theta}, \dot{\theta}) = \frac{n \exp \theta}{(1 + \exp \theta)^2} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \omega(\dot{\theta}, \dot{\theta}) = \frac{n \exp \theta}{(1 + \exp \theta)^2} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.
\]

A basis for \( \mathcal{K}\mathcal{F}(T\mathcal{B}(n,q)) \) is easily seen to be

\[
1, \quad \frac{\cos(\dot{\theta}/2)}{\cosh(\theta/2)}, \quad \frac{\sin(\dot{\theta}/2)}{\cosh(\theta/2)},
\]

where \( \tanh(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} \) and \( \cosh(x) = \frac{\exp(x) + \exp(-x)}{2} \).

As a Lie algebra, the space \( \mathcal{K}\mathcal{F}(T\mathcal{B}(n,q)) \), endowed with the natural Poisson bracket associated to \( \omega \), is isomorphic to the Lie algebra \( \mathfrak{u}(2) \) of the group of unitary matrices \( \mathfrak{u}(2) \) via the isomorphism

\[
1 \mapsto \frac{1}{2n} \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}, \quad \tanh(\theta/2) \mapsto \frac{1}{2n} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},
\]

\[
\frac{\cos(\dot{\theta}/2)}{\cosh(\theta/2)} \mapsto \frac{1}{2n} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \frac{\sin(\dot{\theta}/2)}{\cosh(\theta/2)} \mapsto \frac{1}{2n} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.
\]

Clearly (see \cite{12}), the group \( \Gamma(\mathcal{B}(n,q)) \) is isomorphic to \( \mathbb{Z} \), its natural action on \( T\mathcal{B}(n,q) \) being \( \kappa \cdot (\theta, \dot{\theta}) = (\theta, \dot{\theta} + 4k\pi) \), which is obviously free and proper, and the quotient \( T\mathcal{B}(n,q)/\Gamma(\mathcal{B}(n,q)) \) is diffeomorphic to \( (S^2)^\times \) via the map

\[
[(\theta, \dot{\theta})] \mapsto \left( \tanh(\theta/2), \frac{\cos(\dot{\theta}/2)}{\cosh(\theta/2)}, \frac{\sin(\dot{\theta}/2)}{\cosh(\theta/2)} \right),
\]

where \( [(\theta, \dot{\theta})] := Z \cdot (\theta, \dot{\theta}) = \{ (\theta, \dot{\theta} + 4k\pi) \in \mathbb{R}^2 \mid k \in \mathbb{Z} \} \).

A direct calculation shows that if the canonical Kähler structure of \( (S^2)^\times \) is weighted by \( n \), then \cite{10} defines a map which is an isomorphism of Kähler manifolds, whence Proposition 9.1

The canonical projection \( \pi_{\mathcal{B}(n,q)} : (S^2)^\times \cong \mathcal{B}(n,q)^C \rightarrow \mathcal{B}(n,q) \) can be naturally extended to the whole sphere \( S^2 \) provided we adjoint two elements to \( \mathcal{B}(n,q) \), namely the Dirac measures \( \delta_0 \) and \( \delta_n \) defined, for \( k \in \{0, ..., n\} \), by

\[
\delta_0(k) = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases}, \quad \delta_n(k) = \begin{cases} 1 & \text{if } k = n \\ 0 & \text{if } k \neq n \end{cases}.
\]

Let us denote \( \mathcal{B}(n,q) := \mathcal{B}(n,q) \cup \{ \delta_0, \delta_n \} \) (disjoint union). Clearly, the map \( \pi_{\mathcal{B}(n,q)} : (S^2)^\times \rightarrow \mathcal{B}(n,q) \) extends uniquely as a continuous map \( \pi_{\mathcal{B}(n,q)} : S^2 \rightarrow \mathcal{B}(n,q) \), with \( \pi_{\mathcal{B}(n,q)}(-1,0,0) := \delta_0, \quad \pi_{\mathcal{B}(n,q)}(1,0,0) := \delta_n \), making the following diagram commutative:

\[
\begin{array}{ccc}
(S^2)^\times & \xrightarrow{\pi_{\mathcal{B}(n,q)}} & \mathcal{B}(n,q) \\
\downarrow{\pi_{\mathcal{B}(n,q)}} & & \\
S^2 & \xrightarrow{j} & \mathcal{B}(n,q)
\end{array}
\]
Proposition 9.2. Let \( f : S^2 \to \mathbb{R} \) be a smooth function. Then \( f \) is a Kähler function if and only if there exist \( X \in \mathcal{A}(n,q) \) and \( \phi \in SO(3) \) such that \( f \) can be written

\[
f(x, y, z) = \int_{\Omega} X(k) [(\pi_\mathcal{B}(n,q) \circ \phi)(x, y, z)](k) dk, \quad ((x, y, z) \in S^2)
\]

where \( \pi_\mathcal{B}(n,q) : S^2 \to \mathcal{B}(n,q) \) is the canonical projection coming from the Kählerification of \( \mathcal{B}(n,q) \).

Proof. Let \( X(k) = \alpha + \beta \cdot k \in \mathcal{A}(n,q) \) be arbitrary and let \( \phi \in SO(3) \) be an isometry such that, using a matrix representation,

\[
\phi = \begin{pmatrix} a & b & c \\ * & * & * \\ * & * & *
\end{pmatrix}
\]

(in particular, the real numbers \( a, b, c \) satisfy \( \|(a, b, c)\| = 1 \), i.e. \( a^2 + b^2 + c^2 = 1 \)).

