

Article

The Entropy-Based Quantum Metric

Roger Balian

Institut de Physique Théorique, CEA/Saclay, F-91191 Gif-sur-Yvette Cedex, France;

E-Mail: roger@balian.fr

Received: 15 May 2014; in revised form: 25 June 2014 / Accepted: 11 July 2014 /

Published: 15 July 2014

Abstract: The von Neumann entropy $S(\hat{D})$ generates in the space of quantum density matrices \hat{D} the Riemannian metric $ds^2 = -d^2S(\hat{D})$, which is physically founded and which characterises the amount of quantum information lost by mixing \hat{D} and $\hat{D} + d\hat{D}$. A rich geometric structure is thereby implemented in quantum mechanics. It includes a canonical mapping between the spaces of states and of observables, which involves the Legendre transform of $S(\hat{D})$. The Kubo scalar product is recovered within the space of observables. Applications are given to equilibrium and non equilibrium quantum statistical mechanics. There the formalism is specialised to the relevant space of observables and to the associated reduced states issued from the maximum entropy criterion, which result from the exact states through an orthogonal projection. Von Neumann's entropy specialises into a relevant entropy. Comparison is made with other metrics. The Riemannian properties of the metric $ds^2 = -d^2S(\hat{D})$ are derived. The curvature arises from the non-Abelian nature of quantum mechanics; its general expression and its explicit form for q-bits are given, as well as geodesics.

Keywords: quantum entropy; metric; q-bit; information; geometry; geodesics; relevant entropy

1. A Physical Metric for Quantum States

Quantum physical quantities pertaining to a given system, termed as “observables” \hat{O} , behave as non-commutative random variables and are elements of a C*-algebra. We will consider below systems for which these observables can be represented by n -dimensional Hermitean matrices in a finite-dimensional Hilbert space \mathcal{H} . In quantum (statistical) mechanics, the “state” of such a system encompasses the expectation values of all its observables [1]. It is represented by a density matrix \hat{D} ,

which plays the rôle of a probability distribution, and from which one can derive the expectation value of \hat{O} in the form

$$\langle \hat{O} \rangle = \text{Tr} \hat{D} \hat{O} = (\hat{D}; \hat{O}). \quad (1)$$

Density matrices should be Hermitean ($\langle \hat{O} \rangle$ is real for $\hat{O} = \hat{O}^\dagger$), normalised (the expectation value of the unit observable is $\text{Tr} \hat{D} = 1$) and non-negative (variances $\langle \hat{O}^2 \rangle - \langle \hat{O} \rangle^2$ are non-negative). They depend on $n^2 - 1$ real parameters. If we keep aside the multiplicative structure of the set of operators and focus on their linear vector space structure, Equation (1) appears as a linear mapping of the space of observables onto real numbers. We can therefore regard the observables and the density operators \hat{D} as elements of two dual vector spaces, and expectation values (1) appear as scalar products.

It is of interest to define a metric in the space of states. For instance, the distance between an exact state \hat{D} and an approximation \hat{D}_{app} would then characterise the quality of this approximation. However, all physical quantities come out in the form (1) which lies astride the two dual spaces of observables and states. In order to build a metric having physical relevance, we need to rely on another meaningful quantity which pertains only to the space of states.

We note at this point that quantum states are probabilistic objects that gather information about the considered system. Then, the amount of missing information is measured by von Neumann's entropy

$$S(\hat{D}) \equiv -\text{Tr} \hat{D} \ln \hat{D}. \quad (2)$$

Introduced in the context of quantum measurements, this quantity is identified with the thermodynamic entropy when \hat{D} is an equilibrium state. In non-equilibrium statistical mechanics, it encompasses, in the form of "relevant entropy" (see Section 5 below), various entropies defined through the maximum entropy criterion. It is also introduced in quantum computation. Alternative entropies have been introduced in the literature, but they do not present all the distinctive and natural features of von Neumann's entropy, such as additivity and concavity.

As $S(\hat{D})$ is a concave function, and as it is the sole physically meaningful quantity apart from expectation values, it is natural to rely on it for our purpose. We thus define [2] the distance ds between two neighbouring density matrices \hat{D} and $\hat{D} + d\hat{D}$ as the square root of

$$ds^2 = -d^2 S(\hat{D}) = \text{Tr} d\hat{D} d \ln \hat{D}. \quad (3)$$

This Riemannian metric is of the Hessian form since the metric tensor is generated by taking second derivatives of the function $S(\hat{D})$ with respect to the $n^2 - 1$ coordinates of \hat{D} . We may take for such coordinates the real and imaginary parts of the matrix elements, or equivalently (Section 6) some linear transform of these (keeping aside the norm $\text{Tr} \hat{D} = 1$).

