



Information geometry, dynamics and discrete quantum mechanics

Author

Reginatto, Marcel, Hall, Michael JW

Published

2013

Conference Title

BAYESIAN INFERENCE AND MAXIMUM ENTROPY METHODS IN SCIENCE AND ENGINEERING

DOI

[10.1063/1.4820006](https://doi.org/10.1063/1.4820006)

Downloaded from

<http://hdl.handle.net/10072/56648>

Griffith Research Online

<https://research-repository.griffith.edu.au>



Information geometry, dynamics and discrete quantum mechanics

Marcel Reginatto and Michael J. W. Hall

Citation: [AIP Conference Proceedings](#) **1553**, 246 (2013); doi: 10.1063/1.4820006

View online: <http://dx.doi.org/10.1063/1.4820006>

View Table of Contents: <http://scitation.aip.org/content/aip/proceeding/aipcp/1553?ver=pdfcov>

Published by the [AIP Publishing](#)

Information geometry, dynamics and discrete quantum mechanics

Marcel Reginkatto* and Michael J. W. Hall†

*Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany

†Centre for Quantum Dynamics, Griffith University, Brisbane, QLD 4111, Australia

Abstract. We consider a system with a discrete configuration space. We show that the geometrical structures associated with such a system provide the tools necessary for a reconstruction of discrete quantum mechanics once dynamics is brought into the picture. We do this in three steps. Our starting point is *information geometry*, the natural geometry of the space of probability distributions. Dynamics requires additional structure. To evolve the P^k , we introduce coordinates S^k canonically conjugate to the P^k and a *symplectic structure*. We then seek to extend the metric structure of information geometry, to define a geometry over the full space of the P^k and S^k . Consistency between the metric tensor and the symplectic form forces us to introduce a *Kähler geometry*. The construction has notable features. A complex structure is obtained in a natural way. The canonical coordinates $\psi^k = \sqrt{P^k} e^{iS^k}$ of the Kähler space are precisely the wave functions of quantum mechanics. The full group of unitary transformations is obtained. Finally, one may associate a Hilbert space with the Kähler space, which leads to the standard version of quantum theory. We also show that the metric that we derive here using purely geometrical arguments is precisely the one that leads to Wootters' expression for the statistical distance for quantum systems.

Keywords: quantum mechanics, information geometry, symplectic geometry, Kähler geometry

PACS: 03.65.Ta, 02.40.Tt, 02.40.Yy

INTRODUCTION

In this paper, we consider systems with a discrete configuration space. In the presence of uncertainty, the state of a classical system will be described by a probability $P = (P^1, \dots, P^n)$, where n is the number of available states. An analysis complementing the one carried out in Ref. [1], for continuous systems, leads to the remarkable result that the geometrical structures associated with such a system provide the tools necessary for a reconstruction of discrete quantum mechanics once dynamics is brought into the picture. The reformulation of Ref. [1] presented here is nontrivial: it requires new geometrical insights because assumptions that are natural in the infinite dimensional case (e.g., “spatial locality,” the action of the Galilean group, etc.) are no longer available in the discrete case.

This reconstruction of discrete quantum mechanics has notable features. We show that the natural geometry of the space of probabilities in motion (i.e., taking dynamics into consideration) is a Kähler geometry. The canonical coordinates $\psi^k = \sqrt{P^k} e^{iS^k}$ of the Kähler space are precisely the wave functions of quantum mechanics. The full group of unitary transformations is obtained. Finally, one may associate a Hilbert space with the Kähler space, which leads to the standard version of quantum theory. We note that the Kähler space metric that we derive here using purely geometrical arguments is precisely

the one that leads to Wootters' expression for the statistical distance for quantum systems [2], which he derived using a completely different argument based on the concept of distinguishability. Comparison of our approach to papers of Mehrafarin [3] and Goyal [4, 5], which also start from classical information geometry, will be made elsewhere.

