frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 07 February 2023
DOI 10.3389/fphy.2023.1042257

OPEN ACCESS

EDITED BY
Florio M. Ciaglia,
Universidad Carlos III de Madrid de Madrid,
Spain

REVIEWED BY
Sorin Dragomir,
University of Basilicata, Italy
Fabio Di Cosmo,
Universidad Carlos III de Madrid de Madrid,
Spain

*CORRESPONDENCE
Jan Naudts,
jan.naudts@uantwerpen.be

SPECIALTY SECTION
This article was submitted to Statistical and Computational Physics,
a section of the journal
Frontiers in Physics

RECEIVED 12 September 2022
ACCEPTED 06 January 2023
PUBLISHED 07 February 2023

CITATION
Naudts J (2023), Exponential arcs in manifolds of quantum states.
Front. Phys. 11:1042257.
doi: 10.3389/fphy.2023.1042257

Exponential arcs in manifolds of quantum states

Jan Naudts*

Physics Department, Universiteit Antwerpen, Antwerp, Belgium

The manifold under consideration consists of the faithful normal states on a sigma-finite von Neumann algebra in standard form. Tangent planes and approximate tangent planes are discussed. A relative entropy/divergence function is assumed to be given. It is used to generalize the notion of an exponential arc connecting one state to another. The generator of the exponential arc is shown to be unique up to an additive constant. In the case of Araki's relative entropy, every self-adjoint element of the von Neumann algebra generates an exponential arc. The generators of the composed exponential arcs are shown to add up. The metric derived from Araki's relative entropy is shown to reproduce the Kubo–Mori metric. The latter is the metric used in linear response theory. The e- and m-connections describe a dual pair of geometries. Any finite number of linearly independent generators determines a submanifold of states connected to a given reference state by an exponential arc. Such a submanifold is a quantum generalization of a dually flat statistical manifold.

KEYWORDS

exponential arcs, quantum statistical manifold, quantum divergence function, Araki's relative entropy, dually flat geometry, Tomita–Takesaki theory, linear response theory, Kubo–Mori theory

1 Introduction

The goal of the present paper is to show that the theory of quantum statistical manifolds can be formulated without reference to density matrices. It is tradition to describe the statistical state of a quantum model by a density matrix. In many cases this suffices, in particular when the Hilbert space of wave functions is finite-dimensional. However, even simple models such as the quantum harmonic oscillator or the hydrogen atom require an infinite-dimensional Hilbert space. This involves handling of unbounded operators which cause considerable technical complications. These complications are avoided in the present work.

A one-to-one correspondence between density matrices and quantum states is usually accepted. The quantum states form the sample space of the statistical description. An alternative description emerged in the past century, which introduced the notion of a mathematical state on an algebra of observables which can be realized as an algebra of bounded operators on Hilbert space. See for instance [1–5].

Equilibrium states of quantum statistical mechanics are described by the quantum analogue of the probability distribution of Gibbs, which is a density matrix ρ of the form

$$\rho = \frac{1}{Z} e^{-\beta H},$$

with H a Hermitian matrix, β a parameter the inverse temperature, and Z a function of β used to normalize density matrix ρ so that its trace equals 1. Models described in this way can belong to a quantum exponential family. They possess an intriguing property called the Kubo–Martin–Schwinger (KMS) condition [6]. The KMS condition describes a symmetry property of the time evolution of quantum states. This symmetry coincides with the symmetry between left and right multiplication of operators, which is studied in the Tomita–Takesaki theory [7]. [5] can be used as a reference text for this theory.

The notion of a statistical manifold is studied in information geometry ([8–12]). It is a manifold of probability distributions. The quantum analogue is described in Chapter 7 of [11] as a manifold of k by k density matrices. The book of Petz [13] reviews several aspects of quantum statistics, including the basics of quantum information and quantum information geometry.

The generalization of Amari's dually flat geometry from statistical models with a finite number of parameters to Banach manifolds of mutually equivalent probability measures started with the work of [14]. Non-commutative versions were formulated by [15–19].

The convex set $\mathbb{M}$ of faithful normal states on a σ-finite von Neumann algebra is in general *not* a Banach manifold. The point of view taken in the present work is that the set $\mathbb{M}$ should, by definition, be a quantum statistical manifold. This raises the question of how to transfer common notions of differential geometry and of Banach manifolds to this quantum setting. The present work contributes to this effort.

The relative entropy of Umegaki [20] is the starting point to implement Amari's dually flat geometry on the quantum manifold. It should be noted that relative entropy is called a divergence function in mathematical literature. Araki [21–23] generalizes Umegaki's relative entropy to the context of mathematical states on an algebra of bounded operators on a Hilbert space. The use of Araki's relative entropy replacing that of Umegaki's is the core of the present work.

Exponential arcs were introduced in [24, 25] and used in [26]. These arcs can be considered one-parameter exponential families embedded in the manifold. The maximal exponential model centered at a given probability distribution p equals the set of all probability distributions connected to p by an open exponential arc. Exponential arcs were studied in the quantum setting by [27]. Here, the definition is generalized. The exponential arcs are used to define quantum statistical manifolds as submanifolds of the manifold of all quantum states.

The Radon–Nikodym Theorem plays an important role in probability theory. For each measure absolutely continuous with respect to the reference measure, there exists an essentially unique probability distribution function. The problem that arises in the non-commutative context is the non-uniqueness of the Radon–Nikodym derivative. This leads to different definitions of the relative entropy and of the exponential arcs. First attempts to reformulate the theory of the quantum statistical manifold in terms of states on a C^*-algebra are found in [28,29] and in [27]. These two approaches differ in the choice of the Radon–Nikodym derivative. In the present work, the definition of an exponential arc is generalized so that it depends explicitly on the choice of relative entropy and in that way on the choice of the Radon–Nikodym derivative.

The alternative approach of [30] relies on the Lie Theory for the group of bounded operators with bounded inverse. The state space is partitioned into the disjoint union of the orbits of an action of the Lie group. Under mild conditions, it is shown that the orbits are Banach manifolds. The restriction to bounded operators implies that the orbits do not connect quasi-equivalent states when the Radon–Nikodym derivatives are unbounded operators.

Sections 2–4 give a short introduction on KMS states, on the theory of the modular operator, and on positive cones. Section 5 gives a definition of the manifold $\mathbb{M}$ under study as the convex set of faithful normal states on a sigma-finite von Neumann algebra. The tangent space consists of linear functionals on the algebra. Its extent depends on the chosen topology, and it is not obvious how to find a good compromise. Therefore, the notion of approximate tangent vectors is considered in Section 6.

A dense subset of the manifold $\mathbb{M}$ consists of states majorized by a multiple of the reference state. This subset of states is mentioned in Section 7 because it is easier to handle.

