Quantum theory from the geometry of evolving probabilities

Marcel Reginatto

Physikalisch-Technische Bundesanstalt,

Bundesallee 100, 38116 Braunschweig, Germany

Michael J. W. Hall

Theoretical Physics, RSPE, Australian National University,

Canberra ACT 0200, Australia and

Centre for Quantum Dynamics,

Griffith University, Brisbane,

QLD 4111, Australia

Abstract

We consider the space of probabilities $\{P(x)\}$, where the x are coordinates of a configuration space. Under the action of the translation group, $P(x) \to P(x+\theta)$, there is a natural metric over the parameters θ given by the Fisher-Rao metric. This metric induces a metric over the space of probabilities. Our next step is to set the probabilities in motion. To do this, we introduce a canonically conjugate field S and a symplectic structure; this gives us Hamiltonian equations of motion. We show that it is possible to extend the metric structure to the *full space* of the (P, S), and this leads in a natural way to introducing a Kähler structure; i.e., a geometry that includes compatible symplectic, metric and complex structures.

The simplest geometry that describes these spaces of evolving probabilities has remarkable properties: the natural, canonical variables are precisely the wave functions of quantum mechanics; the Hamiltonian for the quantum free particle can be derived from a representation of the Galilean group using purely geometrical arguments; and it is straightforward to associate with this geometry a Hilbert space which turns out to be the Hilbert space of quantum mechanics. We are led in this way to a reconstruction of quantum theory based solely on the geometry of probabilities in motion.

I. PROBABILITIES, TRANSLATIONS, AND INFORMATION GEOMETRY

Our starting point is an n-dimensional configuration space, with coordinates $x \equiv \{x^1, \ldots, x^n\}$ and probability densities $P(x)$ satisfying $P(x) \geq 0$ and $\int d^n x P(x) = 1$.

If we consider the action of the translation group T on $P(x)$, $T : P(x) \to P(x+\theta)$, there is a natural metric γ_{jk} on the space of parameters θ: the Fisher-Rao metric [1],

$$\gamma_{jk} = \frac{\alpha}{2} \int d^n x \frac{1}{P(x+\theta)} \frac{\partial P(x+\theta)}{\partial \theta^j} \frac{\partial P(x+\theta)}{\partial \theta^k}, \tag{1}$$

where α is a constant. The line element $d\sigma^2 = \gamma_{jk} \Delta^j \Delta^k$ (where $|\Delta^k| << 1$) defines a distance between the two probability distributions $P(x+\theta)$ and $P(x+\theta+\Delta)$.

It will be convenient to consider another form of the metric. In particular, using the equality $\frac{\partial P(x+\theta)}{\partial \theta^j} = \frac{\partial P(x+\theta)}{\partial x^j}$, and making the change of integration variables $x \to x + \theta$, the metric is proportional to the Fisher information matrix,

$$\gamma_{jk} = \frac{\alpha}{2} \int d^n x \frac{1}{P(x)} \frac{\partial P(x)}{\partial x^j} \frac{\partial P(x)}{\partial x^k}. \tag{2}$$

It follows that the line element $d\sigma^2 = \gamma_{jk} \Delta^j \Delta^k$ induces a line element ds^2 in the space of probability densities. Introducing the notation $P_x = P(x)$, $\delta P_x \equiv \frac{\partial P(x)}{\partial x^j} \Delta^j$, we have

$$ds^2 = \frac{\alpha}{2} \int d^n x \frac{1}{P_x} \delta P_x \, \delta P_x = \int d^n x \, d^n x' \, g_{PP}(x, x') \, \delta P_x \, \delta P_{x'} \,, \tag{3}$$

The equivalence of the two line elements can be checked by direct substitution. Eq. (3) was introduced by Jeffreys [2]. The induced metric g_{PP} is diagonal, and given by

$$g_{PP}(x, x') = \frac{\alpha}{2 P_x} \delta(x - x'). \tag{4}$$

II. SYMPLECTIC GEOMETRY AND OBSERVABLES

We now set the probabilities in motion. To do this, we assume the dynamics of $P(x)$ are generated by an action principle. Hence, we introduce an auxiliary field S which is canonically conjugate to P, and a corresponding Poisson bracket for any two functionals $F[P, S]$ and $G[P, S]$:

$$\{F, G\} = \int d^n x \left\{ \frac{\delta F}{\delta P} \frac{\delta G}{\delta S} - \frac{\delta F}{\delta S} \frac{\delta G}{\delta P} \right\}. \tag{5}$$