A simple calculation shows that for \( (x, y, z) \in S^2 \),

\[
\int_{\Omega} X(k) [(\pi_\mathcal{B}(n,q) \circ \phi)(x, y, z)](k) dk = \alpha + \beta \frac{n}{2} + \beta \frac{n}{2}(ax + by + cz)
\]

which is a Kähler function on \( S^2 \) since it is a linear combination of 1, \( x, y, z \).

Reciprocally, if \( u_0, u, v, w \in \mathbb{R} \) with \( (u, v, w) \neq 0 \), then the equation \( \int_{\Omega} X(k) [(\pi_\mathcal{B}(n,q) \circ \phi)(x, y, z)](k) dk = u_0 + ux + vy + wz \), with unknowns \( \alpha, \beta, a, b, c \), admits as a solution

\[
\alpha = u_0 \pm \|(u, v, w)\|, \quad \beta = \mp \frac{2}{n} \|(u, v, w)\|, \quad (a, b, c) = \mp \frac{1}{\|(u, v, w)\|}(u, v, w),
\]

where \( \pm = - \) if \( \beta > 0 \) and \( \pm = + \) if \( \beta < 0 \), and where \( \| . \| \) is the Euclidean norm. If \( (u, v, w) = 0 \), then a solution is given by \( \alpha = u_0 \) and \( \beta = 0 \) (\( \phi \) being arbitrary). The proposition follows.

Following our discussion at the end of \( \$ \) we want to define the spectrum \( \text{spec}(f) \) of a Kähler function \( f(x, y, z) = \int_{\Omega} X(k) [(\pi_\mathcal{B}(n,q) \circ \phi)(x, y, z)](k) dk \) on \( S^2 \) as \( \text{Im}(X) \), and its associated probability on \( \text{spec}(f) \) as \( P_{f,(x,y,z)}(A) = \int_{X^{-1}(A)} [(\pi_\mathcal{B}(n,q) \circ \phi)(x, y, z)](k) dk \). For this to be consistent, we need to check that these formulas are independent of the decomposition of \( f \) given in Proposition 9.2.

Proposition 9.3. Let \( f(x, y, z) = u_0 + ux + vy + wz \) be a Kähler function on \( S^2 \). Then the spectrum \( \text{spec}(f) \) and the probability \( P_{f,(x,y,z)} \) are well defined, and we have \( \text{spec}(f) = \{\lambda_0, \ldots, \lambda_n\} \), where

\[
\lambda_k = u_0 + \frac{2}{n} \|(u, v, w)\| \cdot \left( -\frac{n}{2} + k \right)
\]
Proof of Proposition 9.3. Let $(u,v,w) \in \mathbb{R}^3$ be a vector whose Euclidean norm is $j := n/2$, and let $f : S^2 \rightarrow \mathbb{R}$ be the Kähler function defined by $f(x,y,z) := ux + vy + wz$. Then,

\[
P_{f,(x,y,z)}(\lambda_k) = \frac{1}{2\pi n} \left( \frac{n}{k} \right) \left(1 + \frac{ux + vy + wz}{\|u,v,w\|} \right)^k \left(1 - \frac{ux + vy + wz}{\|u,v,w\|} \right)^{n-k}.
\]

(112)

Corollary 9.4. Let $(u,v,w) \in \mathbb{R}^3$ be a vector whose Euclidean norm is $j := n/2$, and let $f : S^2 \rightarrow \mathbb{R}$ be the Kähler function defined by $f(x,y,z) := ux + vy + wz$. Then,

- $\text{spec}(f) = \{-j, -j+1, \ldots, j-1, j\}$,
- $P_{f,(x,y,z)}(-j+k) = \left\{\frac{n}{k}\right\} \left(\cos^2(\theta/2)\right)^k \left(\sin^2(\theta/2)\right)^{n-k},$ 

where $\theta$ is an angle satisfying $\frac{ux + vy + wz}{\|u,v,w\|} = \cos(\theta)$.

Remark 9.5. As mentioned in the introduction, [113] gives the probability that a spin-$j$ particle entering a second Stern-Gerlach device with maximum spin state (see Footnote 9) is deflected into the $(-j+k)$-th outgoing beam, where $\theta$ is the angle between the two magnetic fields produced by the two Stern-Gerlach devices (see for example [Mar62]). We will see subsequently how to obtain the probabilities corresponding to an incoming particle when the eigenvalue of its spin operator along the magnetic field of the first Stern-Gerlach device is arbitrary.

Proof of Proposition 114. Let $f(x,y,z) = \int_{\Omega} X(k) \left[ (\pi_{\Omega,q}) \circ \phi(x,y,z) \right] (k) dk$ be a Kähler function on $S^2$ with $X(k) = \alpha + \beta \cdot k$ $(\alpha, \beta \in \mathbb{R})$ and $\phi$ having a matrix representation as in [108]. We have to show that $\text{spec}(f) := \text{Im}(\alpha + \beta \cdot k) = \{\alpha, \alpha + \beta, \ldots, \alpha + n \cdot \beta\}$ is independent of the decomposition of $f$. For this, we need to check that if $f$ can be written $f(x,y,z) = \sum_{k=0}^{n} (\overline{\alpha} + \overline{\beta} \cdot k) [ (\pi_{\Omega,q}) \circ \phi(x,y,z) ] (k)$ with different $\overline{\alpha}, \overline{\beta} \in \mathbb{R}$ and a different $\phi \in SO(3)$ (with different $\overline{\alpha}, \overline{\beta}, \overline{\beta} \in \mathbb{R}$), then $\text{Im}(\alpha + \beta \cdot k) = \text{Im}(\overline{\alpha} + \overline{\beta} \cdot k)$. To this end, observe that if $\sum_{k=0}^{n} (\alpha + \beta \cdot k) [ (\pi_{\Omega,q}) \circ \phi(x,y,z) ] (k) = \sum_{k=0}^{n} (\overline{\alpha} + \overline{\beta} \cdot k) [ (\pi_{\Omega,q}) \circ \phi(x,y,z) ] (k)$ for all $(x,y,z) \in S^2$, then

\[
\alpha + \beta n \frac{2}{n} = \overline{\alpha} + \overline{\beta} n \frac{2}{n} \quad \text{and} \quad \beta n \frac{2}{n} \cdot (a,b,c) = \overline{\beta} n \frac{2}{n} \cdot (\overline{a}, \overline{b}, \overline{c}).
\]