Section 8 gives a new definition of exponential arcs. It generalizes existing concepts and is broad enough to cover different approaches. The definition depends on the choice of a relative entropy/divergence function. Such an exponential arc can be seen as a one-dimensional sub-manifold and as a straightforward example of a quantum statistical manifold. Duality properties well-known for models of information geometry are elaborated in Section 9.

The important example of the algebra of n-by-n matrices is considered in Section 10.

Starting with Section 11 the paper specializes to the case of Araki's relative entropy. It is shown in Section 13 that each self-adjoint element h of the von Neumann algebra defines an exponential arc defined relative to Araki's relative entropy and starting at the reference state ω. The initial derivative of the arc exists as a Fréchet derivative and belongs to the tangent plane $T_\omega\mathbb{M}$. The inner product between two such tangent vectors reproduces the metric which is used in the Kubo–Mori Theory of linear response. This is shown in Section 14. The exponential arcs are shown to be geodesics for the e-connection which is, by definition, the dual of the m-connection.

Section 16 applies the results obtained so far to show that manifolds generated by a finite number of exponential arcs have the properties one expects from a quantum statistical manifold.

A few points of concern are discussed in the final Section 17.

2 KMS states

Equilibrium states of quantum statistical mechanics satisfy the KMS condition. In the GNS representation, an equilibrium state becomes a faithful state on a σ-finite von Neumann algebra of operators on a complex Hilbert space. The state is defined by a normalized cyclic and separating vector in the Hilbert space.

The state of a model of statistical physics can be described by a mathematical state on a C^*-algebra $\mathfrak{A}$. It can be represented by a normalized vector Ω (a wave function) in a Hilbert space $\mathcal{H}$. This is known as the GNS (Gelfand–Naimark–Segal) representation theorem. Observable quantities are represented by self-adjoint operators on $\mathcal{H}$. The quantum expectation $\langle x\rangle$ of operator x is then given by

$$\langle x\rangle = (x\Omega, \Omega), \tag{1}$$

with in the right-hand side the scalar product of the two vectors $x\Omega$ and Ω. It should be noted that the mathematical convention is followed that the scalar product (inner product) is linear in its first argument and conjugate-linear in the second argument. In Dirac's *bra-ket* notation, it reads

$$\langle x\rangle = \langle\Omega|x\Omega\rangle.$$

For convenience, one works with a von Neumann algebra $\mathfrak{M}$ of bounded operators on the Hilbert space $\mathcal{H}$. Observables of interest, when unbounded, are represented by operators affiliated with $\mathfrak{M}$. The state ω on the C^*-algebra extends to a vector state on $\mathfrak{M}$ again denoted ω. It is given by

$$\omega(x) = (x\Omega, \Omega), \qquad x \in \mathfrak{M}.$$

The vector Ω is cyclic for $\mathfrak{M}$, which means that the subspace $\mathfrak{M}\Omega$ is dense in the Hilbert space $\mathcal{H}$. It is also assumed in what follows that the state ω is faithful, i.e., $\omega(x^*x) = 0$ implies $x = 0$. This implies that Ω is a separating vector for $\mathfrak{M}$, i.e., $x\Omega = 0$ implies $x = 0$ for any x in $\mathfrak{M}$, and hence it is a cyclic vector for the commutant $\mathfrak{M}'$ of $\mathfrak{M}$, the algebra of all operators commuting with all of $\mathfrak{M}$.

Equilibrium states of statistical mechanics are characterized by the KMS (Kubo–Martin–Schwinger) condition [6]. Roughly speaking, this condition states that the quantum time evolution of the model has an analytic extension into the complex plane. This is made more precise in what follows.

The time evolution is described by a strongly continuous one-parameter group $t \in \mathbb{R} \mapsto u_t$ of unitary operators which leave the algebra $\mathfrak{M}$ unchanged, i.e., $x \in \mathfrak{M}$ implies that $x_t = u_t^* x u_t$ belongs to $\mathfrak{M}$ for all t. The operators u_t are determined by a self-adjoint operator H

$$u_t = e^{-itH},$$

which is the generator of the time evolution in the GNS representation. The time derivative of x_t satisfies

$$i\frac{\mathrm{d}}{\mathrm{d}t}x_t = [x_t, H]_-.$$

This equation has the same form as Heisenberg's equation of motion.

The KMS condition requires that for any pair x, y of operators in $\mathfrak{M}$, there exists a complex function $F(w)$, defined and continuous on the strip $-\beta \le \mathrm{Im}\ w \le 0$ and analytical inside with boundary values

$$F(t) = (x_t y\Omega, \Omega) \quad \text{and} \quad F(t - i\beta) = (y x_t \Omega, \Omega), \qquad t \in \mathbb{R}.$$

In the mathematics literature, the parameter β, which is the inverse temperature of the model, is usually taken equal to 1 or -1.

An immediate consequence of the KMS condition being satisfied is that the state ω is invariant. Indeed, take y equal to the identity operator. Then, one has $F(t—i\beta) = F(t)$ for all t in $\mathbb{R}$. If in addition, x is self-adjoint, then $F(t)$ is a real function. From the Schwarz reflection principle, one then concludes that $F(w)$ is a constant function. This implies $\omega(x_t) = \omega(x)$ for all self-adjoint x and hence for all x. The GNS theorem then guarantees that the vector Ω can be taken to be invariant, i.e., $u_t\Omega = \Omega$ for all t.

3 The modular operator

The quantum-mechanical time evolution coincides with the modular automorphism group of Tomita–Takesaki theory.

The KMS condition, when satisfied, expresses a symmetry which is present in the context of non-commuting operators. The symmetry is the inversion of the order of multiplication of operators. In non-commutative groups, the modular function links left and right Haar measures. The analogue in functional analysis is studied in the theory of the modular operator, also called the Tomita–Takesaki theory [7].

The operator $e^{-\beta H}$ with H the generator of the quantum-time evolution is traditionally denoted as Δ_Ω. It is the modular operator of the Tomita–Takesaki theory. It is in general an unbounded operator such that $\mathfrak{M}\Omega$ is in the domain of the definition of the square root $\Delta_\Omega^{1/2}$ of Δ_Ω. Hence, the expression

$$F(w) = \left(x\Delta_\Omega^{iw/\beta} y\Omega, \Omega\right), \qquad x, y \in \mathfrak{M}, \tag{2}$$

is well-defined for $0 \ge \mathrm{Im}\ w \ge -\beta/2$. The other half of the strip $0 \ge \mathrm{Im}\ w \ge -\beta$ is covered by the Schwarz reflection principle. Indeed, if x and y are self-adjoint, then one can show with the Tomita–Takesaki theory that the map $t \mapsto F(t - i\beta/2)$ is a real function. Hence, the principle can be applied to obtain $F(w) = \overline{F(w - i\beta)}$.