The equations of motion for P and S then have the form

$$\dot{P} = \{P, \mathcal{H}\} = \frac{\delta \mathcal{H}}{\delta S}, \qquad \dot{S} = \{S, \mathcal{H}\} = -\frac{\delta \mathcal{H}}{\delta P}, \tag{6}$$

where $\mathcal{H}$ is the ensemble Hamiltonian that generates time translations.

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

$$\{F, G\} = \int d^n x \, d^n x' \, (\delta F/\delta P_x, \ \delta F/\delta S_x) \, \Omega(x, x') \begin{pmatrix} \delta G/\delta P_{x'} \\ \delta G/\delta S_{x'} \end{pmatrix}, \tag{7}$$

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

$$\Omega(x, x') = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \delta(x - x') \ . \tag{8}$$

We thus have a symplectic structure and a corresponding *symplectic geometry* [3].

The fundamental variables of our phase space are the probabilities P_x and the auxiliary function S_x. We now introduce the notion of an *observable* on this phase space, as any functional $A[P, S]$ that satisfies certain requirements. For example, the infinitesimal canonical transformation generated by any observable A must preserve the normalization and positivity of P. This implies the two conditions [4]

$$A[P, S + c] = A[P, S], \qquad \delta A/\delta S = 0 \text{ if } P(x) = 0. \tag{9}$$

Note that the first condition implies gauge invariance of the theory under $S \to S + c$. A more general condition that might be imposed on observables, which leads to a natural statistical interpretation, is that they are homogeneous of degree one with respect to P,

$$A[\lambda P, S] = \lambda A[P, S]. \tag{10}$$

If we now differentiate both sides of Eq. (10) with respect to λ and set $\lambda = 1$, we get [5]

$$A[P, S] = \int d^n x \, P(\delta A/\delta P) := \langle \delta A/\delta P \rangle, \tag{11}$$

i.e., A can be calculated by integrating over a local density. The main motivation for introducing the homogeneity condition is that it consistently allows observables to be interpreted both as generators of canonical transformations and as expectation values.

It is possible to give a physical interpretation to the canonically conjugate variable S. Notice that $\int dx \, P\nabla S$ is the canonical infinitesimal generator of translations, since

$$\delta P(x) = \delta \mathbf{x} \cdot \left\{ P, \int dx \, P\nabla S \right\} = -\delta \mathbf{x} \cdot \nabla P, \tag{12}$$

$$\delta S(x) = \delta \mathbf{x} \cdot \left\{ S, \int dx \, P\nabla S \right\} = -\delta \mathbf{x} \cdot \nabla S, \tag{13}$$

under action of the generator, and therefore $P\nabla S$ can be considered a *local momentum density* [6].

III. KÄHLER GEOMETRY

We now want to consider the following question: Can we extend the metric $g_{PP}(x, x')$ in Eq. (4), which is only defined on the subspace of probabilities P, to the full phase space of P and S? It can be done, but certain conditions which ensure the compatibility of the metric and symplectic structures have to be satisfied (see the Appendix for a proof). These conditions amount to requiring that the space have a Kähler structure. We show in this way 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^a_b and impose the following conditions [7],

$$\Omega_{ab} = g_{ac} J^c_b , \tag{14}$$

$$J^a_c g_{ab} J^b_d = g_{cd} , \tag{15}$$

$$J^a_b J^b_c = -\delta^a_c . \tag{16}$$

Eq. (14) is a compatibility equation between Ω_{ab} and g_{ab}, Eq. (15) is the condition that the metric should be Hermitian, and Eq. (16) is the condition that J^a_b should be a complex structure. We will now derive the local solutions to these equations.