(115)

Taking into account the fact that $\| (a,b,c) \| = \| (\overline{a}, \overline{b}, \overline{c}) \| = 1$, one immediately sees that $|\beta| = |\overline{\beta}|$, and we are led to the following three possibilities:

\[
\beta = \overline{\beta} = 0, \quad \alpha = \overline{\alpha} \quad \text{or} \quad \beta \neq 0, \quad \beta = \overline{\beta}, \quad \alpha = \overline{\alpha}, \quad (a,b,c) = (\overline{a}, \overline{b}, \overline{c})
\]

(116)

or \( \beta \neq 0, \beta = -\overline{\beta}, \alpha = \overline{\alpha} - n \cdot \beta, \quad (a,b,c) = -(\overline{a}, \overline{b}, \overline{c}). \)

The only ambiguity is for the last case for which we have :

\[
\text{Im}(\overline{\alpha} + \overline{\beta} \cdot k) = \text{Im}(\alpha + \beta \cdot n - \beta \cdot k) = \{\alpha + \beta \cdot n - \beta \cdot k \mid k = 0, \ldots, n\}
\]

\[
= \{\alpha + \beta \cdot (n-k) \mid k = 0, \ldots, n\} = \{\alpha + \beta \cdot k \mid k = 0, \ldots, n\}
\]

(117)

Hence $\text{Im}(\alpha + \beta \cdot k) = \text{Im}(\overline{\alpha} + \overline{\beta} \cdot k)$. It follows that $\text{spec}(f)$ is well defined.

In terms of $u_0, u, v, w$, and assuming $\beta > 0$ for simplicity (the case $\beta \leq 0$ leads to the same result), we have, using (110),

\[
\text{spec}(f) = \text{Im}(\alpha + \beta \cdot k) = \{\alpha + k \cdot \beta \mid k = 0, \ldots, n\}
\]

\[
= \left\{ u_0 - \|(u,v,w)\| + k \cdot \frac{2}{n} \|(u,v,w)\| \mid k = 0, \ldots, n \right\}
\]

(115)

\[
= \left\{ u_0 + \frac{2}{n} \|(u,v,w)\| \cdot \left( -\frac{n}{k} + 2 \right) \mid k = 0, \ldots, n \right\}.
\]

(118)
Observe that the map \( \Psi \) is defined on the whole sphere, i.e. also for \( \alpha \) where \( \Psi(\alpha, \beta) \). This may be done as follows.

As we already mentioned, not all the possibilities in the Stern-Gerlach experiment are exhausted with \( \{0, \pi\} \). But of course, we would like to recover all these probabilities following our "statistical approach".

Let \( j: B(n, q) \hookrightarrow TP_{n+1}^\times \) be the canonical inclusion. The composition of its derivative \( j_* : T^2B(n, q) \hookrightarrow TP_{n+1}^\times \) with the map \( \tau: TP_{n+1}^\times \to \mathbb{P}(\mathbb{C}^{n+1})^\times \) considered in \( \Box \) yields a map \( T^2B(n, q) \to \mathbb{P}(\mathbb{C}^{n+1}) \) that we would like to describe. To this end, recall that the elements of \( B(n, q) \) can be parameterized by the natural parameter \( \theta \in \mathbb{R} \) as \( p(k; \theta) := \exp \left\{ \ln \left( \frac{1}{k} \right) + k\theta - n \ln \left( 1 + \exp(\theta) \right) \right\} \) (see \( \Box \)), and that \( T^2B(n, q) \) is identified with \( \mathbb{R}^2 \) via the map \( \theta \theta_b \mapsto (\theta, \dot{\theta}) \).

**Lemma 9.6.** In terms of the natural parameter \( \theta \), the map \( \tau \circ j_* : T^2B(n, q) \to \mathbb{P}(\mathbb{C}^{n+1}) \) reads:

\[
\tau(\theta, \dot{\theta}) = \left[ p(0; \theta)^{1/2}, p(1; \theta)^{1/2} e^{i\dot{\theta}/2}, \ldots, p(n; \theta)^{1/2} e^{i\dot{\theta} n/2} \right].
\]

**Proof.** Take a smooth curve \( \theta(t) \) in \( \mathbb{R} \) and set \( \theta := \theta(0) \) and \( \dot{\theta} := \frac{d}{dt}\big|_{\theta(0)} \theta(t) \). For \( k \in \{0, \ldots, n\} \), we have:

\[
\begin{align*}
\frac{d}{dt}\big|_{\theta(0)} p(k; \theta(t)) &= \frac{d}{dt}\big|_{\theta(0)} \exp\left\{ \ln \left( \frac{n}{k} \right) + k\theta(t) - n \ln \left( 1 + \exp(\theta(t)) \right) \right\} \\
&= \frac{d}{dt}\big|_{\theta(0)} \left( \ln \left( \frac{n}{k} \right) + k\theta(t) - n \ln \left( 1 + \exp(\theta(t)) \right) \right) : p(k, \theta) \\
&= \dot{\theta} \left( k - \frac{\exp(\theta)}{1 + \exp(\theta)} \right) : p(k, \theta),
\end{align*}
\]

from which we see that \( j_* (\theta, \dot{\theta}) \) corresponds, in the exponential representation of \( TP_{n+1}^\times \) (see \( \Box \)), to the vector \( [u(\theta)]_{k} : (\theta, \dot{\theta}) \), where \( u(\theta) \in \mathbb{R}^{n+1} \) is defined, for \( k \in \{0, \ldots, n\} \), by

\[
u(\theta)_{k} = \dot{\theta} \left( k - \frac{\exp(\theta)}{1 + \exp(\theta)} \right).
\]