The unitary time evolution operator u_t can be written as

$$u_t = \Delta_\Omega^{it/\beta}.$$

The time evolution of an operator x in the Heisenberg picture is then given by

$$x_t = \tau_t^\Omega x = \Delta_\Omega^{-it/\beta} x \Delta_\Omega^{it/\beta}.$$

The action $t \mapsto \tau_t^\Omega$ of the group $\mathbb{R}, +$ is called the modular automorphism group.

The modular conjugation operator J of the Tomita–Takesaki Theory represents the symmetry which is at the basis of the theory. It is a conjugate-linear operator satisfying $J = J^*$ and $J^2 = \mathbb{I}$. Operator x belongs to the von Neumann algebra $\mathfrak{M}$ if and only if JxJ belongs to the commutant algebra $\mathfrak{M}'$. The latter is the space of operators commuting with all operators in $\mathfrak{M}$. The product $J\Delta_\Omega^{1/2}$ is denoted as S_Ω and has the property of

$$S_\Omega x\Omega = x^*\Omega, \qquad x \in \mathfrak{M}.$$

4 Dual cones

The natural positive cone $\mathcal{P}_\Omega$ is needed in subsequent sections. One reason for making use of it is that there exists a one-to-one correspondence between normal states on $\mathfrak{M}$ and normalized vectors in $\mathcal{P}_\Omega$.

Section 4 of [22]introduces the cones V_Ω^α, $0 \le \alpha \le 1/2$, of the vectors in $\mathcal{H}$. The self-dual cone $V_\Omega^{1/4}$ is called the natural positive cone and is denoted as $\mathcal{P}_\Omega$.

By definition, V_Ω^α is the closure of the cone

$$\{\Delta_\Omega^\alpha x\Omega\colon\ x \in \mathfrak{M}, x \ge 0\}.$$

The cone $V_\Omega^{1/2}$ is used in [27] to introduce exponential arcs. It is equal to the closure of the set

$$\{y\Omega\colon\ y \in \mathfrak{M}', y \ge 0\}.$$

To see this note that

$$\begin{aligned}\Delta_\Omega^{1/2} x\Omega &= JS_\Omega x\Omega\\ &= Jx^*\Omega\\ &= y\Omega\end{aligned}$$

with $y = Jx^*J$. The latter is an arbitrary element of the commutant $\mathfrak{M}'$.

The following characterization of the natural positive cone $\mathcal{P}_\Omega$ is found in Section 2.5 of [5].

Proposition 1: The cone $\mathcal{P}_\Omega = V_\Omega^{1/4}$ equals the closure of the set of vectors

$$\{xJx\Omega:\ x \in \mathfrak{M}\}. \tag{3}$$

This result can be understood as follows. Take Φ in $\mathcal{P}$ of the form (3), i.e., $\Phi = xJx\Omega$ with x in $\mathfrak{M}$. Let

$$y = \Delta_\Omega^{-1/4} x \Delta_\Omega^{1/4}. \tag{4}$$

This expression can be inverted to

$$x = \Delta_\Omega^{1/4} y \Delta_\Omega^{-1/4}$$

so that

$$\begin{aligned}\Phi = xJx\Omega &= \Delta_\Omega^{1/4} y \Delta_\Omega^{-1/4} J \Delta_\Omega^{1/4} y \Delta_\Omega^{-1/4}\Omega\\ &= \Delta_\Omega^{1/4} y J \Delta_\Omega^{1/2} y\Omega\\ &= \Delta_\Omega^{1/4} y S_\Omega y\Omega\\ &= \Delta_\Omega^{1/4} y y^*\Omega.\end{aligned}$$

Assume now that one could prove that the operator y defined by (4) belongs to $\mathfrak{M}$; then, the above calculation would show that Φ is of the form $\Phi = \Delta_\Omega^{1/4} a\Omega$ with $a = yy^*$ a positive element of $\mathfrak{M}$. The actual proof of the proposition uses that $\tau_t^\Omega x = \Delta_\Omega^{-it} x \Delta_\Omega^{it}$ belongs to $\mathfrak{M}$.

The cone $\mathcal{P}_\Omega$ is independent [22] of the choice of the cyclic and separating vector Ω in $\mathcal{P}_\Omega$, and the isometry J is the same for all these choices. For this reason, it is said to be universal.

From (3), it is easy to see that each vector in $\mathcal{P}_\Omega$ is an eigenvector with eigenvalue 1 of the modular conjugation operator J. Indeed, one has

$$xJx\Omega = x(JxJ)\Omega = (JxJ)x\Omega = J(xJx\Omega).$$

Here, use is made of $J\Omega = \Omega$ and the fact that the operators x and JxJ commute with each other.

5 A manifold of quantum states

A manifold $\mathbb{M}$ of vector states on the von Neumann algebra $\mathcal{M}$ is defined. Tangent vector fields are Fréchet derivatives of paths in $\mathbb{M}$.

Introduce the notation ω_Φ for the vector state defined by the normalized vector Φ in $\mathcal{H}$. It is given by

$$\omega_\Phi(x) = (x\Phi, \Phi), \qquad x \in \mathfrak{M}.$$

A manifold $\mathbb{M}$ of states on the von Neumann algebra $\mathfrak{M}$ is defined by

$$\mathbb{M} = \{\omega_\Phi:\ \Phi \in \mathcal{P}_\Omega,\ \text{normalized, cyclic and separating for } \mathfrak{M}\}.$$

The equilibrium state $\omega = \omega_\Omega$ is taken as a reference point in $\mathbb{M}$. The subset $\mathcal{P}_\Omega$ of $\mathcal{H}$ is the natural positive cone introduced in the previous section.

The topology on the manifold $\mathbb{M}$ is that of the operator norm. One has

$$\|\omega_\Phi - \omega_\Psi\| = \sup\{|\omega_\Phi(x) - \omega_\Psi(x)|:\ x \in \mathfrak{M}, \|x\| \le 1\}.$$

Several topologies can be defined on the algebra $\mathfrak{M}$. Particularly relevant is the σ-weak topology. For what follows, it is important to know that in the present context, a state ω on $\mathfrak{M}$ is said to be normal if and only if it is σ-weakly continuous and if and only if it is a vector state. See for instance, Theorems 2.4.21 and 2.5.31 of [5].