We saw before that the metric over the subspace of probabilities is diagonal and given by $g_{PP}(x, x') = \frac{\alpha}{2P_x}\delta(x - x')$. We assume that the full metric g_{ab} is also diagonal; that is, of the form $g_{ab}(x, x') = g_{ab}(x)\delta(x - x')$ (this assumption corresponds to a locality assumption). Then g_{ab} is a real, symmetric matrix of the form

$$g_{ab} = \begin{pmatrix} \frac{\alpha}{2P_x} & g_{PS} \\ g_{SP} & g_{SS} \end{pmatrix} \delta(x - x'). \tag{17}$$

The elements $g_{PS} = g_{SP}$ and g_{SS} still need to be determined.

Since $\Omega_{ab}(x, x')$ in Eq. (8) is also diagonal, Eq. (14) implies that $J^c_b(x, x')$ is diagonal; i.e. $J^c_b(x, x') = J^c_b(x)\delta(x - x')$. Using Eq. (16), one can show that J^c_b depends on two arbitrary

5

functionals (which we write as A_x and C_x in the equation below) and can be written in the form

$$J^c_{\ b} = \begin{pmatrix} A_x & C_x(1 + A_x^2) \\ -\frac{1}{C_x} & -A_x \end{pmatrix} \delta(x - x'). \tag{18}$$

It is not difficult to get expressions for $g_{ab}(x, x')$ and $J^c_{\ b}(x, x')$ from the two remaining equations, Eqs. (14) and (15). These final expressions depend only on the arbitrary functional A_x, with

$$g_{ab} = \begin{pmatrix} \frac{\alpha}{2P_x} & A_x \\ A_x & \frac{2P_x}{\alpha}(1 + A_x^2) \end{pmatrix} \delta(x - x'), \tag{19}$$

$$J^a_{\ b} = \begin{pmatrix} A_x & \frac{2P_x}{\alpha}(1 + A_x^2) \\ -\frac{\alpha}{2P_x} & -A_x \end{pmatrix} \delta(x - x'). \tag{20}$$

IV. COMPLEX COORDINATES

Different choices of A_x in Eqs. (19) and (20) correspond in general to different Kähler geometries. From the mathematical point of view, the simplest one among these is a *flat* Kähler space, and it follows from the simplest choice, $A_x = 0$. To show this, we carry out a complex transformation. We set $A_x = 0$ and consider the Kähler structure given (up to a product with $\delta(x - x')$) by

$$\Omega_{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad g_{ab} = \begin{pmatrix} \frac{\alpha}{2P_x} & 0 \\ 0 & \frac{2P_x}{\alpha} \end{pmatrix}, \qquad J^a_{\ b} = \begin{pmatrix} 0 & \frac{2P_x}{\alpha} \\ -\frac{\alpha}{2P_x} & 0 \end{pmatrix}. \tag{21}$$

The complex coordinate transformation that is required is nothing but the Madelung transformation, $\psi = \sqrt{P} \exp(iS/\alpha)$, $\psi^* = \sqrt{P} \exp(-iS/\alpha)$. In terms of the new variables, Eqs. (21) take the standard flat-space form [7]

$$\Omega_{ab} = \begin{pmatrix} 0 & i\alpha \\ -i\alpha & 0 \end{pmatrix}, \qquad g_{ab} = \begin{pmatrix} 0 & \alpha \\ \alpha & 0 \end{pmatrix}, \qquad J^a_{\ b} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}. \tag{22}$$

This shows that the simplest geometrical formulation of the space of probabilities in motion has a natural set of fundamental variables – where, identifying the constant α with $\hbar$, these fundamental variables are precisely the *wave functions* that we encounter in quantum mechanics. This is a remarkable result, because we have not introduced *any* assumptions that concern quantum mechanics. This result was derived using only geometrical assumptions.

V. THE FREE PARTICLE

We now turn our attention to the description of a free particle using the formalism that we have developed. We consider the case where the configuration space is the Euclidean space R^3. This space has a natural metric h_{ab} given by $h_{ab} = \delta_{ab}$ in Cartesian coordinates.