The lemma is now a simple consequence of \( \Box \) together with the definition of \( \tau \) (see \( \Box \)) and the homogeneity of the homogeneous coordinates of the complex projective space \( \mathbb{P}(\mathbb{C}^{n+1}) \).

Recall that the group \( \Gamma(B(n, q)) \) is isomorphic \( \mathbb{Z} \) and that its action on \( T^2B(n, q) \) is given by \( k \cdot (\theta, \dot{\theta}) = (\theta, \dot{\theta} + 4k\pi) \). Clearly, \( \tau \circ j_* \) is \( \mathbb{Z} \)-invariant, and since \( T^2B(n, q) / \Gamma(B(n, q)) \cong (S^2)^\times \) (see \( \Box \)), we get a map \( (S^2)^\times \to \mathbb{P}(\mathbb{C}^{n+1}) \) which can be conveniently described by the following parametrization of the sphere,

\[
x = \cos(\alpha), \quad y = \sin(\alpha) \cos(\beta), \quad z = \sin(\alpha) \sin(\beta),
\]

where \( \alpha \in [0, \pi], \beta \in [0, 2\pi] \). With these parameters, the map \( (S^2)^\times \to \mathbb{P}(\mathbb{C}^{n+1}) \) reads

\[
(\alpha, \beta) \mapsto \left[ \Psi(\alpha, \beta) \right],
\]

where \( \Psi(\alpha, \beta) \) is the vector in \( \mathbb{C}^{n+1} \) whose \( k \)-th component is \( (k = 0, \ldots, n) : \)

\[
\Psi(\alpha, \beta)_{k} := \left( \begin{array}{c} n \end{array} \right)^{1/2} (\cos(\alpha/2))^{k} (\sin(\alpha/2))^{n-k} e^{i\beta k}.
\]

Observe that the map \( \Psi \) is defined on the whole sphere, i.e. also for \( \alpha = 0 \) and \( \alpha = \pi \) (we agree that \( \alpha = 1 \)), and that \( \langle \Psi, \Psi \rangle = 1 \), i.e. \( \Psi \) is normalized (here \( \Psi := \Psi(\alpha, \beta) \)).
By construction, if \( pr : \mathbb{C}^{n+1} - \{0\} \to \mathbb{P}(\mathbb{C}^{n+1}) \) is the canonical projection, then we have the following commutative diagram,

\[
\begin{array}{ccc}
S^2 & \xrightarrow{\Psi} & \mathbb{C}^{n+1} \\
\downarrow{\pi_{(n,q)}} & & \downarrow{pr} \\
\mathcal{B}(n,q) & \xrightarrow{\gamma} & \mathbb{P}(\mathbb{C}^{n+1})
\end{array}
\]

(125)

where \( \gamma : \mathcal{B}(n,q) \hookrightarrow \mathbb{P}(\mathbb{C}^{n+1}) \) is the canonical injection, and where \( r : \mathbb{C}^{n+1} - \{0\} \to \mathbb{P}(\mathbb{C}^{n+1}) \) is defined by

\[
r(z_1, ..., z_{n+1})(k) := \frac{|z_k|^2}{\sum_{j=1}^{n+1} |z_j|^2}, \text{ where } z = (z_1, ..., z_{n+1}).
\]

Regarding \( S^2 \) as an embedded submanifold of \( \mathbb{P}(\mathbb{C}^{n+1}) \) via the map \( pr \circ \Psi \), we have the following proposition.

**Proposition 9.7.** Every Kähler function \( f \) on \( S^2 \) extends uniquely as a Kähler function \( \hat{f} \) on \( \mathbb{P}(\mathbb{C}^{n+1}) \), and the resulting linear map \( \mathcal{K}(S^2) \to \mathcal{K}(\mathbb{P}(\mathbb{C}^{n+1})) \), \( f \mapsto \hat{f} \) satisfies

\[
\left\{ f, g \right\} = \frac{1}{4} \{ \hat{f}, \hat{g} \} 
\]

(126)

for all \( f, g \in \mathcal{K}(S^2) \).

We will show Proposition 9.7 with a series of Lemmas.

Let \( f : S^2 \to \mathbb{R} \) be a Kähler function on the sphere, that is, a linear combinaison of \( 1, x, y, z : f(x, y, z) = u_0 + ux + vy + wz \). According to the characterization of Kähler functions on \( \mathbb{P}(\mathbb{C}^{n+1}) \) given in Lemma 7.6, the function \( f \) possesses a Kähler extension on \( \mathbb{P}(\mathbb{C}^{n+1}) \) if and only if there exists a Hermitian matrix \( Q(f) \in \text{Her}i(\mathbb{C}^{n+1}) \) such that, in terms of the parameters \( \alpha \in [0, \pi] \) and \( \beta \in [0, 2\pi] \) introduced in (122),

\[
u_0 + u \cos(\alpha) + v \sin(\alpha) \cos(\beta) + w \sin(\alpha) \sin(\beta) = \langle \Psi(\alpha, \beta), Q(f) \cdot \Psi(\alpha, \beta) \rangle
\]

(127)

for all \( \alpha \in [0, \pi] \) and all \( \beta \in [0, 2\pi] \).