Any tangent vector is a σ-weakly continuous linear functional on the von Neumann algebra $\mathfrak{M}$. Let $t \mapsto \gamma_t$ be a Fréchet differentiable map defined on an open interval of $\mathbb{R}$ with values in the manifold $\mathbb{M}$. The derivative

$$\dot\gamma_t = \frac{\mathrm{d}}{\mathrm{d}t}\gamma_t$$

is required to exist as a Fréchet derivative, i.e., it satisfies

$$\|\gamma_t - \gamma_s - (t-s)\dot\gamma_t\| = \mathrm{o}(t-s).$$

From the normalization, $\gamma_t(1) = 1$ for all t in the domain of the map, one obtains $\dot\gamma_t(1) = 0$. From $\gamma_t(x^*) = \overline{\gamma_t(x)}$, one obtains $\dot\gamma_t(x^*) = \overline{\dot\gamma_t(x)}$. Hence, the linear functional $\dot\gamma_t$ is Hermitian.

There are several ways to define the tangent space $T_\omega\mathbb{M}$ at the point ω in $\mathbb{M}$. Intuitively, a tangent vector is a derivative, defined in some sense, of a path $t \mapsto \gamma_t$ in $\mathbb{M}$ passing through the point ω. The states of the manifold $\mathbb{M}$ belong to the space $\mathfrak{M}_*$ of all σ-weakly continuous linear functionals on the algebra $\mathfrak{M}$ (see Proposition 2.4.18 of [5]). Hence, one expects that tangent vectors belong to $\mathfrak{M}_*$ as well.

In this section, the requirement is made that the path $t \mapsto \gamma_t$ is Fréchet-differentiable. This may be too restrictive. In what follows, we adopt the definition that the tangent space $T_\omega\mathbb{M}$ consists of all Hermitian χ in $\mathfrak{M}_*$, satisfying $\chi(1) = 0$. It should be noted that it is well-possible that for certain elements χ of this space, there is no smooth curve passing through ω with the property that the derivative at ω equals χ.

6 Approximate tangents

Approximate tangent vectors can be defined in an intrinsic manner.

An alternative definition of the tangent space starts from the following observation.

Proposition 2: The set $\mathcal{T}_\omega$ defined by

$$\mathcal{T}_\omega = \left\{\lambda(\phi - \psi):\ \phi, \psi \in \mathbb{M}, \lambda \in \mathbb{R} \text{ and } \omega = \frac{1}{2}(\phi + \psi)\right\}.$$

is a linear subspace of the tangent space $T_\omega\mathbb{M}$.

Proof:

Let ϕ and ψ be two states in $\mathbb{M}$ such that $\omega = \frac{1}{2}(\phi + \psi)$. Construct a Fréchet-differentiable path γ by

$$\gamma_t = (1-t)\psi + t\phi, \qquad t \in (0,1).$$

The state γ_t belongs to the manifold $\mathbb{M}$ because the latter is a convex set. In particular, one has $\omega = \gamma_{1/2}$ and $\phi - \psi = \dot\gamma_{1/2}$ is a tangent vector. This shows that $\phi - \psi$ and hence also $\lambda(\phi - \psi)$ belongs to $T_\omega\mathbb{M}$. One concludes that $\mathcal{T}_\omega \subset T_\omega\mathbb{M}$.

Assume now that $\lambda(\phi - \psi)$ and $\lambda'(\phi' - \psi')$ both belong to $\mathcal{T}_\omega$. We have to show that

$$\lambda(\phi - \psi) + \lambda'(\phi' - \psi')$$

belongs to $\mathcal{T}_\omega$. If $\lambda = 0$ or $\lambda' = 0$, then the claim is clearly satisfied. Without restriction, assume $\lambda = 1$.

If $\lambda' > 0$, then choose

$$\phi'' = \frac{1}{1+\lambda'}(\phi + \lambda'\phi') \quad \text{and} \quad \psi'' = \frac{1}{1+\lambda'}(\psi + \lambda'\psi').$$

Both ϕ'' and ψ'' belong to $\mathbb{M}$ because the latter is a convex set. One verifies that $\phi'' + \psi'' = 2\omega$ and

$$(1+\lambda')(\phi''-\psi'') = \phi - \psi + \lambda'(\phi'-\psi').$$

This shows that the latter sum belongs to $\mathcal{T}_\omega$.

In the case that $\lambda' < 0$, one chooses

$$\phi'' = \frac{1}{1-\lambda'}(\phi - \lambda'\psi') \quad \text{and} \quad \psi'' = \frac{1}{1-\lambda'}(\psi - \lambda'\phi')$$

to reach the same conclusion. This finishes the proof that $\mathcal{T}_\omega$ is a linear subspace of $T_\omega\mathbb{M}$.

We introduce the notations

$$\mathcal{R}_{\omega,\epsilon} = \bigcup \mathcal{T}_\phi \text{ with } \phi \in \mathbb{M} \text{ such that } \|\phi-\omega\| < \epsilon.$$

and

$$\mathcal{T}_\omega^{\text{approx}} = \bigcap_{\epsilon>0} \overline{\mathcal{R}_{\omega,\epsilon}}.$$

The construction of $\mathcal{T}_\omega^{\text{approx}}$ is analogous to the construction of the approximate tangent space in Chapter 3 of [31]. Clearly, $\mathcal{T}_\omega \subset \mathcal{T}_\omega^{\text{approx}}$. Further properties are derived below.

Proposition 3: If γ is a Fréchet-differentiable path in $\mathbb{M}$, then $\dot\gamma_t$ belongs to $\mathcal{T}_\omega^{\text{approx}}$ with $\omega = \gamma_t$.

Proof:

Let γ be a Fréchet-differentiable path in $\mathbb{M}$. Without restriction of generality, assume that $\gamma_0 = \omega$. For any $\epsilon > 0$ and $\delta > 0$, there exists $t \neq 0$ such that

$$\|(\gamma_t + \gamma_{-t})/2 - \gamma_0\| < \epsilon.$$

and

$$\|(\gamma_t - \gamma_{-t})\big/2t - \dot\gamma_0\| < \delta. \tag{5}$$

Then, ϕ defined by

$$\phi = \frac{1}{2}(\gamma_t + \gamma_{-t})$$

satisfies $\|\phi—\omega\| < \epsilon$, and $(\gamma_t—\gamma_{-t})/2t$ belongs to $\mathcal{R}_{\omega,\epsilon}$. Hence, (5) shows that the tangent vector $\dot\gamma_0$ belongs to the closure of $\mathcal{R}_{\omega,\epsilon}$. Because $\epsilon > 0$ is arbitrary, it also belongs to the intersection, which is $\mathcal{T}_\omega^{\text{approx}}$.

Lemma 1: $\mathcal{R}_{\omega,\epsilon}$ is a linear subspace of $T_\omega\mathbb{M}$.

