

The Uniqueness of the Chentsov Metric

M. R. Grasselli* R. F. Streater
Dept. of Maths., King's College London, Strand, WC2R 2LS

13/07/2000

Abstract

We show that on the space of faithful density matrices, the only monotone metrics, for which the exponential and mixture affine connections are mutually dual, are constant multiples of the Bogoliubov-Kubo-Mori (*BKM*) metric.

Introduction

Information manifolds are equipped with two natural flat connections: the mixture connection, obtained from the linear structure of trace class operators themselves, and the exponential connection, obtained when combinations of states are performed by adding their logarithms [4]. Following Amari [1, 2], we consider duality to be the fundamental structure. Thus, given these two connections, we find in §3 all the Riemannian metrics that make them dual. In §4, we combine this result with Petz's characterisation [9] of monotone metrics to find that the *BKM* metric is, up to a factor, the unique monotone metric with respect to which the exponential and mixture connections are dual.

The Exponential and Mixture Connections

Let \mathcal{H}^N be a finite dimensional complex Hilbert space, \mathcal{A} the subset of self-adjoint matrices and \mathcal{M} the set of all invertible density matrices on \mathcal{H}^N . Then \mathcal{A} is an N^2 -dimensional real vector space and \mathcal{M} can be regarded as

*Supported by a grant from CAPES-Brazil.

an n -dimensional manifold with $n = N^2 - 1$. Defining the 1-embedding of \mathcal{M} into \mathcal{A} as

$$\ell_1 : \mathcal{M} \rightarrow \mathcal{A} \quad \rho \mapsto \log \rho,$$

we can use the affine structure of \mathcal{A} to obtain an affine structure on \mathcal{M} . Since then \mathcal{M} is flat, we can identify \mathcal{M} with its tangent space. At each point $\rho \in \mathcal{M}$, consider the subspace $\mathcal{A}_\rho = \{A \in \mathcal{A} : \text{Tr}(\rho A) = 0\}$ of \mathcal{A} , called the space of ‘scores’; we define the isomorphism

$$(\ell_1)_{*(\rho)} : T_\rho \mathcal{M} \rightarrow \mathcal{A}_\rho \quad v \mapsto (\ell_1 \circ \gamma)'(0),$$

where $\gamma : (-\varepsilon, \varepsilon) \rightarrow \mathcal{M}$ is a curve in the equivalence class of the tangent vector v . We call this isomorphism the 1-representation of the tangent space $T_\rho \mathcal{M}$. If $(\theta^1, \dots, \theta^n)$ is a coordinate system for \mathcal{M} , then the 1-representation of the basis $\left\{ \frac{\partial}{\partial \theta^1} \Big|_\rho, \dots, \frac{\partial}{\partial \theta^n} \Big|_\rho \right\}$ of $T_\rho \mathcal{M}$ is $\left\{ \frac{\partial \log \rho}{\partial \theta^1}, \dots, \frac{\partial \log \rho}{\partial \theta^n} \right\}$. The 1-representation of a vector field X on \mathcal{M} is therefore the \mathcal{A} -valued function $(X)^{(1)}$ given by $(X)^{(1)}(\rho) = (\ell_1)_{*(\rho)} X_\rho$.

The exponential- or 1-connection is that got using the 1-embedding and the following parallel transport [4]

$$\tau_{\rho_0, \rho_1}^{(1)} : T_{\rho_0} \mathcal{M} \rightarrow T_{\rho_1} \mathcal{M} \quad v \mapsto (\ell_1)_{*(\rho_1)}^{-1} \left((\ell_1)_{*(\rho_0)} v - \text{Tr}[\rho_1 (\ell_1)_{*(\rho_0)} v] \right).$$

Giving the parallel transport in a ‘hood of ρ is equivalent to specifying the covariant derivative. The 1-representation of the 1-covariant derivative is

$$\left(\nabla_{\frac{\partial}{\partial \theta^i}}^{(1)} \frac{\partial}{\partial \theta^j} \right)^{(1)} = \frac{\partial^2 \log \rho}{\partial \theta^i \partial \theta^j} - \text{Tr} \left(\rho \frac{\partial^2 \log \rho}{\partial \theta^i \partial \theta^j} \right) \quad (1)$$

The construction above corresponds to making \mathcal{M} into an affine space and endowing it with the *natural* flat connection induced by the linear structure of the space of scores, which then provide us with an affine coordinate system.