2. Interpretation in the Context of Quantum Information

The simplest example, related to quantum information theory, is that of a q-bit (two-level system or spin $\frac{1}{2}$) for which $n = 2$. Its states, represented by 2×2 Hermitean normalised density matrices \hat{D} , can conveniently be parameterised, on the basis of Pauli matrices, by the components $r_\mu = D_{12} + D_{21}$,

$i(D_{12} - D_{21}), D_{11} - D_{22}$ ($\mu = 1, 2, 3$) of a 3-dimensional vector \mathbf{r} lying within the unit Poincaré–Bloch sphere ($r \leq 1$). From the corresponding entropy

$$S = \frac{1+r}{2} \ln \frac{2}{1+r} + \frac{1-r}{2} \ln \frac{2}{1-r}, \tag{4}$$

we derive the metric

$$ds^2 = \frac{1}{1-r^2} \left(\frac{\mathbf{r} \cdot d\mathbf{r}}{r} \right)^2 + \frac{1}{2r} \ln \frac{1+r}{1-r} \left\| \frac{\mathbf{r} \times d\mathbf{r}}{r} \right\|^2, \tag{5}$$

which is a natural Riemannian metric for q-bits, or more generally for positive 2×2 matrices. The metric tensor characterizing (5) diverges in the vicinity of pure states $r = 1$, due to the singularity of the entropy (2) for vanishing eigenvalues of \hat{D} . However, the distance between two arbitrary (even pure) states \hat{D}' and \hat{D}'' measured along a geodesic is always finite. We shall see (Equation (29)) that for $n = 2$ the geodesic distance s between two neighbouring pure states \hat{D}' and \hat{D}'' , represented by unit vectors \mathbf{r}' and \mathbf{r}'' making a small angle $\delta\varphi \sim |\mathbf{r}' - \mathbf{r}''|$, behaves as $\delta s^2 \sim \delta\varphi^2 \ln(4\sqrt{\pi}/\delta\varphi)$. The singularity of the metric tensor manifests itself through this logarithmic factor.

Identifying von Neumann’s entropy to a measure of missing information, we can give a simple interpretation to the distance between two states. Indeed, the concavity of entropy expresses that some information is lost when two statistical ensembles described by different density operators merge. By mixing two equal size populations described by the neighbouring distributions $\hat{D}' = \hat{D} + \frac{1}{2}\delta\hat{D}$ and $\hat{D}'' = \hat{D} - \frac{1}{2}\delta\hat{D}$ separated by a distance δs , we lose an amount of information given by

$$\Delta S \equiv S(\hat{D}) - \frac{S(\hat{D}') + S(\hat{D}'')}{2} \sim \frac{\delta s^2}{8}, \tag{6}$$

and thereby directly related to the distance δs defined by (3). The proof of this equivalence relies on the expansion of the entropies $S(\hat{D}')$ and $S(\hat{D}'')$ around \hat{D} , and is valid when $\text{Tr} \delta\hat{D}^2$ is negligible compared to the smallest eigenvalue of \hat{D} . If \hat{D}' and \hat{D}'' are distant, the quantity $8\Delta S$ cannot be regarded as the square of a distance that would be generated by a local metric. The equivalence (6) for neighbouring states shows that ds^2 is the metric that is the best suited to measure losses of information by mixing.

The singularity of δs^2 at the edge of the positivity domain of \hat{D} may suggest that the result (6) holds only within this domain. In fact, this equivalence remains nearly valid even in the limit of pure states because ΔS itself involves a similar singularity. Indeed, if the states $\hat{D}' = |\psi' \rangle \langle \psi'|$ and $\hat{D}'' = |\psi'' \rangle \langle \psi''|$ are pure and close to each other, the loss of information ΔS behaves as $8\Delta S \sim \delta\varphi^2 \ln(4/\delta\varphi)$ where $\delta\varphi^2 \sim 2 \text{Tr} \delta\hat{D}^2$. This result should be compared to various geodesic distances between pure quantum states, which behave as $\delta s^2 \sim \delta\varphi^2 \ln(4\sqrt{\pi}/\delta\varphi)$ for the present metric, and as $\delta s_{\text{BH}}^2 = 4\delta s_{\text{FS}}^2 \sim \delta\varphi^2 \sim \text{Tr}(\hat{D}' - \hat{D}'')^2$ for the Bures – Helstrom and the quantum Fubini – Study metrics, respectively (see Section 7; these behaviours hold not only for $n = 2$ but for arbitrary n since only the space spanned by $|\psi' \rangle$ and $|\psi'' \rangle$ is involved). Thus, among these metrics, only $ds^2 = -d^2S$ can be interpreted in terms of information loss, whether the states \hat{D}' and \hat{D}'' are pure or mixed.