Proof:

Take χ and ξ in $\mathcal{R}_{\omega,\epsilon}$. There exist ϕ and ψ in $\mathbb{M}$ such that $\chi \in \mathcal{T}_\phi$ and $\xi \in \mathcal{T}_\psi$ with $\|\phi—\omega\| < \epsilon$ and $\|\psi—\omega\| < \epsilon$. Therefore, there exist real λ, μ and states ϕ_1, Φ_2, Ψ_1, Ψ_2 in $\mathbb{M}$ such that

$$\chi = \lambda(\phi_1 - \phi_2) \quad \text{and} \quad \phi = \frac{1}{2}(\phi_1 + \phi_2)$$

and

$$\xi = \mu(\psi_1 - \psi_2) \quad \text{and} \quad \psi = \frac{1}{2}(\psi_1 + \psi_2).$$

If $\lambda = 0$ or $\mu = 0$, then $\chi + \xi$ belongs to $\mathcal{R}_{\omega,\epsilon}$ without further argument. Assume, therefore, that $\lambda \neq 0$ and $\mu \neq 0$. If $\lambda\mu > 0$, then $\chi + \xi$ belongs to $\mathcal{T}_\pi$, with $\pi = (1-\alpha)\phi + \alpha\psi$ and α given by

$$\alpha = \frac{\mu}{\lambda+\mu}.$$

Indeed, let

$$\begin{aligned}\pi_1 &= (1-\alpha)\phi_1 + \alpha\psi_1,\\ \pi_2 &= (1-\alpha)\phi_2 + \alpha\psi_2.\end{aligned}$$

Then both π_1 and π_2 belong to $\mathbb{M}$ and satisfy

$$\begin{aligned}\pi_1 + \pi_2 &= 2(1-\alpha)\phi + 2\alpha\psi\\ &= 2\pi\end{aligned}$$

and

$$\begin{aligned}(\lambda+\mu)(\pi_1 - \pi_2) &= (\lambda+\mu)\left[\frac{1-\alpha}{\lambda}\chi + \frac{\alpha}{\mu}\xi\right]\\ &= \chi + \xi.\end{aligned}$$

In addition,

$$\begin{aligned}\|\pi - \omega\| &= \|(1-\alpha)(\phi-\omega) + \alpha(\psi-\omega)\|\\ &\le \|(1-\alpha)(\phi-\omega)\| + \|\alpha(\psi-\omega)\|\\ &< \epsilon.\end{aligned}$$

One concludes that in this case, $\chi + \xi$ belongs to $\mathcal{R}_{\omega,\epsilon}$.

The case that $\lambda\mu < 0$ is similar. That $\chi \in \mathcal{R}_{\omega,\epsilon}$ implies $\lambda\chi \in \mathcal{R}_{\omega,\epsilon}$ is straightforward. One can conclude that $\mathcal{R}_{\omega,\epsilon}$ is a linear space. It clearly is a subspace of $T_\omega\mathbb{M}$.

Proposition 4: $\mathcal{T}_\omega^{\text{approx}}$ is a closed linear subspace of $T_\omega\mathbb{M}$.

Proof:

The lemma shows that $\mathcal{R}_{\omega,\epsilon}$ is a linear subspace of $T_\omega\mathbb{M}$, which is a space closed in norm. Hence, also the norm closure of $\mathcal{R}_{\omega,\epsilon}$ is a subset of this space and therefore also of $T_\omega\mathbb{M}$.

7 Majorized states

The subset of states majorized by a multiple of the reference state ω is considered.

Definition 1: A state ϕ on $\mathfrak{M}$ *is* said to be majorized by a multiple of the state ω if there exists a positive constant λ such that

$$\phi(x^*x) \le \lambda\omega(x^*x) \quad \text{for all } x \in \mathfrak{M}.$$

Take $a' \neq 0$ in the commutant algebra $\mathfrak{M}'$ and let

$$\Phi = \frac{1}{\|a'\Omega\|}a'\Omega.$$

Then, the state ω_Φ is majorized by a multiple of the state ω. Indeed, one has for any positive x in $\mathfrak{M}$

$$\begin{aligned}\omega_\Phi(x) &= \frac{(xa'\Omega, a'\Omega)}{(a'\Omega, a'\Omega)}\\ &= \frac{\left(a'^*a'x^{1/2}\Omega, x^{1/2}\Omega\right)}{(a'\Omega, a'\Omega)}\\ &\le \frac{\|a'^*a'\|}{(a'\Omega, a'\Omega)}\omega(x).\end{aligned}$$

It is well-known that all states majorized by a multiple of the state ω are obtained in this way. This is the content of the following proposition.

Proposition 5: If the vector state ω_Φ is majorized by a multiple of the state ω, then there exists a unique element a$'$ of the commutant $\mathfrak{M}'$ such that $\Phi = a'\Omega$.

Proof:

An operator a' is densely defined by

$$a'x\Omega = x\Phi, \qquad x \in \mathfrak{M}.$$

It satisfies $a'\Omega = \Phi$. It is well-defined because $x\Omega = 0$ implies

$$\|x\Phi\|^2 = \omega_\Phi(x^*x) \le \text{constant}\,\omega(x^*x) = \text{constant}\,\|x\Omega\|^2 = 0$$

so that $x\Phi = 0$.

The operator a' is bounded because

$$\|a'x\Omega\|^2 = \phi(x^*x) \le \text{constant}\,\omega(x^*x) = \text{constant}\,\|x\Omega\|^2.$$

The operator a' commutes with any x in $\mathfrak{M}$ because

$$a'x(y\Omega) = xy\Phi = x(a'y\Omega) = xa'(y\Omega)$$

and Ω is cyclic for $\mathfrak{M}$.

The operator a' is unique. Indeed, assuming b' in $\mathfrak{M}'$ satisfies $\Phi = b'\Omega$. Then, one has for all x in $\mathfrak{M}$

$$0 = x(a' - b')\Omega = (a' - b')x\Omega.$$

Hence, $a' - b'$ vanishes on $\mathfrak{M}\Omega$ which is dense in the Hilbert space because Ω is cyclic for $\mathfrak{M}$. Because $a' - b'$ is a bounded and hence continuous operator, it vanishes everywhere so that $a' = b'$.

Item (8) of Theorem 3 of [22] implies the following.

Proposition 6: If a vector state ω_Φ, defined by a vector Φ in the natural positive cone $\mathcal{P}_\Omega$, is dominated by a multiple of the state ω, then there exists a unique element a in the algebra $\mathfrak{M}$ such that $\Phi = a\Omega$ and

$$\omega_\Phi(x) = \omega(a^*xa), \qquad x \in \mathfrak{M}.$$

Proof:

Proposition 5 shows that a' in the commutant $\mathfrak{M}'$ exists such that $\Phi = a'\Omega$. Because Φ and Ω both belong to $\mathcal{P}_\Omega$, one has $\Phi = J\Phi = Ja'J\Omega$.

Let $a = Ja'J$. From $J\mathfrak{M}'J = \mathfrak{M}$, it follows that a belongs to $\mathfrak{M}$. This shows the existence.