**Lemma 9.8.** There exists a unique Hermitian matrix \( Q(f) \) such that (127) holds for all \( \alpha \in [0, \pi] \) and all \( \beta \in [0, 2\pi] \). It is explicitly given by

\[
Q(f)_{kk} = u_0 - u \cdot \frac{2}{n} \left( \frac{n}{2} - k \right), \quad Q(f)_{l,l+1} = \frac{1}{n} \sqrt{(n-l)(1+l)} \cdot (v-iw), \quad Q(f)_{ab} = 0,
\]

(128)

where \( k = 0, ..., n, \ l = 0, ..., n-1, \ a = 0, ..., n-2 \) and \( b \) is such that \( a + 2 \leq b \leq n \).

**Proof.** For \( a, b, k \in \{0, 1, ..., n\} \) and \( \alpha \in [0, \pi] \), set

\[
\begin{align*}
C_{ab}(\alpha) & := \binom{n}{a}^{1/2} \binom{n}{b}^{1/2} \left( \cos(\alpha/2) \right)^{a+b} \left( \sin(\alpha/2) \right)^{2n-(a+b)} \cdot Q(f)_{ab}, \\
A_{k}(\alpha) & := \sum_{j=0}^{n-k} C_{j,k+j}(\alpha).
\end{align*}
\]

28
Observe that $C_{ab} = C_{ba}$ (complex conjugate). By definition of $\Psi$ (see (124)), we have

$$
\langle \Psi(\alpha, \beta), Q(f) \cdot \Psi(\alpha, \beta) \rangle
= \sum_{a,b=0}^{n} \left( \frac{n}{a} \right)^{1/2} \left( \frac{n}{b} \right)^{1/2} \left( \cos(\alpha/2) \right)^{a+b} \left( \sin(\alpha/2) \right)^{2n-(a+b)} e^{i\beta(b-a)} Q(f)_{ab}
$$

$$
= \sum_{a,b=0}^{n} C_{ab}(\alpha) e^{i\beta(b-a)} = \sum_{k=0}^{n} \left( \sum_{a=0}^{n-k} C_{a,k+a}(\alpha) \right) e^{ik\beta} + \sum_{k=0}^{n} \left( \sum_{a=0}^{n-k} C_{a,k-a}(\alpha) \right) e^{-ik\beta}
$$

$$
= \sum_{a=0}^{n} C_{aa}(\alpha) + \sum_{k=1}^{n} A_k(\alpha) e^{-ik\beta} + \sum_{k=1}^{n} \overline{A}_k(\alpha) e^{ik\beta}
$$

$$
= A_0(\alpha) + 2 \sum_{k=1}^{n} \text{Rel}(A_k(\alpha)) \cdot \cos(k\beta) - 2 \sum_{k=1}^{n} \text{Im}(A_k(\alpha)) \cdot \sin(k\beta),
$$

and thus (127) may be rewritten:

$$
u_0 + u \cos(\alpha) + v \sin(\alpha) \cos(\beta) + w \sin(\alpha) \sin(\beta)
= A_0(\alpha) + 2 \sum_{k=1}^{n} \text{Rel}(A_k(\alpha)) \cdot \cos(k\beta) - 2 \sum_{k=1}^{n} \text{Im}(A_k(\alpha)) \cdot \sin(k\beta).
$$

Since the functions $\cos(k\beta)$ and $\sin(k'\beta)$ ($k = 0, \ldots, n$ and $k' = 1, \ldots, n$) are linearly independent, we obtain:

$$
A_0(\alpha) = u_0 + u \cos(\alpha), \quad 2 \text{Rel}(A_1(\alpha)) = v \sin(\alpha), \quad 2 \text{Rel}(A_k(\alpha)) = 0 \quad \text{for all } k \geq 2,
$$

and

$$
-2 \text{Im}(A_1(\alpha)) = w \sin(\alpha), \quad -2 \text{Im}(A_k(\alpha)) = 0 \quad \text{for all } k \geq 2.
$$

From this set of equations, one sees that $\text{Rel}(A_k(\alpha)) = \text{Im}(A_k(\alpha)) = 0$ for all $k \geq 2$, i.e. $A_k(\alpha) = 0$ for all $k \geq 2$. In view of the definition of $A_k(\alpha)$, we thus have

$$
\sum_{a=0}^{n-k} \left( \frac{n}{a} \right)^{1/2} \left( \frac{n}{a+k} \right)^{1/2} \left( \cos(\alpha/2) \right)^{2a+k} \left( \sin(\alpha/2) \right)^{2n-(2a+k)} Q(f)_{a,k+a} = 0.
$$

It is not difficult to show that the functions $\cos(\alpha/2)^k \sin(\alpha/2)^{N-k}$ ($k = 0, \ldots, N$, $N \in \mathbb{N}$) are linearly independent, and thus for all $k \geq 2$ and for all $a$ such that $0 \leq a \leq n-k$,

$$
Q(f)_{a,k+a} = Q(f)_{k,a,a} = 0.
$$

Hence, except for the three “central diagonals”, all entries of $Q(f)$ vanish; this corresponds to the third equation in (123).