At the other extreme, around the most disordered state $\hat{D} = \hat{I}/n$, in the region $\| n\hat{D} - \hat{I} \| \ll 1$, the metric becomes Euclidean since $ds^2 = \text{Tr} d\hat{D}d \ln \hat{D} \sim n \text{Tr}(d\hat{D})^2$ (for $n = 2$, $ds^2 = d\mathbf{r}^2$). For a given shift $d\hat{D}$, the qualitative change of a state \hat{D} , as measured by the distance ds , gets larger and larger as the state \hat{D} becomes purer and purer, that is, when the information contents of \hat{D} increases.

3. Geometry of Quantum Statistical Mechanics

A rich geometric structure is generated for both states and observables by von Neumann’s entropy through introduction of the metric $ds^2 = -d^2S$. Now, this metric (3) supplements the algebraic structure of the set of observables and the above duality between the vector spaces of states and of observables, with scalar product (1). Accordingly, we can define naturally within the space of states scalar products, geodesics, angles, curvatures.

We can also regard the coordinates of $d\hat{D}$ and $d \ln \hat{D}$ as covariant and contravariant components of the same infinitesimal vector (Section 6). To this aim, let us introduce the mapping

$$\hat{D} \equiv \frac{e^{\hat{X}}}{\text{Tr } e^{\hat{X}}} \tag{7}$$

between \hat{D} in the space of states and \hat{X} in the space of observables. The operator \hat{X} appears as a parameterisation of \hat{D} . (The normalisation of \hat{D} entails that \hat{X} , defined within an arbitrary additive constant operator $X_0 \hat{I}$, also depends on $n^2 - 1$ independent real parameters.) The metric (3) can then be re-expressed in terms of \hat{X} in the form

$$ds^2 = \text{Tr } d\hat{D}d\hat{X} = \text{Tr} \int_0^1 d\xi \hat{D} e^{-\xi\hat{X}} d\hat{X} e^{\xi\hat{X}} d\hat{X} - (\text{Tr } \hat{D}d\hat{X})^2 = d^2 \ln \text{Tr } e^{\hat{X}} = d^2 F, \tag{8}$$

where we introduced the function

$$F(\hat{X}) \equiv \ln \text{Tr } e^{\hat{X}} \tag{9}$$

of the observable \hat{X} (The addition of $X_0 \hat{I}$ to \hat{X} results in the addition of the irrelevant constant X_0 to F). This mapping provides us with a natural metric in the space of observables, from which we recover the scalar product between $d\hat{X}_1$ and $d\hat{X}_2$ in the form of a Kubo correlation in the state \hat{D} . The metric (8) has been quoted in the literature under the names of Bogoliubov–Kubo–Mori.

4. Covariance and Legendre Transformation

We can recover the above geometric mapping (7) between \hat{D} and \hat{X} , or between the covariant and contravariant coordinates of $d\hat{D}$, as the outcome of a Legendre transformation, by considering the function $F(\hat{X})$. Taking its differential $dF = \text{Tr } e^{\hat{X}} d\hat{X} / \text{Tr } e^{\hat{X}}$, we identify the partial derivatives of $F(\hat{X})$ with the coordinates of the state $\hat{D} = e^{\hat{X}} / \text{Tr } e^{\hat{X}}$, so that \hat{D} appears as conjugate to \hat{X} in the sense of Legendre transformations. Expressing then \hat{X} as function of \hat{D} and inserting into $F - \text{Tr } \hat{D}\hat{X}$, we recognise that the Legendre transform of $F(\hat{X})$ is von Neumann’s entropy $F - \text{Tr } \hat{D}\hat{X} = S(\hat{D}) = -\text{Tr } \hat{D} \ln \hat{D}$. The conjugation between \hat{D} and \hat{X} is embedded in the equations

$$dF = \text{Tr } \hat{D}d\hat{X} ; \quad dS = -\text{Tr } \hat{X}d\hat{D} . \tag{10}$$

Legendre transformations are currently used in equilibrium thermodynamics. Let us show that they come out in this context directly as a special case of the present general formalism. The entropy of thermodynamics is a function of the extensive variables, energy, volume, particle numbers, etc. Let us focus for illustration on the energy U , keeping the other extensive variables fixed. The thermodynamic entropy $S(U)$, a function of the single variable U , generates the inverse temperature as $\beta = \partial S / \partial U$.

Its Legendre transform is the Massieu potential $F(\beta) = S - \beta U$. In order to compare these properties with the present formalism, we recall how thermodynamics comes out in the framework of statistical mechanics. The thermodynamic entropy $S(U)$ is identified with the von Neumann entropy (2) of the Boltzmann–Gibbs canonical equilibrium state \hat{D} , and the internal energy with $U = \text{Tr} \hat{D} \hat{H}$. In the relation (7), the operator \hat{X} reads $\hat{X} = -\beta \hat{H}$ (within an irrelevant additive constant). By letting U or β vary, we select within the spaces of states and of observables a one-dimensional subset. In these restricted subsets, \hat{D} is parameterised by the single coordinate U , and the corresponding \hat{X} by the coordinate $-\beta$. By specialising the general relations (10) to these subsets, we recover the thermodynamic relations $dF = -U d\beta$ and $dS = \beta dU$. We also recover, by restricting the metric (3) or (8) to these subsets, the current thermodynamic metric $ds^2 = -(\partial^2 S / \partial U^2) dU^2 = -dU d\beta$.