The element a is unique because the correspondence between vector states on $\mathfrak{M}$ and vectors in $\mathcal{P}_\Omega$ is one-to-one and Ω is a separating vector for $\mathfrak{M}$.

If $\mathfrak{M}$ is a commutative algebra, then a^*a is the Radon–Nikodym derivative of the state ω_Φ with respect to the reference state ω.

The subset of states of $\mathbb{M}$ majorized by a multiple of the state ω is dense in $\mathbb{M}$ in the sense that for any state ϕ in $\mathbb{M}$, there exists a sequence $(a_n)_n$ of elements of $\mathfrak{M}$ with the property that $a_n\Omega$ is a Cauchy sequence and

$$\phi(x) = \lim_{n\to\infty} \omega(a_n^*xa_n), \qquad x \in \mathfrak{M}.$$

See Propositions 1.5 and 2.5 of [32].

Proposition 7: *A tangent vector χ belongs* to the subspace $\mathcal{T}_\omega$ of the tangent space $T_\omega\mathbb{M}$ if and only if it is proportional to the difference of two states ϕ and ψ in $\mathbb{M}$, both majorized by a multiple of the state ω.

Proof:

If χ belongs to $\mathcal{T}_\omega$, then by definition, there exist states ϕ and ψ in $\mathbb{M}$ such that $\chi = \lambda(\phi - \psi)$ and $\phi + \psi = 2\omega$. The latter implies that both ϕ and ψ are majorized by 2ω.

Conversely, assume that ϕ and ψ in $\mathbb{M}$ are both majorized by a multiple of the state ω and let $\chi = \lambda(\phi - \psi)$. This implies the existence of $\mu \ge 1$ and $\nu \ge 1$ such that $\phi \le \mu\omega$ and $\psi \le \nu\omega$.

Without restriction, assume that $\lambda > 0$.

Introduce

$$\phi' = \omega + \rho\chi \quad \text{and} \quad \psi' = \omega - \rho\chi$$

with ρ still to be chosen. By construction, it holds that $\phi' + \psi' = 2\omega$ and $\phi' - \psi' = 2\rho\chi$. Hence, if ϕ' and ψ' are states in $\mathbb{M}$ and $\rho \neq 0$, then one can conclude that χ belongs to $\mathcal{T}_\omega$.

From

$$\chi(x^*x) \le \lambda\phi(x^*x) \le \lambda\mu\omega(x^*x)$$

and

$$\chi(x^*x) \ge -\lambda\psi(x^*x) \ge -\lambda\nu\omega(x^*x),$$

one obtains

$$\begin{aligned}\phi'(x^*x) &\ge [1 - \rho\lambda\nu]\omega(x^*x),\\ \psi'(x^*x) &\ge [1 - \rho\lambda\mu]\omega(x^*x).\end{aligned}$$

Let ρ be equal to the inverse of the maximum of $\lambda\mu$ and $\lambda\nu$ to prove the positivity of the functionals ϕ' and ψ'. Normalization $\phi'(1) = \psi'(1) = 1$ follows from $\chi(1) = 0$. The functions are σ-weakly continuous as well. Hence, they are states in $\mathbb{M}$. This ends the proof that χ belongs to $\mathcal{T}_\omega$.

8 Exponential arcs

[27] introduces the notion of an exponential arc in the Hilbert space, inspired by the notion of exponential arcs in probability space as introduced by [24, 25]. Here, a definition is given which depends on the choice of a relative entropy.

In the present context, a divergence function $D(\phi\|\psi)$ is a real function of two states ϕ and ψ in the manifold $\mathbb{M}$. It cannot be negative, and it vanishes if and only if the two arguments are equal. A value of $+\infty$ is allowed. An energy function is an affine function $\mathfrak{h}$ defined on a convex subset of the set of normal states on the algebra $\mathfrak{M}$.

The following definition of an exponential arc in the manifold $\mathbb{M}$ assumes that a divergence function $D(\phi\|\psi)$ is given.

Definition 2: An exponential arc γ is a path in the manifold

$$t \in [0, 1] \mapsto \gamma_t \in \mathbb{M}$$

for which there exists an energy function $\mathfrak{h}$ such that

- γ_t is in the domain of $\mathfrak{h}$;
- The divergence $D(\gamma_s\|\gamma_t)$ between any two points of the arc is finite;
- For any state ψ in the domain of $\mathfrak{h}$, one has

$$D(\psi\|\gamma_t) = D(\psi\|\gamma_0) + D(\gamma_0\|\gamma_t) + t(\mathfrak{h}(\gamma_0) - \mathfrak{h}(\psi)), \qquad 0 \le t \le 1. \tag{6}$$

The energy function $\mathfrak{h}$ is the generator of the exponential arc. The arc is said to connect the state γ_1 to the state γ_0.

A subclass of energy functions is formed by the functions $\mathfrak{h}$ for which there exists a self-adjoint operator h in the von Neumann algebra $\mathfrak{M}$ so that

$$\mathfrak{h}(\psi) = \psi(h), \qquad \psi \in \mathbb{M}. \tag{7}$$

In such a case, h is called the generator as well. The exponential arcs defined in [27] agree with the above definition with a generator defined by an unbounded operator affiliated with the commutant algebra $\mathfrak{M}'$.

Proposition 8: Expression (6) implies

$$D(\gamma_s\|\gamma_0) + D(\gamma_0\|\gamma_s) = s(\mathfrak{h}(\gamma_s) - \mathfrak{h}(\gamma_0)). \tag{8}$$

and

$$D(\psi\|\gamma_t) = D(\psi\|\gamma_s) + D(\gamma_s\|\gamma_t) + (t-s)(\mathfrak{h}(\gamma_s) - \mathfrak{h}(\psi)). \tag{9}$$

It should be noted that with $s = 0$, expression (9) reduces to (6).
Proof:
Take $\psi = \gamma_s$ in (6) to find

$$D(\gamma_s\|\gamma_t) = D(\gamma_s\|\gamma_0) + D(\gamma_0\|\gamma_t) + t(\mathfrak{h}(\gamma_0) - \mathfrak{h}(\gamma_s)), \qquad 0 \le s, t \le 1. \tag{10}$$

In particular, with $s = t$, this implies (8).
To prove (9), use (10) to write the right-hand side as

$$\begin{aligned}\text{r.h.s.} = {} & D(\psi\|\gamma_s) + D(\gamma_s\|\gamma_0) + D(\gamma_0\|\gamma_t) + t(\mathfrak{h}(\gamma_0) - \mathfrak{h}(\psi)) \\ & - s(\mathfrak{h}(\gamma_s) - \mathfrak{h}(\psi)).\end{aligned}$$

Next, eliminate $D(\gamma_0\|\gamma_t)$ and $D(\psi\|\gamma_s)$ with the help of (6). This gives

$$\begin{aligned}\text{r.h.s.} &= D(\psi\|\gamma_s) + D(\gamma_s\|\gamma_0) + D(\psi\|\gamma_t) - D(\psi\|\gamma_0) - s(\mathfrak{h}(\gamma_s) - \mathfrak{h}(\psi)) \\ &= D(\gamma_s\|\gamma_0) + D(\psi\|\gamma_t) + D(\gamma_0\|\gamma_s) + s(\mathfrak{h}(\gamma_0) - \mathfrak{h}(\gamma_s)) \\ &= D(\psi\|\gamma_t).\end{aligned}$$

To obtain the last line, use (8).