INFORMATION GEOMETRY

The starting point is a classical system which has associated with it n different states. The probability that the system is in state i is given by P^i , $i = 1, \dots, n$, with $P^i \geq 0$ and $\sum_i P^i = 1$. There is a natural line element given by

$$ds^2 = G_{ij} dP^i dP^j = \frac{\alpha}{2P^i} \delta_{ij} dP^i dP^j \quad (1)$$

where α is a constant. The value of this constant can not be determined *a priori*; it is usually set to $\frac{1}{2}$. We do not make this assumption here but instead allow α to be a free parameter. The metric G_{ij} is known as the *information metric*,

$$G_{ij} = \frac{\alpha}{2P^i} \delta_{ij}. \quad (2)$$

The line element of Eq. (1) leads to a concept of distance on a probability space. This distance seems to have been introduced into statistics by Bhattacharyya [6, 7] as a way of providing a measure of divergence for multinomial probabilities [8]. Wootters calls it the *statistical distance* [2]. To derive the statistical distance from the metric G_{ij} , consider two points in probability space, P_A and P_B , joined by a curve $P^i(t)$, $0 \leq t \leq 1$, and write the expression for the length l of the curve in the form

$$l = \int_0^1 dt \sqrt{G_{ij} \frac{dP^i(t)}{dt} \frac{dP^j(t)}{dt}}. \quad (3)$$

The statistical distance is defined as the shortest distance between P_A and P_B . To compute the statistical distance, it is convenient to do the change of coordinates $X^i = \sqrt{P^i}$. Then

$$l = \sqrt{2\alpha} \int_0^1 dt \sqrt{\sum_{i=1}^n \left[\frac{dX^i(t)}{dt} \right]^2}. \quad (4)$$

Since the curve $P(t)$ is assumed to lie in the probability space, it must satisfy the condition $\sum_{i=1}^n P^i(t) = \sum_{i=1}^n [X^i(t)]^2 = 1$; that is, the curve must lie on a unit n -dimensional sphere in the X space. The shortest distance on the n -dimensional sphere is equal to the angle between the unit vectors X_A and X_B . This leads immediately to

$$d(P_A, P_B) = \sqrt{2\alpha} \cos^{-1} \left(\sum_{i=1}^n X_A^i X_B^i \right) = \sqrt{2\alpha} \cos^{-1} \left(\sum_{i=1}^n \sqrt{P_A^i} \sqrt{P_B^i} \right). \quad (5)$$

which agrees with the expressions in the papers of Bhattacharyya and Wootters provided α is set to the standard value of $\frac{1}{2}$.

The statistical distance does not play a fundamental role in our discussion. Nevertheless, we make reference to it here because it clarifies the relation of our work to that of Wootters.

DYNAMICS, SYMPLECTIC GEOMETRY, AND OBSERVABLES

We now set the probabilities in motion. We assume that the dynamics of P^i are generated by an action principle and we introduce additional coordinates S^i which are canonically conjugate to the P^i and a corresponding Poisson bracket for any two functions $F(P, S)$ and $G(P, S)$,

$$\{F, G\} = \sum_i \left(\frac{\partial F}{\partial P^i} \frac{\partial G}{\partial S^i} - \frac{\partial F}{\partial S^i} \frac{\partial G}{\partial P^i} \right). \quad (6)$$

As is well known, the Poisson bracket can be rewritten geometrically as

$$\{F, G\} = (\partial F / \partial P, \partial F / \partial S) \Omega \begin{pmatrix} \partial G / \partial P \\ \partial G / \partial S \end{pmatrix}, \quad (7)$$

where Ω is the corresponding symplectic form, given in this case by

$$\Omega = \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix}, \quad (8)$$

where $\mathbf{1}$ is the unit matrix in n dimensions. We thus have a symplectic structure and a corresponding *symplectic geometry* (which is why this formulation of the dynamics is a natural one for a geometric approach). The equations of motion for P^i and S^i are given by $\dot{P}^i = \{P^i, H\}$, $\dot{S}^i = \{S^i, H\}$ where H is the Hamiltonian that generates time translations.

The observables of the theory are functions $A(P, S)$ of the coordinates P^i and S^i . Certain restrictions are imposed on them, so not every function is an observable. For example, the infinitesimal canonical transformation generated by any observable A must preserve the normalization and positivity of P . This implies the two conditions [9]

$$A(P, S + c) = A(P, S), \quad \partial A / \partial S^i = 0 \text{ if } P^i = 0. \quad (9)$$

Note that the first condition implies gauge invariance of the theory under $S^i \rightarrow S^i + \chi$, where χ is a constant [9].