More generally, we can consider the Boltzmann–Gibbs states of equilibrium statistical mechanics as the points of a manifold embedded in the full space of states. The thermodynamic extensive variables, which parameterise these states, are the expectation values of the conserved macroscopic observables, that is, they are a subset of the expectation values (1) which parameterise arbitrary density operators. Then the standard geometric structure of thermodynamics simply results from the restriction of the general metric (3) to this manifold of Boltzmann–Gibbs states. The commutation of the conserved observables simplifies the reduced thermodynamic metric, which presents the same features as a Fisher metric (see Section 6).

5. Relevant Entropy and Geometry of the Projection Method

The above ideas also extend to non-equilibrium quantum statistical mechanics [2–4]. When introducing the metric (3), we indicated that it may be used to estimate the quality of an approximation. Let us illustrate this point with the Nakajima–Zwanzig–Mori–Robertson projection method, best introduced through maximum entropy. Consider some set $\{\hat{A}_k\}$ of “relevant observables”, whose time-dependent expectation values $a_k \equiv \langle \hat{A}_k \rangle = \text{Tr} \hat{D} \hat{A}_k$ we wish to follow, discarding all other variables. The exact state \hat{D} encodes the variables $\{a_k\}$ that we are interested in, but also the expectation values (1) of the other observables that we wish to eliminate. This elimination is performed by associating at each time with \hat{D} a “reduced state” \hat{D}_R which is equivalent to \hat{D} as regards the set $a_k = \text{Tr} \hat{D}_R \hat{A}_k$, but which provides no more information than the values $\{a_k\}$. The former condition provides the constraints $\langle \hat{A}_k \rangle = a_k$, and the latter condition is implemented by means of the maximum entropy criterion: One expresses that, within the set of density matrices compatible with these constraints, \hat{D}_R is the one which maximises von Neumann’s entropy (2), that is, which contains solely the information about the relevant variables a_k . The least biased state \hat{D}_R thus defined has the form $\hat{D}_R = e^{\hat{X}_R} / \text{Tr} e^{\hat{X}_R}$, where $\hat{X}_R \equiv \sum_k \lambda_k \hat{A}_k$ involves the time-dependent Lagrange multipliers λ_k , which are related to the set a_k through $\text{Tr} \hat{D}_R \hat{A}_k = a_k$.

The von Neumann entropy $S(\hat{D}_R) \equiv S_R\{a_k\}$ of this reduced state \hat{D}_R is called the “relevant entropy” associated with the considered relevant observables \hat{A}_k . It measures the amount of missing information, when only the values $\{a_k\}$ of the relevant variables are given. During its evolution, \hat{D} keeps track of the initial information about all the variables $\langle \hat{O} \rangle$ and its entropy $S(\hat{D})$ remains constant in time. It is therefore smaller than the relevant entropy $S(\hat{D}_R)$ which accounts for the loss of information about the

irrelevant variables. Depending on the choice of relevant observables $\{\hat{A}_k\}$, the corresponding relevant entropies $S_R\{a_k\}$ encompass various current entropies, such as the non-equilibrium thermodynamic entropy or Boltzmann’s H-entropy.

The same structure as the one introduced above for the full spaces of observables and states is recovered in this context. Here, for arbitrary values of the parameters λ_k , the exponents $\hat{X}_R = \sum_k \lambda_k \hat{A}_k$ constitute a subspace of the full vector space of observables, and the parameters $\{\lambda_k\}$ appear as the coordinates of \hat{X}_R on the basis $\{\hat{A}_k\}$. The corresponding states \hat{D}_R , parameterised by the set $\{a_k\}$, constitute a subset of the space of states, the manifold \mathcal{R} of “reduced states”(Note that this manifold is not a hyperplane, contrary to the space of relevant observables; it is embedded in the full vector space of states, but does not constitute a subspace). By regarding $S_R\{a_k\}$ as a function of the coordinates $\{a_k\}$, we can define a metric $ds^2 = -d^2 S_R\{a_k\}$ on the manifold \mathcal{R} , which is the restriction of the metric (3). Its alternative expression $ds^2 = \sum_k da_k d\lambda_k = d^2 F_R\{\lambda_k\}$, where $F_R\{\lambda_k\} \equiv \ln \text{Tr} \exp \sum_k \lambda_k \hat{A}_k$, is a restriction of (8). The correspondence between the two parameterisations $\{a_k\}$ and $\{\lambda_k\}$ is again implemented by the Legendre transformation which relates $S_R\{a_k\}$ and $F_R\{\lambda_k\}$.