Now let \mathcal{A}_0 be the subspace of traceless operators in \mathcal{A} . Consider the -1 -embedding

$$\ell_{-1} : \mathcal{M} \rightarrow \mathcal{A} \quad \rho \mapsto \rho,$$

and define, at each $\rho \in \mathcal{M}$, the -1 -representation of tangent vectors as

$$(\ell_{-1})_{*(\rho)} : T_\rho \mathcal{M} \rightarrow \mathcal{A}_0 \quad v \mapsto (\ell_{-1} \circ \gamma)'(0),$$

where $\gamma : (-\varepsilon, \varepsilon) \rightarrow \mathcal{M}$ is again a curve in the equivalence class of the tangent vector v . In coordinates, the -1 -representation of the basis $\left\{ \frac{\partial}{\partial \theta^1} \Big|_\rho, \dots, \frac{\partial}{\partial \theta^n} \Big|_\rho \right\}$

of $T_\rho\mathcal{M}$ is $\left\{\frac{\partial\rho}{\partial\theta^1},\dots,\frac{\partial\rho}{\partial\theta^n}\right\}$. As before, the -1 -representation of a vector field X on \mathcal{M} is an \mathcal{A}_0 -valued function denoted by X^- .

We obtain the mixture or -1 -connection by defining the flat parallel transport

$$\tau_{\rho_0,\rho_1}^{(-1)} : T_{\rho_0}\mathcal{M} \rightarrow T_{\rho_1}\mathcal{M} \qquad v \mapsto (\ell_{-1})_{*(\rho_1)}^{-1} \left((\ell_{-1})_{*(\rho_0)} v \right).$$

Duality and the *BKM* Metric

Two connections ∇ and ∇^* on a Riemannian manifold (\mathcal{M}, g) are dual with respect to g if and only if

$$Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X^* Z), \tag{2}$$

for any vector fields X, Y, Z on \mathcal{M} [1]. Equivalently, if $\tau_{\gamma(t)}$ and $\tau_{\gamma(t)}^*$ are the respective parallel transports along a curve γ on \mathcal{M} , then ∇ and ∇^* are dual with respect to g if and only if

$$g(Y, Z) = g\left(\tau_{\gamma(t)} Y, \tau_{\gamma(t)}^* Z\right). \tag{3}$$

Given any connection ∇ on (\mathcal{M}, g) , we can always find a unique connection ∇^* such that ∇ and ∇^* are dual with respect to g . On the other hand, given two connections ∇ and ∇^* , we can ask what are the possible Riemannian metrics g with respect to which they are dual. In particular, we want explore this question for the case of the exponential and mixture connections on a manifold of density matrices.

Another concept of duality [1] is that of dual coordinate systems, regardless of any connection. Two coordinate systems $\theta = (\theta^i)$ and $\eta = (\eta_i)$ on a Riemannian manifold (\mathcal{M}, g) are dual with respect to g if and only if their natural bases for $T_p\mathcal{M}$ are *biorthogonal* at every point $p \in \mathcal{M}$, that is, $g(\partial/\partial\theta^i, \partial/\partial\eta_j) = \delta_j^i$. Equivalently, $\theta = (\theta^i)$ and $\eta = (\eta_i)$ are dual with respect to g if and only if $g_{ij} = \partial\eta_i/\partial\theta^j$ and $g^{ij} = \partial\theta_i/\partial\eta^j$, at every point $p \in \mathcal{M}$, where, as usual, $g^{ij} = (g_{ij})^{-1}$.