To describe a free particle, we look for a realization of the Galilean group in terms of the algebra of Poisson brackets defined by Eq. (5). The Galilean group has 10 generators:

   A_i : space displacements,

   H : time displacements,

   L_i : space rotations,

   G_i : Galilean transformations ("boosts"),

where $i = 1, 2, 3$. The generator H transforms as a scalar, while A_i, L_i, and G_i transform as vectors. These generators have to satisfy the Poisson bracket relations [8]

$$\{H, A_i\} = 0, \tag{23}$$

$$\{H, L_i\} = 0, \tag{24}$$

$$\{H, G_i\} = -A_i, \tag{25}$$

$$\{L_i, A_j\} = \epsilon_{ijk} A_k, \tag{26}$$

$$\{L_i, L_j\} = \epsilon_{ijk} L_k, \tag{27}$$

$$\{L_i, G_j\} = \epsilon_{ijk} G_k, \tag{28}$$

$$\{A_i, A_j\} = 0, \tag{29}$$

$$\{A_i, G_j\} = -m\delta_{ij}, \tag{30}$$

$$\{G_i, G_j\} = 0, \tag{31}$$

where m is the mass of the particle.

We represent the generators by observables. For space displacements and rotations one finds, by considering the corresponding infinitesimal transformations of P and S, that

$$A_i = \int d^3x \, P \, (\partial_i S), \qquad L_i = \int d^3x \, P \, (\epsilon_{ijk} \, x_j \, \partial_k S), \tag{32}$$

up to additive constants (cf. Eqs. (12) and (13)). Further, for the Galilean boost transformations it is natural to choose the observables

$$G_i = \int d^3x \, P \, (mx_i - t\partial_i S), \tag{33}$$

7

where t is the time. This follows from the standard definition $G_i = (mQ_i - tA_i)$ and the choice $Q_i = \int d^3x \, P \, x_i$ for the position observable. One can check that Eqs. (26-31) are satisfied when we make these choices. Note that the generators satisfy the homogeneity condition, Eq. (11), and have clear interpretations as expectation values.

The one remaining step is to find an observable that satisfies Eqs. (23-25). The first two equations, Eqs. (23-24), will be satisfied by any scalar H, since a scalar is invariant under translations and rotations. Using $G_k = (mQ_k - tA_k)$ and $\{H, A_k\} = 0$, Eq. (25) then simplifies to

$$\{H, G_i\} = -m \int d^3x \frac{\delta H}{\delta S} x_i = -A_i = -\int d^3x \, P \, \partial_i S. \tag{34}$$

Eq. (34) does not have a unique solution, and different choices of H will correspond to different theories of the free particle.

It is remarkable that the *only* scalar that can be constructed using purely geometrical quantities *does* satisfy Eq. (34). Therefore, from the point of view of the geometry of probabilities in motion, it is the natural choice for the generator of time translations. It is straightforward to calculate this scalar. Using an argument similar to the one used in the first section of this paper, we can show that the metric over the fields P and S given in Eq. (21) induces a corresponding metric over the space of parameters,

$$g_{jk} = \frac{2}{\alpha} \int d^3x \, P \, \left(\frac{\partial S}{\partial x^j} \frac{\partial S}{\partial x^k} + \frac{\alpha^2}{4P^2} \frac{\partial P}{\partial x^j} \frac{\partial P}{\partial x^k} \right). \tag{35}$$

To get a scalar, we need to contract g_{jk} with a tensor with two upper indices. The only geometrical object that we have available for this is the inverse metric of the Euclidean configuration space, $h^{ab} = \delta^{ab}$. If we contract g_{jk} with h^{ab}, multiply by the constant $\frac{\alpha}{4m}$, and set $\alpha = \hbar$, the result is

$$H = \frac{\alpha}{4m} \delta^{jk} g_{jk} = \frac{1}{2m} \int d^n x \, \left[P \left| \nabla S \right|^2 + \frac{\hbar^2}{4P} \left| \nabla P \right|^2 \right] = \frac{\hbar^2}{2m} \int d^n x \, \left| \nabla \psi \right|^2. \tag{36}$$

This is the average energy (and the ensemble Hamiltonian) of a free particle in *quantum mechanics*. This leads therefore to the quantum theory of a free particle. In particular, Eq. (6) is equivalent to the Schrödinger equation $i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi$.