Corollary 1: *If* $t \mapsto \gamma_t$ is an exponential arc with generator $\mathfrak{h}$ that connects γ_1 to γ_0, then for any s, t in [0, 1], the map $\epsilon \mapsto \gamma_{(1-\epsilon)s+\epsilon t}$ is an exponential arc with generator $(t-s)\mathfrak{h}$ that connects γ_t to γ_s.

Corollary 2: If $t \mapsto \gamma_t$ is an exponential arc with generator $\mathfrak{h}$ that connects γ_1 to γ_0, then $t \mapsto \gamma_{1-t}$ is an exponential arc with generator $-\mathfrak{h}$, connecting the state γ_0 to the state γ_1.

The following two propositions deal with the uniqueness of an exponential arc and of its generator.

Proposition 9: Let ω and ϕ be two states in $\mathbb{M}$. Fix an energy function $\mathfrak{h}$. There is at most one exponential arc $t \mapsto \gamma_t$ with generator $\mathfrak{h}$ that connects ϕ to ω.

Proof:

Assume both $t \mapsto \gamma_t$ and $t \mapsto \delta_t$ are exponential arcs connecting the state ϕ to the state ω. Subtract (6) from the same expression with γ_t replaced by δ_t and take $s = 0$. This gives

$$D(\psi\|\delta_t) - D(\psi\|\gamma_t) = D(\omega\|\delta_t) - D(\omega\|\gamma_t). \tag{11}$$

Take ψ equal to δ_t. Then, one obtains

$$0 \ge -D(\delta_t\|\gamma_t) = D(\omega\|\delta_t) - D(\omega\|\gamma_t).$$

On the other hand, with $\psi = \gamma_t$, one obtains

$$0 \le D(\gamma_t\|\delta_t) = D(\omega\|\delta_t) - D(\omega\|\gamma_t).$$

The two expressions together yield

$$D(\omega\|\delta_t) - D(\omega\|\gamma_t) = 0.$$

This implies $D(\gamma_t\|\delta_t) = 0$. By the basic property of a divergence, one concludes that $\gamma_t = \delta_t$.

Proposition 10: If the exponential arc $t \mapsto \gamma_t$ has two generators $\mathfrak{h}$ and $\mathfrak{k}$, then these generators differ by a constant on their common domain of definition.

Proof:

It follows from (6) that

$$\mathfrak{h}(\gamma_s) - \mathfrak{h}(\psi) = \mathfrak{k}(\gamma_s) - \mathfrak{k}(\psi), \qquad s \in [0,1] \tag{12}$$

for all states ψ in the intersection of the domains of $\mathfrak{h}$ and $\mathfrak{k}$. This implies that a constant c exists so that

$$\mathfrak{k}(\psi) = \mathfrak{h}(\psi) + c$$

for all ψ in the common domain.

The requirement (6) is a stability condition. The generator $\mathfrak{h}$ is a perturbation which shifts the state γ_0 to the state ψ. This interpretation will become clear further on. The effect on the relative entropy of the shift along the arc $t \mapsto \gamma_t$ is linear. In the standard case, the relative entropy is based on the logarithmic function. This justifies calling the path $t \mapsto \gamma_t$ an exponential arc.

It should be noted that the Pythagorean relation [33, 34]

$$D(\psi\|\gamma_t) = D(\psi\|\gamma_s) + D(\gamma_s\|\gamma_t)$$

is satisfied for all ψ with the same *energy* as the state γ_s, i.e., with

$$\mathfrak{h}(\psi) = \mathfrak{h}(\gamma_s).$$

If the divergence function is interpreted as the square of a pseudo-distance, then the aforementioned relation states that for an arbitrary state ψ, the point γ_s of the arc which has the same energy is the point with minimal distance.

9 The scalar potential

> *The exponential arc has a dual structure similar to that found in information geometry [10, 11].*

Given an exponential arc $t \mapsto \gamma_t$ with generator $\mathfrak{h}$, introduce the potential Φ_γ defined by

$$\Phi_\gamma(t) = D(\gamma_0\|\gamma_t) + t\mathfrak{h}(\gamma_0).$$

Its Legendre transform is given by

$$\Phi^*_\gamma(\alpha) = \sup\{\alpha t - \Phi_\gamma(t) : 0 \le t \le 1\}.$$

Proposition 11: For any exponential arc $t \mapsto \gamma_t$ with generator $\mathfrak{h}$, one has

(a) The function $t \mapsto \mathfrak{h}(\gamma_t)$ is strictly increasing;
(b) $\Phi_\gamma(t) = \Phi_\gamma(s) + D(\gamma_s\|\gamma_t) + (t-s)\mathfrak{h}(\gamma_s)$;
(c) The line $t \mapsto \Phi_\gamma(s) + (t-s)\mathfrak{h}(\gamma_s)$ is tangent to the potential Φ_γ at the point $t = s$; this implies that the potential $\Phi_\gamma(s)$ is a strictly convex function, continuous on the open interval (0, 1);
(d) The following identity holds:

$$\Phi_\gamma(s) + \Phi^*_\gamma(\mathfrak{h}(\gamma_s)) = s\mathfrak{h}(\gamma_s), \qquad s \in [0,1].$$

Proof:

(a) Take $\psi = \gamma_t$ in (6). This gives

$$0 = D(\gamma_t\|\gamma_s) + D(\gamma_s\|\gamma_t) + (t-s)(\mathfrak{h}(\gamma_s) - \mathfrak{h}(\gamma_t)).$$

Because divergences cannot be negative, this implies that $t \mapsto \mathfrak{h}(\gamma_t)$ is non-decreasing. Assume now that $\mathfrak{h}(\gamma_s) = \mathfrak{h}(\gamma_t)$. Then, it follows that

$$0 = D(\gamma_t\|\gamma_s) = D(\gamma_s\|\gamma_t).$$

The latter implies that $s = t$. One concludes that $s < t$ implies a strict inequality $\mathfrak{h}(\gamma_s) < \mathfrak{h}(\gamma_t)$.