The projection method relies on the mapping $\hat{D} \mapsto \hat{D}_R$ which associates \hat{D}_R to \hat{D} . It consists in replacing the Liouville–von Neumann equation of motion for \hat{D} by the corresponding dynamical equation for \hat{D}_R on the manifold \mathcal{R} , or equivalently for the coordinates $\{a_k\}$ or for the coordinates $\{\lambda_k\}$, a programme that is in practice achieved through some approximations. This mapping is obviously a projection in the sense that $\hat{D} \mapsto \hat{D}_R \mapsto \hat{D}_R$, but moreover the introduction of the metric (3) shows that the vector $\hat{D} - \hat{D}_R$ in the space of states is perpendicular to the manifold \mathcal{R} at the point \hat{D}_R . This property is readily shown by writing, in this metric, the scalar product $\text{Tr} d\hat{D} d\hat{X}'$ of the vector $d\hat{D} = \hat{D} - \hat{D}_R$ by an arbitrary vector $d\hat{X}'$ in the tangent plane of \mathcal{R} . The latter is conjugate to any combination $d\hat{X}'$ of observables \hat{A}_k , and this scalar product vanishes because $\text{Tr} \hat{D} \hat{A}_k = \text{Tr} \hat{D}_R \hat{A}_k$. Thus the mapping $\hat{D} \mapsto \hat{D}_R$ appears as an orthogonal projection, so that the relevant state \hat{D}_R associated with \hat{D} may be regarded as its best possible approximation on the manifold \mathcal{R} .

6. Properties of the Metric

The metric tensor can be evaluated explicitly in a basis where the matrix \hat{D} is diagonal. Denoting by D_i its eigenvalues and by dD_{ij} the matrix elements of its variations, we obtain from (3)

$$ds^2 = \text{Tr} \int_0^\infty d\xi \left(\frac{d\hat{D}}{\hat{D} + \xi} \right)^2 = \sum_{ij} \frac{\ln D_i - \ln D_j}{D_i - D_j} dD_{ij} dD_{ji} . \tag{11}$$

(For $D_i = D_j$, whether or not $i = j$, the ratio is defined as $1/D_i$ by continuity.) In the same basis, the form (8) of the metric reads

$$ds^2 = \frac{1}{Z} \sum_{ij} \frac{e^{X_i} - e^{X_j}}{X_i - X_j} dX_{ij} dX_{ji} - \left(\frac{\sum_i e^{X_i} dX_{ii}}{Z} \right)^2 , \tag{12}$$

with $Z = \sum_i e^{X_i}$ (For $X_i = X_j$, the ratio is e^{X_i}). The singularity of the metric (11) in the vicinity of vanishing eigenvalues of \hat{D} , in particular near pure states (end of Section 2), is not apparent in the representation (12) of this metric, because the mapping from \hat{D} to \hat{X} sends the eigenvalue X_i to $-\infty$ when D_i tends to zero.

Let us compare the expression (11) with the corresponding classical metric, which is obtained by starting from Shannon’s entropy instead of von Neumann’s entropy. For discrete probabilities p_i , we have then $S\{p_i\} = -\sum_i p_i \ln p_i$ and hence the same definition $ds^2 = -d^2S\{p_i\}$ as above of an entropy-based metric yields $ds^2 = \sum_i dp_i^2/p_i$, which is identified with the Fisher information metric. The present metric thus appears as the extension to quantum statistical mechanics of the Fisher metric when the latter is interpreted in terms of entropy. In fact, the terms of (11) which involve the diagonal elements $i = j$ of the variations $d\hat{D}$ reduce to dD_{ii}^2/D_i . This result was expected since density matrices behave as probability distributions if both \hat{D} and $d\hat{D}$ are diagonal.

Let us more generally consider in (11), instead of solely diagonal variations dD_{ii} , variations dD_{ij} with indices i and j such that $|D_i - D_j| \ll D_i + D_j$. The expansion of D_i and D_j around $\frac{1}{2}(D_i + D_j)$ in the corresponding ratios of (11) yields $(\ln D_i - \ln D_j)/(D_i - D_j) \sim 2/(D_i + D_j)$. The considered terms of (11) are therefore the same as in the Bures–Helstrom metric

$$ds_{\text{BH}}^2 = \sum_{ij} \frac{2}{D_i + D_j} dD_{ij} dD_{ji}, \tag{13}$$

introduced long ago as an extension to matrices of the Fisher metric [5]. We thus recover this Bures–Helstrom metric as an approximation of the present entropy-based metric $ds^2 = -d^2S(\hat{D})$. For $n = 2$, ds_{BH}^2 is obtained from the expression (5) of ds^2 by omitting the factor $\tanh^{-1} r/r$ entering the second term.