The next theorem gives a characterisation of dual coordinate systems in terms of potential functions, thus introducing convexity theory and the related duality with respect to Legendre transforms.

Theorem 4 (Amari, 1985) *When a Riemannian manifold (\mathcal{M}, g) has a pair of dual coordinate systems (θ, η) , there exist potential functions $\Psi(\theta)$*

and $\Phi(\eta)$ such that

$$g_{ij}(\theta) = \frac{\partial^2 \Psi(\theta)}{\partial \theta^i \partial \theta^j} \quad \text{and} \quad g^{ij} = \frac{\partial^2 \Phi(\eta)}{\partial \eta_i \partial \eta_j}.$$

Conversely, when either potential function Ψ or Φ exists from which the metric is derived by differentiating it twice, there exist a pair of dual coordinate systems. The dual coordinate systems and the potential functions are related by the following Legendre transforms

$$\theta^i = \frac{\partial \Phi(\eta)}{\partial \eta_i}, \quad \eta_i = \frac{\partial \Psi(\theta)}{\partial \theta^i}$$

and

$$\Psi(\theta) + \Phi(\eta) - \theta^i \eta_i = 0$$

In contrast to the case of dual connections, dual coordinate systems do not necessarily exist on every Riemannian manifold [1]. When the additional property of flatness is required, the following theorem provides a link between the two concepts of duality. We say that a connection ∇ on manifold \mathcal{M} is flat if \mathcal{M} admits a global ∇ -affine coordinate system. This is equivalent to its curvature and torsion both being zero.

Theorem 5 (Amari) *Suppose that ∇ and ∇^* are two flat connections on a manifold \mathcal{M} . If they are dual with respect to a Riemannian metric g on \mathcal{M} , then there exists a pair (θ, η) of dual coordinate systems such that θ is ∇ -affine and η is a ∇^* -affine.*

We now consider the uniqueness of the Riemannian metric on \mathcal{M} . Using either the 1 or the -1 representation of the tangent bundle $T\mathcal{M}$, we define a Riemannian metric on \mathcal{M} by a smooth assignment of an inner product $\langle \cdot, \cdot \rangle_\rho$ in $\mathcal{A} \subset B(\mathcal{H}^N)$ for each point $\rho \in \mathcal{M}$. If $A^{(+1)}, B^{(+1)}$ and $A^{(-1)}, B^{(-1)}$ are, respectively, the 1 and -1 representations of $A, B \in T_\rho \mathcal{M}$, then the *BKM* metric is

$$g_\rho^B(A, B) = \text{Tr} \left(A^{(-1)} B^{(+1)} \right) = \int_0^\infty \text{Tr} [(\rho + \alpha)^{-1} A^-(\rho + \alpha)^{-1} B^+] d\alpha. \quad (6)$$

The (1) and (-1) connections are dual with respect to the *BKM* metric [6, 5]. A natural question is whether it is the only metric with this property. The next theorem tells us the consequences of duality alone.

Theorem 7 *If the connections $\nabla^{(1)}$ and $\nabla^{(-1)}$ are dual with respect to a Riemannian metric g on \mathcal{M} , then there exists a constant $n \times n$ matrix M , such that $(g_\rho)_{ij} = \sum_{k=1}^n M_{ik}(g_\rho^B)_{kj}$.*

Proof: Since the two connections are flat, by theorem 5, there exist dual coordinate systems (θ, η) such that θ is $\nabla^{(1)}$ -affine and η is $\nabla^{(-1)}$ -affine. Thus, applying theorem 4, there exist a potential function $\Psi(\theta)$ such that

$$g_{ij}(\theta) = \frac{\partial^2 \Psi(\theta)}{\partial \theta^i \partial \theta^j} \quad \text{and} \quad \eta_i = \frac{\partial \Psi(\theta)}{\partial \theta^i}.$$