KÄHLER GEOMETRY

We now want to consider the following question: Can we extend the metric G_{ij} in Eq. (2), which is only defined on the n -dimensional subspace of probabilities P^i , to the full $2n$ -dimensional phase space of the P^i and S^i ? It can be done, but certain conditions which ensure the compatibility of the metric and symplectic structures have to be satisfied. These conditions are equivalent to requiring that the space have a Kähler structure (see the Appendix of Ref. [1] for a proof). We are led then to the beautiful result that the natural geometry of the space of probabilities in motion is a *Kähler geometry*.

A Kähler structure brings together metric, symplectic and complex structures in a harmonious way. To define such a space, introduce a complex structure J_b^a and impose the following conditions [10],

$$\Omega_{ab} = g_{ac}J_b^c, \quad (10)$$

$$J_c^a g_{ab} J_d^b = g_{cd}, \quad (11)$$

$$J_b^a J_c^b = -\delta_c^a. \quad (12)$$

Eq. (10) is a compatibility equation between the symplectic structure Ω_{ab} and the metric g_{ab} , Eq. (11) is the condition that the metric should be Hermitian, and Eq. (12) is the condition that J_b^a should be a complex structure.

We derive the solutions to these equations. The metric over the subspace of probabilities is the information metric, Eq. (2). The metric over the full space will take the form

$$g_{ab} = \begin{pmatrix} \mathbf{G} & \mathbf{E} \\ \mathbf{E}^T & \mathbf{F} \end{pmatrix}, \quad (13)$$

where $\mathbf{G} = \text{diag}(\frac{\alpha}{2P^i})$, and \mathbf{E} and \mathbf{F} are $n \times n$ matrices that need to be determined using the Kähler conditions and the expression for Ω_{ab} , Eq. (8). A matrix calculation leads to general forms for the metric g_{ab} and the complex structure J_b^a . These are

$$g_{ab} = \begin{pmatrix} \mathbf{G} & \mathbf{A}^T \\ \mathbf{A} & (\mathbf{1} + \mathbf{A}^2)\mathbf{G}^{-1} \end{pmatrix}, \quad J_b^a = \begin{pmatrix} \mathbf{A} & (\mathbf{1} + \mathbf{A}^2)\mathbf{G}^{-1} \\ -\mathbf{G} & -\mathbf{G}\mathbf{A}\mathbf{G}^{-1} \end{pmatrix}. \quad (14)$$

where the $n \times n$ matrix \mathbf{A} satisfies $\mathbf{G}\mathbf{A}\mathbf{G}^{-1} = \mathbf{A}^T$ but is otherwise arbitrary.

The Kähler conditions restrict the form of the metric g_{ab} but leave the matrix \mathbf{A} (and therefore the geometry) undetermined. To fix the geometry, it is necessary to introduce an additional condition. As we show in the Appendix, there is a natural requirement based on the idea of invariance of the metric under the motions generated by the observables of the theory. This condition leads to $\mathbf{A} = 0$ and, as shown below, to a very simple geometry. The condition $\mathbf{A} = 0$ alternatively follows by assuming the extended metric has zero curvature, analogously to the classical information metric G_{ij} .

COMPLEX COORDINATES

We set $\mathbf{A} = 0$ and consider the Kähler structure given by

$$\Omega_{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad g_{ab} = \begin{pmatrix} \mathbf{G} & 0 \\ 0 & \mathbf{G}^{-1} \end{pmatrix}, \quad J_b^a = \begin{pmatrix} 0 & \mathbf{G}^{-1} \\ -\mathbf{G} & 0 \end{pmatrix}. \quad (15)$$