In order to express the properties of the Riemannian metric (3) in a general form, which will exhibit the tensor structure, we use a Liouville representation. There, the observables $\hat{O} = O_\mu \hat{\Omega}^\mu$, regarded as elements of a vector space, are represented by their coordinates O_μ on a complete basis $\hat{\Omega}^\mu$ of n^2 observables. The space of states is spanned by the dual basis $\hat{\Sigma}_\mu$, such that $\text{Tr} \hat{\Omega}^\nu \hat{\Sigma}_\mu = \delta_\mu^\nu$, and the states $\hat{D} = D^\mu \hat{\Sigma}_\mu$ are represented by their coordinates D_μ . Thus, the expectation value (1) is the scalar product $D^\mu O_\mu$. In the matrix representation which appears as a special case, μ denotes a pair of indices i, j , $\hat{\Omega}^\mu$ stands for $|j \rangle \langle i|$, $\hat{\Sigma}_\mu$ for $|i \rangle \langle j|$, O_μ denotes the matrix element O_{ji} and D^μ the element D_{ij} . For the q-bit ($n = 2$) considered in Section 2, we have chosen the Pauli operators $\hat{\sigma}^\mu$ as basis $\hat{\Omega}^\mu$ for observables, and $\frac{1}{2}\hat{\sigma}_\mu$ as dual basis $\hat{\Sigma}_\mu$ for states, so that the coordinates $D^\mu = \text{Tr} \hat{D} \hat{\Omega}^\mu$ of $\hat{D} = \frac{1}{2}(\hat{I} + r^\mu \hat{\sigma}_\mu)$ are the components r^μ of the vector \mathbf{r} (The unit operator \hat{I} is kept aside since \hat{D} is normalised and since constants added to \hat{X} are irrelevant). The function $F\{X\} = \ln \text{Tr} e^{\hat{X}}$ of the coordinates X_μ of the observable \hat{X} , and the von Neumann entropy $S\{D\}$ as function of the coordinates D^μ of the state \hat{D} , are related by the Legendre transformation $F = S + D^\mu X_\mu$, and the relations (10) are expressed by $D^\mu = \partial F / \partial X_\mu$, $X_\mu = -\partial S / \partial D^\mu$. The metric tensor is given by

$$g^{\mu\nu} = \frac{\partial^2 F}{\partial X_\mu \partial X_\nu}, \quad g_{\mu\nu} = -\frac{\partial^2 S}{\partial D^\mu \partial D^\nu}, \tag{14}$$

and the correspondence issued from (7) between covariant and contravariant infinitesimal variations of \hat{X} and \hat{D} is implemented as $dD^\mu = g^{\mu\nu} dX_\nu$, $dX_\mu = g_{\mu\nu} dD^\nu$.

These expressions exhibit the Hessian nature of the metric. This property simplifies the expression of the Christoffel symbol, which reduces to

$$\Gamma_{\mu\nu\rho} = -\frac{1}{2} \frac{\partial^3 S}{\partial D^\mu \partial D^\nu \partial D^\rho}, \tag{15}$$

and which provides a parametric representation $\hat{D}(t)$ of the geodesics in the space of states through

$$\frac{d^2 D^\mu}{dt^2} + g^{\mu\sigma} \Gamma_{\sigma\nu\rho} \frac{dD^\nu}{dt} \frac{dD^\rho}{dt} = 0. \tag{16}$$

Then, the Riemann curvature tensor comes out as

$$R_{\mu\rho\nu\sigma} = g^{\xi\zeta} (\Gamma_{\mu\sigma\xi} \Gamma_{\nu\rho\zeta} - \Gamma_{\mu\nu\xi} \Gamma_{\rho\sigma\zeta}), \tag{17}$$

the Ricci tensor and the scalar curvature as

$$R_{\mu\nu} = g^{\rho\sigma} R_{\mu\rho\nu\sigma}, \quad R = g^{\mu\nu} R_{\mu\nu}, \tag{18}$$

We have noted that the classical equivalent of the entropy-based metric $ds^2 = -d^2S$ is the Fisher metric $\sum_i dp_i^2/p_i$, which as regards the curvature is equivalent to a Euclidean metric. While the space of classical probabilities is thus flat, the above equations show that the space of quantum states is curved. This curvature arises from the non-commutation of the observables, it vanishes for the completely disordered state $\hat{D} = \hat{I}/n$. Curvature can thus be used as a measure of the degree of classicality of a state.