Now there are still three equations in (131) we haven’t used, namely $A_0(\alpha) = u_0 + u \cos(\alpha), \quad 2 \text{Rel}(A_1(\alpha)) = v \sin(\alpha)$ and $-2 \text{Im}(A_1(\alpha)) = w \sin(\alpha)$. Using the definitions of $A_0(\alpha)$ and $A_1(\alpha)$, these equations reads

- $\sum_{a=0}^{n} \left( \frac{n}{a} \right) \left( \cos(\alpha/2) \right)^{2a} \left( \sin(\alpha/2) \right)^{2n-2a} Q(f)_{aa} = u_0 + u \cos(\alpha), \quad (134)$

- $2 \sum_{a=0}^{n-1} \left( \frac{n}{a+1} \right)^{1/2} \left( \frac{n}{a+1} \right)^{1/2} \left( \cos(\alpha/2) \right)^{2a+1} \left( \sin(\alpha/2) \right)^{2n-(2a+1)} Q(f)_{a,a+1} = (v - iw) \cdot \sin(\alpha). \quad (135)$

Using the identity $\sin(\alpha) = 2 \sin(\alpha/2) \cos(\alpha/2)$ as well as

$$
\left( \frac{n}{a} \right)^{1/2} \left( \frac{n}{a+1} \right)^{1/2} = \frac{n}{a} \sqrt{(n-a)(a+1)} = 1 - \sum_{a=0}^{n-1} \left( \frac{n}{a} \right) \left( \cos^2(\alpha/2) \right)^a \left( \sin^2(\alpha/2) \right)^{(n-1)-a}, \quad (136)
$$
one rewrites (135) as
\[
\sum_{a=0}^{n-1} \binom{n-1}{a} \frac{n}{\sqrt{(n-a)(a+1)}} Q(f)_{a,a+1} - (v - iw) \left( \cos(\alpha/2) \right)^{2a} \left( \sin(\alpha/2) \right)^{2n-2a-2} = 0 \tag{137}
\]
from which it follows that
\[
Q(f)_{a,a+1} = \frac{1}{n} \cdot \sqrt{(n-a)(a+1)} (v - iw) \quad \text{for all} \quad 0 \leq a \leq n-1.
\]
Finally, from (134) together with the identity
\[
\sum_{a=0}^{n} \binom{n}{a} \left( \frac{n}{2} - a \right) \binom{n}{a} \left( \cos(\alpha/2) \right)^{2a} \left( \sin(\alpha/2) \right)^{2n-2a-2} = -\frac{n}{2} \cos(\alpha),
\tag{138}
\]
one easily obtains the first equation in (128). The lemma follows.

**Lemma 9.9.** The map \( Q : K(S^2) \to \text{Herm}(C^{n+1}) \) satisfies
\[
Q(\{f,g\}) = -\frac{i}{2} [Q(f), Q(g)]
\tag{139}
\]
for all \( f, g \in K(S^2) \).

**Proof.** By direct computations using (128). \( \square \)

**Lemma 9.10.** For \( f, g \in K(S^2) \), \( \hat{\{f, g\}} = \frac{1}{4} \{\hat{f}, \hat{g}\} \).

**Proof.** The lemma is a consequence of Lemma 9.9 together with Lemma 7.6. Indeed, using the Lie algebra isomorphism \( u(n+1) \to K(\mathbb{P}(C^{n+1})) \), \( A \mapsto \xi^A \) given in Lemma 7.6 and the fact that \( \hat{f} = \xi^{-2iQ(f)} \), we see that
\[
\{\hat{f}, \hat{g}\} = \xi^{-2iQ(\{f,g\})} = \xi^{i[Q(f), Q(g)]} = \xi^{iQ(f)} \xi^{iQ(g)} = \{\xi^{Q(f)}, \xi^{Q(g)}\} = \{\xi^{2Q(f/2)}, \xi^{2Q(-g/2)}\} = \{-\hat{f}/2, -\hat{g}/2\} = \frac{1}{4} \{\hat{f}, \hat{g}\}.
\tag{140}
\]
The lemma follows. \( \square \)

Proposition 9.7 follows from the last three lemmas.

Since \( \text{Vect}_R \{x, y, z\} \cong su(2) \), the restriction of the map \(-1/2i \, Q \) to \( \text{Vect}_R \{x, y, z\} \) yields an unitary representation \( su(2) \to u(n+1) \) which is actually irreducible. Hence, by considering the problem of extending Kähler functions on \( S^2 \) to \( \mathbb{P}(C^{n+1}) \), we have been let to compute the irreducible unitary representations of the Lie algebra \( su(2) \), which is exactly what physicists use to describe the spin of a particle. This means that we can recover all the probabilities in the Stern-Gerlach experiment and that, in fine, all information on the spin is encoded in the binomial distribution \( B(n, q) \).

## 10 Gaussians and the quantum harmonic oscillator

Let \( \mathcal{N}(\mu, 1) \) be the set of all probability density functions defined over \( \Omega := \mathbb{R} \) by
\[
p(\xi; \mu) := \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(\mu - \xi)^2}{2} \right\},
\tag{141}
\]
where \( \xi \in \Omega \) and \( \mu \in \mathbb{R} \).
Since \( p(\xi; \mu) = \exp \{ C(\xi) + \theta \cdot F(\xi) - \psi(\theta) \} \) with \( C(\xi) = -1/2 \cdot \xi^2 \), \( \theta = \mu \), \( F(\xi) = \xi \) and \( \psi(\theta) := 1/2 \cdot \theta^2 + \ln(\sqrt{2\pi}) \), \( \mathcal{N}(\mu, 1) \) is an exponential family whose natural parameter is \( \theta = \mu \).