We now carry out the Madelung transformation, $\psi^i = \sqrt{P^i} \exp(iS^i/\alpha)$, $\bar{\psi}^i = \sqrt{P^i} \exp(-iS^i/\alpha)$. In terms of these complex coordinates, the tensors that define the Kähler geometry, Eqs. (15), take the standard form which is characteristic of a flat-space [10],

$$\Omega_{ab} = \begin{pmatrix} 0 & i\alpha\mathbf{1} \\ -i\alpha\mathbf{1} & 0 \end{pmatrix}, \quad g_{ab} = \begin{pmatrix} 0 & \alpha\mathbf{1} \\ \alpha\mathbf{1} & 0 \end{pmatrix}, \quad J_b^a = \begin{pmatrix} -i\mathbf{1} & 0 \\ 0 & i\mathbf{1} \end{pmatrix}. \quad (16)$$

This shows that the simplest geometrical formulation of the space of probabilities in motion has a natural set of fundamental variables, ψ^i and $\bar{\psi}^i$. If we set the constant α equal to \hbar , these fundamental variables are precisely the *wave functions* of quantum mechanics.

This is a remarkable result because we have not introduced *any* assumptions that concern quantum mechanics, only geometrical arguments.

STATISTICAL DISTANCE IN THE KÄHLER SPACE

We now want to introduce a new expression for statistical distance which will be valid in the Kähler space. In going from the n -dimensional space of probabilities P^i to the full $2n$ -dimensional phase space of the P^i and S^i , the metric has been extended, and this should be taken in consideration when carrying out the generalization of the statistical distance.

Consider two points ψ_A and ψ_B representing states in this space which are joined by a curve $\psi^i(t)$, $0 \leq t \leq 1$. Combining Eq. (3) with Eq. (16), the correct generalization of the expression for the distance l will be given by

$$l = \int_0^1 dt \sqrt{g_{ij} \frac{d\psi^i(t)}{dt} \frac{d\bar{\psi}^j(t)}{dt}} = \sqrt{2\alpha} \int_0^1 dt \left| \left(\frac{d\psi(t)}{dt}, \frac{d\psi(t)}{dt} \right) \right|. \quad (17)$$

where we have introduced the notation $|(\psi, \psi)| = \sqrt{\sum_i \psi^i \bar{\psi}^i}$.

The statistical distance in the Kähler space is defined as the shortest distance computed with Eq. (17). Since the curve $P(t)$ is assumed to lie in the probability space, then $\psi(t)$ must satisfy the condition

$$1 = \sum_i \psi^i(t) \bar{\psi}^i(t), \quad (18)$$

that is, the curve must lie on the unit sphere in the $\{\psi(t), \bar{\psi}(t)\}$ space. The shortest distance on the unit sphere is equal to the angle between the unit vectors ψ_A and ψ_B . This leads immediately to the expression for the statistical distance that appears at the end of section III of Wootters' paper (no equation number) provided α is set to the standard value of $\frac{1}{2}$,

$$d(\psi_A, \psi_B) = \sqrt{2\alpha} \cos^{-1} |(\psi_A, \psi_B)|. \quad (19)$$

In this way we provide a very brief, geometrical derivation of Wootters' expression for the statistical distance in quantum mechanics, which he derived using a completely different argument based on distinguishability [2].

GROUP OF LINEAR UNITARY TRANSFORMATIONS

We now examine the transformations which are allowed by the theory. As pointed out briefly in the discussion on observables, there are some basic conditions that must be

satisfied; we list them and derive the group of transformations that is consistent with such conditions.

The first requirement is that the transformations preserve the normalization of the probability, $\sum_i P^i = \sum_i \psi^i \bar{\psi}^i = 1$.

The second requirement is that the metric be form invariant under those transformations; i.e., that the line element $d\sigma^2 = 2\alpha \sum_j d\bar{\psi}^j d\psi^j$ of the Kähler space is preserved by the transformations.

Requiring normalization of the probability and metric invariance leads to the group of rotations on the $2n$ -dimensional sphere. Such rotations are linear with respect to ψ^j and $\bar{\psi}^j$. For an infinitesimal transformation, it follows that

$$\dot{\psi}^j = -i \frac{\partial H}{\partial \bar{\psi}^j}, \quad \dot{\bar{\psi}}^j = i \frac{\partial H}{\partial \psi^j}, \quad (20)$$

are linear in ψ and $\bar{\psi}$, where H is the Hamiltonian that generates the motion. Then H must be of the form

$$H = E(t) + \sum_{j,k} \left[M_{jk} \bar{\psi}^j \psi^k + N_{jk} \psi^j \psi^k + \bar{N}_{jk} \bar{\psi}^j \bar{\psi}^k \right] \quad (21)$$

where $E(t)$ is an arbitrary function of time, M is Hermitian, and N is symmetric.