7. Geometry of the Space of q-Bits

In the illustrative example of a q-bit, the operator $\hat{X} = \chi_\mu \hat{\sigma}^\mu$ associated with \hat{D} is parameterised by the 3 components of the vector χ_μ ($\mu = 1, 2, 3$), related to \mathbf{r} by $\chi = \tanh^{-1} r$ and $\chi_\mu/\chi = r^\mu/r$. The metric tensor given by (5) is expressed as

$$g_{\mu\nu} = Kr_\mu r_\nu + \frac{\chi}{r} \delta_{\mu\nu}, \quad K \equiv \frac{1}{r} \frac{d\chi}{dr} = \frac{1}{r^2} \left(\frac{1}{1-r^2} - \frac{\chi}{r} \right), \tag{19}$$

$$g^{\mu\nu} = (1-r^2)p^{\mu\nu} + \frac{r}{\chi} q^{\mu\nu}.$$

(We have defined $r_\mu = r^\mu$, $\delta_{\mu\nu} = \delta^\mu_\nu = \delta^{\mu\nu}$ so as to introduce the projectors $r^\mu r^\nu / r^2 \equiv p^{\mu\nu} \equiv \delta^{\mu\nu} - q^{\mu\nu}$ in the Euclidean 3-dimensional space, and thus to simplify the subsequent calculations.) In polar coordinates $\mathbf{r} = (r, \theta, \varphi)$, the infinitesimal distance takes the form

$$ds^2 = dr d\chi + r\chi(d\theta^2 + \sin^2\theta d\varphi^2). \tag{20}$$

We determine from (15) and (19) the explicit form

$$\Gamma_{\mu\nu\rho} = \frac{K}{2} (r_\mu \delta_{\nu\rho} + r_\nu \delta_{\mu\rho} + r_\rho \delta_{\mu\nu}) + \frac{1}{2r} \frac{dK}{dr} r_\mu r_\nu r_\rho \tag{21}$$

of the Christoffel symbol. By raising its first index with $g^{\mu\nu}$ and using polar coordinates, we obtain from (16) the equations of geodesics for $n = 2$. Within the Poincaré–Bloch sphere the geodesics are deduced by rotations from a one-parameter family of curves which lie in the $\theta = \frac{1}{2}\pi$, $|\varphi| \leq \frac{1}{2}\pi$ half-plane and which are symmetric with respect to the $\varphi = 0$ axis. This family is characterized by the equations (where $\chi = \tanh^{-1} r$):

$$\frac{d^2 r}{dt^2} + \frac{r}{1-r^2} \left(\frac{dr}{dt} \right)^2 - \frac{r}{2} \left[1 + \frac{\chi}{r} (1-r^2) \right] \left(\frac{d\varphi}{dt} \right)^2 = 0, \tag{22}$$

$$\frac{d^2\varphi}{dt^2} + \frac{1}{r} \frac{dr}{dt} \frac{d\varphi}{dt} + \frac{1}{\chi} \frac{d\chi}{dt} \frac{d\varphi}{dt} = 0, \tag{23}$$

and the boundary conditions at $t = 0$:

$$r(0) = a, \quad \varphi(0) = 0, \quad \frac{dr(0)}{dt} = 0, \quad \frac{d\varphi(0)}{dt} = \frac{1}{k}, \quad k^2 = a \tanh^{-1} a. \tag{24}$$

Equation (23) provides, using the boundary conditions (24):

$$\frac{d\varphi}{dt} = \frac{k}{r\chi}. \tag{25}$$

Insertion of (25) into (22) gives rise to an equation for $r(t)$, which can be integrated by regarding t as a function of $\zeta = \arcsin r$. One obtains:

$$\left(\frac{dr}{dt}\right)^2 = (1 - r^2) \left(1 - \frac{k^2}{r\chi}\right). \tag{26}$$

The scale of t has been fixed by relating to $r(0)$ the boundary condition (24) for $d\varphi(0)/dt$, a choice which ensures that $ds^2 = drd\chi + r\chi d\varphi^2 = dt^2$, and hence that the parameter t measures the distance along geodesics.

For $k = 0$, we obtain $r = |\sin t|$, $\varphi = \pm\pi/2$. Thus, the longest geodesics are the diameters of the Poincaré–Bloch sphere. We find the value π for their “length”, that is, for the geodesic distance between two orthogonal pure states. At the other extreme, when the middle point $r = a$, $\varphi = 0$ of a geodesic lies close to the surface $r = 1$ of the sphere, the asymptotic form of the equation (26) is solved as

$$t = \pm 2k\sqrt{\pi}e^{-k^2} \operatorname{erf} \xi, \quad \xi = \sqrt{\frac{1}{2} \ln \frac{1-a}{1-r}}, \quad k^2 = \frac{1}{2} \ln \frac{2}{1-a} \tag{27}$$

(by taking ξ as variable instead of r). The determination of the explicit equations of such short geodesic curves is achieved by integrating (25) into

$$\varphi = \frac{t}{k} = \pm 2\sqrt{\pi}e^{-k^2} \operatorname{erf} \xi. \tag{28}$$

From (27) and (28) we can determine the geodesic distance between two neighbouring pure states $\hat{D}' = |\psi' \rangle \langle \psi'|$ and $\hat{D}'' = |\psi'' \rangle \langle \psi''|$ represented by the points $r_{\max} = 1$, $\varphi_{\max} = \pm\frac{1}{2}\delta\varphi$ with $\delta\varphi$ small. At these two points, we have $\xi \rightarrow \infty$, $\operatorname{erf} \xi = 1$, and this determines k in terms of $\frac{1}{2}\delta\varphi$ through (28). The length of the geodesic that joins them, given by (27), is:

$$\delta s^2 = \delta\varphi^2 \ln \frac{4\sqrt{\pi}}{\delta\varphi}, \quad \delta\varphi = \arccos |\langle \psi' | \psi'' \rangle|. \tag{29}$$