The Fisher metric is easily seen to be constant \( h_F(\theta) \equiv 1 \), which implies that \( T\mathcal{N}(\mu, 1) \cong \mathbb{C} \) (isomorphism of Kähler manifolds). Moreover, it is not difficult to see that the space \( \mathcal{K}(\mathbb{C}) \) of Kähler functions on \( \mathbb{C} \)

\[\mathcal{K}(\mathbb{C}) = \text{Vect}_\mathbb{R} \{1, x, y, \frac{x^2 + y^2}{2}\} \tag{142}\]

(here \( x \) and \( y \) are respectively the real and imaginary parts of \( z \in \mathbb{C} \)), with the following commutators

\[\{1, .\} = 0, \quad \{x, y\} = 1, \quad \{x, \frac{x^2 + y^2}{2}\} = y, \quad \{y, \frac{x^2 + y^2}{2}\} = -x. \tag{143}\]

Clearly, \( \Gamma(\mathcal{N}(\mu, 1)) \) is trivial. Hence, \( \mathcal{N}(\mu, 1)^\mathbb{C} = T\mathcal{N}(\mu, 1)/\{e\} \cong \mathbb{C} \), i.e.,

\[\mathcal{N}(\mu, 1)^\mathbb{C} \cong \mathbb{C}. \]

The canonical projection \( \pi_{\mathcal{N}(\mu, 1)} : \mathbb{C} \to \mathcal{N}(\mu, 1) \) is easily seen to be

\[\pi_{\mathcal{N}(\mu, 1)} : \mathbb{C} \to \mathcal{N}(\mu, 1), \quad \pi_{\mathcal{N}(\mu, 1)}(\xi) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{(x - \xi)^2}{2} \right\}, \tag{145}\]

where \( z = x + iy \in \mathbb{C} \).

For the spectral theory, observe that \( \mathcal{A}_{\mathcal{N}(\mu, 1)} = \text{Vect}_\mathbb{R} \{1, \xi\} \), where \( \xi : \mathbb{R} \to \mathbb{R} \) is the identify, and that the group of holomorphic isometries of \( \mathbb{C} \) is the group \( E(2) = \mathbb{R}^2 \times O(2) \) of Euclidean isometries. As a simple calculation shows, the only Kähler functions \( f \) on \( \mathbb{C} \) that can be written as \( f(z) = \int_{\mathbb{R}} X(\xi) \left[ (\pi \circ \phi)(z) \right] d\xi \), where \( X \in \mathcal{A}_{\mathcal{N}(\mu, 1)} \) and \( \phi \in E(2) \), are functions of the form \( f(z) = u_0 + ux + vy \), where \( u_0, u, v \in \mathbb{R} \). For such function, the subset \( \text{spec}(f) := \text{Im}(X) \) is well defined, i.e. independent of the decomposition of \( f \), and so is its associated probability \( P_{f,z} \).

For \( f(z) = u_0 + ux + vy \), calculations yield \( \text{spec}(f) = \mathbb{R} \) if \( \text{grad}(f) \neq 0 \), \( \text{spec}(f) = \{u_0\} \) if \( \text{grad}(f) = 0 \) and\n
\[P_{f,z} = \left\{ \begin{array}{ll}
\frac{1}{\sqrt{2\pi} \|\text{grad}(f)\|} \cdot \exp \left\{ -\frac{(\xi - f(z))^2}{2\|\text{grad}(f)\|^2} \right\} & \text{if } \text{grad}(f) \neq 0, \\
\frac{1}{\delta_{u_0}} & \text{if } \text{grad}(f) = 0,
\end{array} \right. \tag{146}\]

where \( \text{grad}(f) = (u, v) \) denotes the Riemannian gradient of \( f \) and \( \|\text{grad}(f)\| = \sqrt{u^2 + v^2} \) its Euclidean norm.

Let us now relate the above formulas with the quantum harmonic oscillator. Let \( \hbar \) be a nonnegative real constant and let \( \Psi : \mathbb{C} \to C^\infty(\mathbb{R}, \mathbb{C}) \) be the function defined for \( \xi \in \mathbb{R} \) and \( z = x + iy \in \mathbb{C} \) by

\[\Psi(z)(\xi) := \frac{1}{(2\pi\hbar)^{1/4}} \exp \left\{ -\frac{(\xi - x)^2}{4\hbar} \right\} \exp \left\{ -\frac{i}{\hbar} y \xi \right\}. \tag{147}\]

Let us also define a linear map \( \mathbb{Q} \) from the space \( \mathcal{K}(\mathbb{C}) \) to the space of unbounded operators acting on \( L^2(\mathbb{R}, \mathbb{C}) \) by

\[1 \mapsto \text{Id}, \quad x \mapsto x, \quad y \mapsto i\hbar \frac{\partial}{\partial x}, \quad \frac{x^2 + y^2}{2} \mapsto -\frac{\hbar^2}{2} \frac{\partial^2}{\partial x^2} + \frac{1}{2} x^2 - \left( \frac{\hbar^2}{8} + \frac{1}{2} \right). \tag{148}\]

Observe that \( \mathbb{Q} \) is “essentially” the operator which quantizes the classical harmonic oscillator.

31
Proposition 10.1. For all $f \in \mathcal{H}(\mathbb{C})$ and for all $z \in \mathbb{C}$, we have:

$$f(z) = \langle \Psi(z), Q(f) \cdot \Psi(z) \rangle,$$

(149)

where $(\,,\,)$ is the usual $L^2$-scalar product on $L^2(\mathbb{R}, \mathbb{C})$.

Proof. By direct calculations.