The third and final requirement is that we only consider rotations on the $2n$ -dimensional sphere that are compatible with the equations of motion. As pointed out in the discussion on observables, conservation of probability requires that the ensemble Hamiltonian be invariant (up to an additive constant) under $S^j \rightarrow S^j + \chi$, since to first-order [9]

$$0 = \varepsilon \sum_j \dot{P}^j = \varepsilon \sum_j \frac{\partial H}{\partial S^j} = H(P, S + \varepsilon) - H(P, S). \quad (22)$$

This condition, when written in terms of complex coordinates, is equivalent to invariance of the Hamiltonian under $\psi \rightarrow \psi e^{i\chi}$. Using the notation $Q := \sum N_{jk} \psi^j \psi^k$, Eq. (21) leads to the equality

$$[Q e^{2i\chi} + \bar{Q} e^{-2i\chi}] = 0 \quad (23)$$

which must be valid for all χ . Differentiating with respect to χ gives the additional equality

$$2i [Q e^{2i\chi} - \bar{Q} e^{-2i\chi}] = 0. \quad (24)$$

Combining these two expressions leads to $Q = 0$. Since this must hold for all ψ , it follows that $N_{jk} \equiv 0$, i.e., the ensemble Hamiltonian has the Hermitian form

$$H = E(t) + \sum_{j,k} M_{jk} \bar{\psi}^j \psi^k \quad (25)$$

as desired.

This shows that the group of transformations of the theory is precisely the group of *linear unitary transformations*. Note that all that is used here, in moving from all rotations on the $2n$ -dimensional sphere to the subset of unitary transformations, is (i) the conservation of probability and (ii) that the equations of motion follow from an action principle.

HILBERT SPACE FORMULATION

There is a standard construction that associates a complex Hilbert space with any Kähler space. Given two complex vectors ϕ^i and φ^i , define the Dirac product by [11]

$$\begin{aligned}
 \langle \phi | \varphi \rangle &= \frac{1}{2} \sum_i \left\{ (\phi^i, \bar{\phi}^i) \cdot [g + i\Omega] \cdot \begin{pmatrix} \varphi^i \\ \bar{\varphi}^i \end{pmatrix} \right\} \\
 &= \frac{1}{2} \sum_i \left\{ (\phi^i, \bar{\phi}^i) \left[\begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix} + i \begin{pmatrix} 0 & i\mathbf{1} \\ -i\mathbf{1} & 0 \end{pmatrix} \right] \begin{pmatrix} \varphi^i \\ \bar{\varphi}^i \end{pmatrix} \right\} \\
 &= \sum_i \bar{\phi}^i \varphi^i
 \end{aligned} \tag{26}$$

This suggests that the Hilbert space structure of quantum mechanics is perhaps not as fundamental as its geometrical structure.

CONCLUDING REMARKS

We have shown that the Hilbert space formulation of discrete quantum theory emerges from the geometry of probabilities in motion. The basic elements that go into this geometrical reconstruction of discrete quantum mechanics are the natural metric on the space of probabilities (information geometry), the description of dynamics using a Hamiltonian formalism (symplectic geometry), and requirements of consistency (Kähler geometry).

We summarize some of the remarkable features of this construction. The wave functions of quantum mechanics, $\psi^k = \sqrt{P^k} e^{iS^k}$, appear as the natural complex coordinates of the Kähler space that describes the geometry of probabilities in motion. The full group of unitary transformations is derived based on consistency requirements. And, finally, a Hilbert space may be associated with the Kähler space of the theory, which leads to the standard version of quantum theory.