Another choice is given by H in the limit where $\hbar \to 0$. Then we get the average energy (and the ensemble Hamiltonian) of a free particle in *classical mechanics*, and the classical theory of a free particle [4].

VI. THE HILBERT SPACE OF QUANTUM MECHANICS FROM THE GEOMETRIC APPROACH

There is a standard construction that associates a complex Hilbert space with any infinite dimensional Kähler space. Given two complex functions ϕ and φ, define the Dirac product by [9]

$$
\begin{aligned}
\langle \phi | \varphi \rangle &= \frac{1}{2} \int \left\{ (\phi(x^\mu), \phi^*(x^\mu)) \cdot [g + i\Omega] \cdot \begin{pmatrix} \varphi(x^\mu) \\ \varphi^*(x^\mu) \end{pmatrix} \right\} d^n x \\
&= \frac{1}{2} \int \left\{ (\phi(x^\mu), \phi^*(x^\mu)) \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + i \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \right] \begin{pmatrix} \varphi(x^\mu) \\ \varphi^*(x^\mu) \end{pmatrix} \right\} d^n x \\
&= \int \phi^*(x^\mu) \varphi(x^\mu) d^n x
\end{aligned}
\tag{37}
$$

In this way, the Hilbert space structure of quantum mechanics follows from the Kähler geometry. In particular, the complex structure that is needed for the formulation of quantum mechanics arises in a very natural way – it is forced upon us by the geometry.

VII. DISCUSSION

We have shown that quantum theory emerges from the geometry of probabilities in motion. This is a reconstruction of quantum theory that seems rather surprising to us, in that none of the elements that are usually assumed to be characteristic of quantum theory (e.g., a representation in terms of an algebra of operators, uncertainty relations, the assumption of a classical theory that is quantized, etc.) are introduced *a priori*. Instead, our starting points are the natural metric on the space of probabilities (information geometry) and the description of evolution in terms of a Hamiltonian formalism (symplectic geometry), together with requirements of consistency (Kähler geometry) and simplicity (the choice of a flat Kähler space).

There is one important additional assumption that needs further discussion, and that is the choice of the translation group when defining the information metric. Since the information metric relates to probabilities, and these probabilities are associated with measurements, it makes sense to look at the role that the translation group plays in measurement theory, in particular as it relates to the *limitations* imposed by measurement uncertainty. This can

provide some insight on why the translation group is important here, even if our remarks are of necessity qualitative rather than quantitative.

It will be useful to go back to one of the oldest concepts in measurement, the idea of *resolution* as it is understood in optics. Here the translation group already makes an appearance. The image of a star, when observed through a telescope with finite resolution, will be blurred. To give an operational definition of the resolution of the telescope, you take two (perhaps hypothetical) point-sources and consider the pair of blurred images. If it is possible to tell them apart, you say that the telescope can resolve them. A measure of resolution is given for example by the well know Rayleigh criterion. It is clear that resolution means the ability to *distinguish* between two point-sources that are close together. There is another operational definition of resolution which is equivalent: take a single point source, consider its image, and then consider the action of the *translation group* on this one image. Define the resolution of your telescope by the smallest distance that you need to move the image such that the superposition of the original image and the displaced image are distinguishable.

In our case, we are dealing with the space of probabilities on an n-dimensional configuration space rather than the space of blurred images of a telescope, but there is a clear analogy between both situations. Just as is done in optics, we introduce a concept of resolution in this space by considering the ability to distinguish between a probability $P(x)$ and a probability $P(x + \theta)$ *that results from the action of the translation group*. Since the metric of Eq. (2) or, equivalently, Eq. (4), is the one that is associated with the action of the translation group, it will quantify the ability to distinguish between the two in the limit of infinitesimal displacements.