Equation (149) is the exact analogue of (127), but in an infinite dimensional context. Indeed, in [Mola] (work in progress), we regard the space $\mathcal{D} := \{ \rho : M \to \mathbb{R} | \rho \text{ smooth}, \rho > 0 \text{ and } \int_M \rho(x) \cdot d\text{vol}_g = 1 \}$ of smooth density probability functions on a (compact) oriented Riemannian manifold $(M,g)$ as an infinite dimensional (Fréchet) manifold, and we exhibit the analogues of the Fisher metric and exponential connection on $\mathcal{D}$, obtaining via Dombrowski’s construction, an almost Hermitian structure on $T\mathcal{D}$ which allows for an embedding $\mathbb{C} \hookrightarrow T\mathcal{D} \subseteq C^\infty(M,\mathbb{C})$ similar to that of $S^2 \hookrightarrow P(\mathbb{C}^{n+1})$ in [3]. In this approach, the map $\Psi$ is obtained by solving a simple partial differential equation related to a particular description of the tangent bundle of $\mathcal{D}$ (see Mola for details).

Acknowledgements. It is a pleasure to thank Peter Heinzner, Shoshichi Kobayashi, Daniel Sternheimer, Hisung Tze and Tilmann Wurzbacher for many helpful discussions. I would like in particular to express my gratitude to Yoshiaki Maeda for having been such a great “host professor” during my postdoctoral stay in Keio University; I owe him a lot.

This work was done with the financial support of the Japan Society for the Promotion of Science.

References

[AN00] Shun-ichi Amari and Hiroshi Nagaoka. Methods of information geometry, volume 191 of Translations of Mathematical Monographs. American Mathematical Society, Providence, RI, 2000. Translated from the 1993 Japanese original by Daishi Harada.

[AS99] Abhay Ashtekar and Troy A. Schilling. Geometrical formulation of quantum mechanics. In On Einstein’s path (New York, 1996), pages 23–65. Springer, New York, 1999.

[CL84] R. Cirelli and P. Lanzavecchia. Hamiltonian vector fields in quantum mechanics. Nuovo Cimento B (11), 79(2):271–283, 1984.

[CM74] Paul R. Chernoff and Jerrold E. Marsden. Properties of infinite dimensional Hamiltonian systems. Lecture Notes in Mathematics, Vol. 425. Springer-Verlag, Berlin, 1974.

[CMP90] Renzo Cirelli, Alessandro Manià, and Livio Pizzocchero. Quantum mechanics as an infinite-dimensional Hamiltonian system with uncertainty structure. I, II. J. Math. Phys., 31(12):2891–2897, 2898–2903, 1990.

[Dom62] Peter Dombrowski. On the geometry of the tangent bundle. J. Reine Angew. Math., 210:73–88, 1962.

[Gib92] G. W. Gibbons. Typical states and density matrices. J. Geom. Phys., 8(1-4):147–162, 1992.

[Hes84] André Heslot. Une caractérisation des espaces projectifs complexes. C. R. Acad. Sci. Paris Sér. I Math., 298(5):95–97, 1984.

[Hes85] André Heslot. Quantum mechanics as a classical theory. Phys. Rev. D (3), 31(6):1341–1348, 1985.
[Hug95] L. P. Hughston. Geometric aspects of quantum mechanics. In *Twistor theory (Plymouth)*, volume 169 of *Lecture Notes in Pure and Appl. Math.*, pages 59–79. Dekker, New York, 1995.

[Jos02] J. Jost. *Riemannian geometry and geometric analysis*. Universitext. Springer-Verlag, Berlin, third edition, 2002.

[Kaw] K. Kawamura. *Quantum mechanics and operator algebras on the Hilbert ball (The revised)*. arXiv:funct-an/9710002v3.

[Kib79] T. W. B. Kibble. Geometrization of quantum mechanics. *Comm. Math. Phys.*, 65(2):189–201, 1979.

[Kos70] Bertram Kostant. Quantization and unitary representations. I. Prequantization. In *Lectures in modern analysis and applications, III*, pages 87–208. Lecture Notes in Math., Vol. 170. Springer, Berlin, 1970.

[Lan02] S. Lang. *Introduction to differentiable manifolds*. Universitext. Springer-Verlag, New York, second edition, 2002.

[Mar02] Louis Marchildon. *Quantum mechanics*. Advanced Texts in Physics. Springer-Verlag, Berlin, 2002. From basic principles to numerical methods and applications, Translated from the 2000 French original by the author.

[Mola] Mathieu Molitor. Information geometry and the hydrodynamical formulation of quantum mechanics. *In preparation*.

[Molb] Mathieu Molitor. Remarks on the statistical origin of the geometrical formulation of quantum mechanics. To appear in *International Journal of Geometric Methods in Modern Physics*, Vol. 9, No. 3 (DOI: 10.1142/S0219887812200010).

[MR93] Michael K. Murray and John W. Rice. *Differential geometry and statistics*, volume 48 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.

[MT03] Djordje Minic and Chia-Hsiung Tze. Background independent quantum mechanics and gravity. *Phys. Rev. D (3)*, 68(6):061501, 5, 2003.

[MT04] Djordje Minic and Chia-Hsiung Tze. A general theory of quantum relativity. *Phys. Lett. B*, 581(1-2):111–118, 2004.

[Nom60] Katsumi Nomizu. On local and global existence of Killing vector fields. *Ann. of Math.* (2), 72:105–120, 1960.

[NS92] Katsumi Nomizu and Udo Simon. Notes on conjugate connections. In *Geometry and topology of submanifolds, IV (Leuven, 1991)*, pages 152–173. World Sci. Publ., River Edge, NJ, 1992.

[Sou97] J.-M. Souriau. *Structure of dynamical systems*, volume 149 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1997. A symplectic view of physics, Translated from the French by C. H. Cushman-de Vries, Translation edited and with a preface by R. H. Cushman and G. M. Tuynman.