Thus, in spite of its singularity for $r = 1$, the present 3-dimensional metric (5) in the space r, θ, φ defines distances between pure states represented by points on the surface $r = 1$ of the Poincaré–Bloch sphere. However, It should be noted that the presence of the logarithmic factor in (29) forbids such distances to be generated by a 2-dimensional metric in the space θ, φ . In fact, the distance (29) is measured along a geodesic that penetrates the sphere $r = 1$, because no geodesic is tangent to the surface of this sphere nor lies on its surface.

In contrast, all geodesics produced by the Bures–Helstrom metric are tangent to the surface of the sphere, or are its great circles. They are given by Equations (25) and (26), where χ is replaced by r and k by a ; the solution of these equations provides the ellipses

$$r \cos \varphi = a \cos t, \quad r \sin \varphi = \sin t. \quad (30)$$

Here as above, the largest distance π is reached for orthogonal pure states represented by opposite points on the sphere, but now a peculiarity occurs. Whereas the metric $ds^2 = -d^2S$ produces a single geodesic, the diameter joining these two points (with “length” π), the Bures metric produces a double infinity of geodesics, the half-ellipses (30) having as long axis this diameter, and having all the same “length” π . Other pairs of pure states are joined by geodesics which are arcs of great circles, and their Bures distance $\delta s_{\text{BH}} = \delta\varphi$ is identified with the ordinary length of the arc. Here for $n = 2$ as in the general case, the 3-dimensional Bures–Helstrom metric admits a restriction to pure states generated by a 2-dimensional metric, which is identified with the quantum Fubini–Study metric, itself defined only for pure states by $s_{\text{FS}} = \arccos |\langle \psi' | \psi'' \rangle| = \frac{1}{2}s_{\text{BH}}$.

Returning to the metric $ds^2 = d^2S$, the Riemann curvature is obtained from (17) as

$$R^\mu_{\rho\nu\sigma} = \frac{K}{4} \left[\left(r^2 + \frac{r}{\chi} - 1 \right) (q^\mu_\sigma q_{\nu\rho} - q^\mu_\nu q_{\rho\sigma}) + \left(r^2 - \frac{r}{\chi} + 1 \right) (p^\mu_\sigma q_{\nu\rho} - p^\mu_\nu q_{\rho\sigma}) \right. \\ \left. + \frac{r}{\chi} \frac{1}{1-r^2} \left(r^2 - \frac{r}{\chi} + 1 \right) (q^\mu_\sigma p_{\nu\rho} - q^\mu_\nu p_{\rho\sigma}) \right]. \quad (31)$$

Contracting with $g^{\rho\sigma}$ the indices of (30) as in (18), we finally derive the Ricci curvature

$$R^\mu_{\nu} = -\frac{Kr}{2\chi} \left(r^2 \delta^\mu_{\nu} + \frac{\chi - r}{\chi} p^\mu_{\nu} \right), \quad (32)$$

and the scalar curvature

$$R = -\frac{Kr}{2\chi} \left(3r^2 + \frac{\chi - r}{\chi} \right). \quad (33)$$

Both are negative in the whole Poincaré sphere. In the limit $r \rightarrow 0$, the curvature R vanishes as $R \sim -\frac{10}{9}r^2$, as expected from the general argument of Section 2: a weakly polarised spin behaves classically. At the other extreme $r \rightarrow 1$, R behaves as $R \sim -2[(1-r) |\ln(1-r)|]^{-1}$; it diverges, again as expected: pure states have the largest quantum nature.

The metric $ds^2 = -d^2S$, introduced above in the context of quantum mechanics for mixed states (and their pure limit) and information theory, might more generally be useful to characterise distances in spaces of positive matrices.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Thirring, W. Quantum Mechanics of Large Systems. In *A Course of Mathematical Physics*; Volume 4; Springer-Verlag: New York, NY, USA, 1983.

2. Balian, R.; Alhassid, Y.; Reinhardt, H. Dissipation in many-body systems: A geometric approach based on information theory. *Phys. Rep.* **1986**, *131*, 1–146.
3. Balian, R. Incomplete descriptions and relevant entropies. *Am. J. Phys.* **1999**, *67*, 1078–1090.
4. Balian, R. Information in statistical physics. *Stud. Hist. Philos. Mod. Phys.* **2005**, *36*, 323–353.
5. Bures, D. An extension of Kakutani's theorem. *Trans. Am. Math. Soc.* **1969**, *135*, 199–212.

© 2014 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).