Now since θ is $\nabla^{(1)}$ -affine, there exist linearly independent operators $\{1, X_1, \dots, X_n\}$ such that

$$\rho = \exp\left(\theta^1 X_1 + \dots + \theta^n X_n - \tilde{\Psi}(\theta)\right), \quad (8)$$

where $\tilde{\Psi} = \log \text{Tr} \exp(\theta^i X_i)$ is the free energy. Any such set of operators defines a $\nabla^{(-1)}$ -affine coordinate system through the formula $\tilde{\eta}_i = \text{Tr}(\rho X_i)$. Differentiating $\tilde{\Psi}$ with respect to θ^i we obtain

$$\frac{\partial \tilde{\Psi}(\theta)}{\partial \theta^i} = \text{Tr}(\rho X_i) = \tilde{\eta}_i.$$

Thus $\tilde{\eta}_i = \frac{\partial \tilde{\Psi}(\theta)}{\partial \theta^i}$ and $\eta_i = \frac{\partial \Psi(\theta)}{\partial \theta^i}$ are two $\nabla^{(-1)}$ -affine coordinate systems, so they must be related by an affine transformation. So there exist an $n \times n$ matrix M and numbers (a_1, \dots, a_n) such that

$$\eta_i = \sum_{k=1}^n M_{ik} \tilde{\eta}_k + a_i,$$

that is,

$$\frac{\partial \Psi(\theta)}{\partial \theta^i} = \sum_{k=1}^n M_{ik} \frac{\partial \tilde{\Psi}(\theta)}{\partial \theta^k} + a_i,$$

and differentiating this equation with respect to θ^j gives

$$g_{ij}(\theta) = \frac{\partial^2 \Psi(\theta)}{\partial \theta^i \partial \theta^j} = \sum_{k=1}^n M_{ik} \frac{\partial^2 \tilde{\Psi}(\theta)}{\partial \theta^j \partial \theta^k} = \sum_{k=1}^n M_{ik} g_{kj}^B. \quad (9)$$

The Condition of Monotonicity

We say [3] that a metric g on \mathcal{A}_0 (the (-1) representation) is monotone iff

$$g_{S\rho}(SA^-, SA^-) \leq g_\rho(A^-, A^-) \quad (10)$$

for every $\rho \in \mathcal{M}$, $A \in T_\rho\mathcal{M}$, and every completely positive, trace preserving map $S : \mathcal{A} \rightarrow \mathcal{A}$.

Let $\widehat{\mathcal{M}}$ be the manifold of faithful weights (the positive-definite matrices). We can extend g and g^B from $T\mathcal{M}$ to $T\widehat{\mathcal{M}}$ as follows. At $\rho \in \mathcal{M}$ and $A \in T\widehat{\mathcal{M}}$, put $A = A_0\rho + A^-$, where $A_0 = \text{Tr } A - 1$ and $\text{Tr } A^- = 0$. Then put

$$\hat{g}_\rho(A, B) = A_0B_0 + g_\rho(A^-, B^-). \quad (11)$$

For g^B this extension coincides with that given by eq. (6). Then, if g is monotone on $T\mathcal{M}$, its extension is monotone on $T\widehat{\mathcal{M}}$: let S be a trace-preserving CP map on $T\widehat{\mathcal{M}}$; then

$$\widehat{g}_{S\rho}(SA, SA) = A_0^2 + g_{S\rho}(SA^-, SA^-) \leq A_0^2 + g_\rho(A^-, A^-) = \widehat{g}_\rho(A, A).$$

For any metric g on $T\widehat{\mathcal{M}}$, and putting $A^{(-1)} = (A_0, A^-)$, we define the positive operator K_ρ by

$$g_\rho(A, B) = \left\langle A^{(-1)}, K_\rho \left(B^{(-1)} \right) \right\rangle_{HS} = \text{Tr} \left(A^{(-1)} K_\rho \left(B^{(-1)} \right) \right), \quad (12)$$

acting on $T\widehat{\mathcal{M}}$ furnished with the Hilbert-Schmidt scalar product. Petz uses the following operators on $T\widehat{\mathcal{M}}$: $L_\rho A = \rho A$, $R_\rho A = A\rho$, and has [7, 9, 10] given a characterisation of monotone metrics on matrix spaces in terms of operator monotone functions. He proved