Appendix A: Symplectic geometry, compatibility conditions, and Kähler structure

We consider a finite space, but similar relations hold for infinite dimensional spaces. A symplectic vector space is a vector space V that is equipped with a bilinear form $\omega : V \times V \to R$ that is [10]:

a. Skew-symmetric: $\omega(u, v) = -\omega(v, u)$ for all $u, v \, \epsilon \, V$,

b. Non-degenerate: if $\omega(u, v) = 0$ for all $v \, \epsilon \, V$, then $u = 0$.

The standard space is $\Re^{2n}$, and typically ω is chosen to be the matrix

$$\omega_{ab} = \begin{pmatrix} 0 & \mathbf{1}^n \\ -\mathbf{1}^n & 0 \end{pmatrix}, \tag{A1}$$

where $\mathbf{1}^n$ is the unit matrix in n dimensions.

Consider the dual space V^*. The symplectic structure can be identified with an element of $V^* \times V^*$, so that $\omega(u, v) = \omega_{ab} u^a v^b$. Since the spaces V and V^* are isomorphic, there is a ω^{ab} that is the dual of ω_{ab}. This ω^{ab} can be identified with an element of $V \times V$. The convention is to set [10]

$$\omega^{ab} = -(\omega^{-1})^{ab} = \begin{pmatrix} 0 & \mathbf{1}^n \\ -\mathbf{1}^n & 0 \end{pmatrix}, \tag{A2}$$

so that $\omega^{ac} \omega_{cb} = -\delta^a{}_b$.

We assume there is a metric in the space, $g_{ab} = g_{ba}$, and a corresponding inverse metric g^{ab} with $g_{ab} g^{bc} = \delta_{ac}$ (indices are raised and lowered with g_{ab} and g^{ab}) The metric also defines a map $V \to V^*$ to the dual space in an obvious way. Therefore, the space has two linear operators that induce maps $V \to V^*$, ω_{ab} and g_{ab}. They will be related by an equation of the form $\omega_{ab} = g_{ac} j^c{}_b$ for some choice of linear operator $j^c{}_b$. This is Eq. (14), the first of the Kähler conditions.

The relations $\omega^{ac} \omega_{cb} = -\delta^a{}_b$ and $\omega_{ab} = g_{ac} j^c{}_b$ lead to the condition $j^a{}_s j^s{}_c = -\delta^a{}_c$, that is, that $j^a{}_b$ is a complex structure. This is Eq. (16), the third of the Kähler conditions.

Finally, $\omega_{ab} = g_{ac} j^c{}_b$ and $j^a{}_s j^s{}_c = -\delta^a{}_c$, together with the symmetries $-\omega_{cb} = \omega_{bc}$ and $g_{ba} = g_{ab}$, lead to $g_{cd} = j^a{}_c g_{ab} j^b{}_d$ which is Eq. (15), the second of the Kähler conditions.

This shows that consistency requirements imply that a space with both symplectic and metric structures must have a Kähler structure.

---

[1] C. R. Rao, "Differential Metrics in Probability Spaces," in *Differential Geometry in Statistical Inference*, edited by S.-I. Amari, O. E. Barndorff-Nielsen, R. E. Kass, S. L. Lauritzen, and C. R. Rao, Institute of Mathematical Statistics, Hayward, CA, 1987, pp. 217-240.

[2] H. Jeffreys, *Proc. Roy. Soc. A* **186**, 453-461 (1946).

[3] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd ed., Springer-Verlag, New York, 1989.

[4] M. J. W. Hall and M. Reginatto, *Phys. Rev. A* **72**, 062109 (2005).

[5] M. J. W. Hall, *Phys. Rev. A* **78**, 042104 (2008).

[6] M. Reginatto and M. J. W. Hall, *J. Phys.: Conf. Ser.* **174**, 012038 (2009).

[7] S.I. Goldberg, *Curvature and Homology*, Dover Publications, New York, 1982.

[8] R. J. Finkelstein, *Nonrelativistic Mechanics*, W. A. Benjamin, Reading, Massachusetts, 1973.

[9] T. W. B. Kibble, *Commun. math. Phys.* **65**, 189-201 (1979).

[10] J. Stewart, *Advanced General Relativity*, Cambridge University Press, Cambridge, 1990.