We have shown that we can derive Wootters' statistical distance for quantum mechanics, Eq. (17), in a purely geometrical way. We have commented in this brief paper on the connection between our work and that of Wootters; it would also be interesting to investigate the connection of our work to papers of Mehrafarin [3] and Goyal [4, 5] which also take an information-geometrical approach. One of the main differences seems to come from the emphasis that we place on finding a geometrical description for the space of probabilities *in motion*. The use of an action principle to describe the dynamics of the probabilities P^i introduces geometrical structure that is quite powerful: we immediately get a doubling of the dimensionality of the space (i.e., $\{P^i\} \rightarrow \{P^i, S^i\}$) and, in addition, we end up with a complex structure and unitary transformations, all of them ingredients that are essential for quantum mechanics. In our formalism, these are mainly the result of consistency requirements which we impose to ensure the peaceful coexistence of information geometry and symplectic geometry.

It appears then that discrete quantum mechanics stands at the intersection of information geometry, symplectic geometry, and Kähler geometry.

APPENDIX: THE CONDITION $\mathbf{A} = 0$

To fix the matrix \mathbf{A} that appears in Eqs. (14) we need to supplement the Kähler conditions, Eqs. (10-12), with an additional condition. We make use of the fact that motion in the probability space (for example, the trajectory in phase space generated by an observable) preserves the normalization $\sum_i P^i = 1$. To look more closely at this constraint, we introduce the new set of $2n$ real coordinates

$$x^i = \sqrt{2\alpha P^i} \cos(S^i/\alpha), \quad y^i = \sqrt{2\alpha P^i} \sin(S^i/\alpha). \quad (27)$$

One can check that this coordinate transformation is also a canonical transformation, $\{x^i, y^j\} = \delta^{ij}$, so the symplectic form remains invariant.

The constraint $\sum_i P^i = 1$ takes the form $\sum_i \{(x^i)^2 + (y^i)^2\} = 2\alpha$ in the new coordinates. Therefore, motion is restricted to a $2n$ -dimensional sphere in the space of the x^i and y^i . We require that the metric be invariant under rotations (equivalently, invariant under the motions generated by observables); i.e., that it be spherically symmetric. In the coordinates x^i and y^i , a spherically symmetric line element has the general form

$$d\rho^2 = f(r) \sum_i \{(dx^i)^2 + (dy^i)^2\} + g(r) \left\{ \sum_i (x^i dx^i + y^i dy^i) \right\}^2 \quad (28)$$

where f and g are arbitrary functions of $r = \sqrt{\sum_i \{(x^i)^2 + (y^i)^2\}}$. In terms of the original coordinates P^i and S^i , the spherically symmetric metric takes the form

$$d\rho^2 = f(\sum_i P^i) \sum_i \left\{ \frac{\alpha}{2P^i} (dP^i)^2 + \frac{2P^i}{\alpha} (dS^i)^2 \right\} + g(\sum_i P^i) \left\{ \alpha \sum_i dP^i \right\}^2 \quad (29)$$

The important point here is that the metric $d\rho^2$ does *not* have any mixed terms proportional to $dP^i dS^j$, and this means that we must set $\mathbf{A} = 0$ to satisfy our requirement of metric invariance.

REFERENCES

1. M. Reginatto and M. J. W. Hall, *AIP Conf. Proc.* **1443**, 96-103 (2011).
2. W. K. Wootters, *Phys. Rev. D* **23**, 357-362 (1981).
3. M. Mehrafarin, *Int. J. Theor. Phys.* **44**, 429-442 (2005).
4. P. Goyal, *Phys. Rev. A* **78**, 052120 (2008).
5. P. Goyal, *New Journal of Physics* **12**, 023012 (2010).
6. A. Bhattacharyya, *Bull. Calcutta Math. Soc.* **35**, 99-109 (1943).
7. A. Bhattacharyya, *Sanhkyā* **7**, 401-406 (1946).
8. I. J. Good, *J. Stat. Comput. Simulation* **36**, 179-186 (1990).
9. M. J. W. Hall, *J. Phys. A* **37**, 7799-7811 (2004).
10. S.I. Goldberg, *Curvature and Homology*, Dover Publications, New York, 1982.
11. T. W. B. Kibble, *Commun. math. Phys.* **65**, 189-201 (1979).