Theorem 13 (Petz, 1996) *A Riemannian metric g on $\widehat{\mathcal{M}}$ is monotone if and only if*

$$K_\rho = \left(R_\rho^{1/2} f(L_\rho R_\rho^{-1}) R_\rho^{1/2} \right)^{-1},$$

where K_ρ is defined in (12) and $f : R^+ \rightarrow R^+$ is an operator monotone function satisfying $f(t) = t f(t^{-1})$.

In particular, the *BKM* metric is monotone and its corresponding operator monotone function is $f^B(t) = (t - 1)/\log t$. Combining this characterisation with theorem (7), we obtain the following uniqueness result.

Theorem 14 *If the connections $\nabla^{(1)}$ and $\nabla^{(-1)}$ are dual with respect to a monotone Riemannian metric g on \mathcal{M} , then g is a constant multiple of the BKM metric.*

Proof: Let K_ρ^g and K_ρ^B be the operators associated with the monotone metrics g and g^B as in equation (12). Let us extend the matrix M of theorem (7) from $T\mathcal{M}$ to a matrix on $\widehat{\mathcal{M}}$ by $M_{0,k} = 0 = M_{k,0}$, $M_{0,0} = 1$. Then eq. (9) gives the following relation between the extended metrics \hat{g} and \hat{g}^B , in the coordinates $(\theta^0, \theta^1, \dots, \theta^n)$:

$$\hat{g}_\rho(\partial_i, \partial_j) = \sum_{k=0}^n \hat{M}_{ik} \hat{g}_\rho^B(\partial_k, \partial_j). \quad (15)$$

In terms of the kernels, this gives $K_\rho^g = \hat{M} K_\rho^B$. But then Petz's formula for K in theorem (13) leads to $\hat{M} = f^g(L_\rho R_\rho^{-1})^{-1} f^B(L_\rho R_\rho^{-1})$, where f^g and f^B are the operator-monotone functions corresponding to g and g^B respectively. Thus the matrix \hat{M} is a certain function of the operators L_ρ and R_ρ , but is independent of the point ρ . We conclude that it must be a constant multiple of the identity matrix.

References

- [1] S.-I. Amari, **Differential Geometric Methods in Statistics**, *Lecture Notes in Statistics*, **28**, Springer-Verlag, New York, 1985.
- [2] S.-I. Amari, *Information Geometry*, *Contemp. Math.*, **203**, 1997.
- [3] N. N. Čencov, **Statistical Decision Rules and Optimal Inferences**, *Translations of Mathematical Monographs*, American Mathematical Society, Providence, 1982.
- [4] M. R. Grasselli and R. F. Streater, *The Quantum Information Manifold for ε -bounded Forms*, to appear in *Rep. Math. Phys.*, math-hp/9910031.
- [5] H. Hasegawa, *Exponential and mixture families in quantum statistics: dual structures and unbiased parameter estimation.*, *Rep. Math. Phys.*, **39**, 49-68, 1997.
- [6] H. Nagaoka, *Differential Geometric Aspects of Quantum State Estimation and Relative Entropy*, in **Quantum Communication and Measurement**, eds. V. P. Balavkin, O. Hirota and R. L. Hudson, Plenum Press, 1995.

- [7] D. Petz, *Quasi-entropies for Finite Quantum Systems*, Rep. Math. Phys., **23**, 57-65, 1986.
- [8] D. Petz, *Geometry of Canonical Correlation on the State Space of a Quantum System*, J. Math. Phys., **35**, 780-795, 1994.
- [9] D. Petz, *Monotone Metrics on Matrix Spaces*, Lin. Alg. Appl., **244**, 81-96, 1996.
- [10] D. Petz and C. Sudar, *Geometries of Quantum States*, J. Math. Phys., **37**, 2662-2673, 1996.