(b) From the definition of the exponential arc, one obtains

$$D(\gamma_s\|\gamma_t) + (t-s)\mathfrak{h}(\gamma_s) = D(\psi\|\gamma_t) - D(\psi\|\gamma_s) + (t-s)\mathfrak{h}(\psi).$$

Take $\psi = \gamma_0$ in this expression to find

$$\begin{aligned} D(\gamma_s\|\gamma_t) + (t-s)\mathfrak{h}(\gamma_s) &= D(\gamma_0\|\gamma_t) - D(\gamma_0\|\gamma_s) + (t-s)\mathfrak{h}(\gamma_0) \\ &= \Phi_\gamma(t) - \Phi_\gamma(s). \end{aligned}$$

(c) From (b), one obtains

$$\Phi_\gamma(t) - t\mathfrak{h}(\gamma_s) \ge \Phi_\gamma(s) - s\mathfrak{h}(\gamma_s), \qquad 0 \le t \le 1 \tag{13}$$

because $D(\gamma_s\|\gamma_t) \ge 0$ with equality if and only if $s = t$. This implies that $t \mapsto \Phi_\gamma(s) + (t-s)\mathfrak{h}(\gamma_s)$ is a line tangent to the potential $\Phi_\gamma(s)$. By (a), the slope of this line is a strictly increasing function of s. Hence, the potential $\Phi_\gamma(s)$ is a strictly convex function, continuous on the open interval (0, 1).

(d) (13) implies that

$$\begin{aligned} \Phi_\gamma^*(\mathfrak{h}(\gamma_s)) &= \sup_t t\mathfrak{h}(\gamma_s) - \Phi_\gamma(t) \\ &\le s\mathfrak{h}(\gamma_s) - \Phi_\gamma(s). \end{aligned}$$

On the other hand, one can use (b) to obtain

$$\begin{aligned} \Phi_\gamma(s) + \Phi_\gamma^*(\mathfrak{h}(\gamma_s)) &\ge \Phi_\gamma(s) + t\mathfrak{h}(\gamma_s) - \Phi_\gamma(t) \\ &= t\mathfrak{h}(\gamma_s) - [D(\gamma_s\|\gamma_t) + (t-s)\mathfrak{h}(\gamma_s)] \\ &= -D(\gamma_s\|\gamma_t) + s\mathfrak{h}(\gamma_s). \end{aligned}$$

The optimal choice $t = s$ yields the lower bound $s\mathfrak{h}(\gamma_s)$.

A dual parameter η of the exponential arc γ, dual to the parameter t, is the value $\mathfrak{h}(\gamma_t)$ of the generator $\mathfrak{h}$. By item (a) of the proposition, it is a strictly increasing function of t. It is almost equal everywhere to the derivative $\dot{\Phi}_\gamma(t)$ of the value of the potential along the path.

10 The matrix case

If ρ and σ are two density matrices, then the obvious definition of an exponential arc connecting σ to ρ is

$$t \mapsto \sigma_t = \exp(\log\rho + t(\log\sigma - \log\rho) - \zeta(t))$$

with normalization $\zeta(t)$ given by

$$\zeta(t) = \log \textbf{Tr} \exp(\log\rho + t(\log\sigma - \log\rho)).$$

It is shown below that the corresponding states given by

$$\phi_t(x) = \textbf{Tr}\, \sigma_t x, \qquad x \in \mathfrak{A}$$

form an exponential arc for the relative entropy of Umegaki [20] in the GNS-representation of the state σ_0.

Fix a non-degenerate density matrix ρ of size n-by-n. It is a positive-definite matrix with trace Tr ρ equal to 1.

Umegaki's relative entropy for the pair of density matrices σ, τ is given by

$$D(\sigma\|\tau) = \mathrm{Tr}\,\sigma(\log\sigma - \log\tau).$$

Assume now a map

$$t \mapsto \sigma_t = \exp(\log\rho + th - \zeta(t)) \tag{14}$$

with normalization $\zeta(t)$ and with h given by

$$h = \log\sigma - \log\rho.$$

This is the obvious definition of an exponential arc in terms of density matrices. The corresponding potential is

$$\begin{aligned} \Phi_\sigma(t) &= D(\sigma_0\|\sigma_t) + t\mathfrak{h}(\sigma_0) \\ &= \zeta(t) \end{aligned}$$

with

$$\mathfrak{h}(\tau) = \mathbf{Tr}\,\tau h = \mathbf{Tr}\,\tau(\log\sigma - \log\rho).$$

The map (14) is also an exponential arc in the sense of Definition 2. To see this, consider any density matrix τ and calculate

$$\begin{aligned} D(\tau\|\sigma_t) - D(\tau\|\sigma_s) - D(\sigma_s\|\sigma_t) = &\ \mathbf{Tr}\,\tau(\log\tau - \log\sigma_t) \\ &- \mathbf{Tr}\,\tau(\log\tau - \log\sigma_s) \\ &- \mathbf{Tr}\,\sigma_s(\log\sigma_s - \log\sigma_t) \\ = &\ -(t-s)\,\mathbf{Tr}\,(\tau - \sigma_s)h \\ = &\ (t-s)(\mathfrak{h}(\gamma_s) - \mathfrak{h}(\tau)). \end{aligned}$$

This is of the form (6) except that the relative entropy is expressed in terms of density matrices in $\mathfrak{M}$ instead of vector states in the GNS representation of the state defined by the density matrix ρ.

An explicit construction of the GNS representation is possible. See for instance, the appendix of [28]. Let $\omega = \sigma_0$ denote the state determined by the density matrix ρ

$$\omega(A) = \mathbf{Tr}\,\rho A$$

for any n-by-n matrix A with entries in $\mathbb{C}$. Such a matrix A is represented on the Hilbert space $\mathcal{H} = \mathbb{C}^n \otimes \mathbb{C}^n$ by the operator $A \otimes \mathbb{I}$, where $\mathbb{I}$ is the n-by-n identity matrix. The von Neumann algebra $\mathfrak{M}$ is the space of operators $A \otimes \mathbb{I}$.

The matrix ρ can be diagonalized. This gives the spectral representation

$$\rho = \sum_i p_i e_i,$$

where $(e_i)_i$ is an orthonormal basis in $\mathbb{C}^n$. Let

$$\Omega = \sum_i \sqrt{p_i} e_i \otimes e_i.$$

It is a normalized vector in $\mathcal{H}$. One readily verifies that

$$\omega(A) = (A \otimes \mathbb{I}\Omega, \Omega)$$

for any n-by-n matrix A. In this way, any density matrix ρ defines a vector Ω in $\mathcal{H}$. The vector Ω is cyclic and separating for $\mathfrak{M}$ if ρ is non-degenerate. Hence, there is a one-to-one correspondence between non-degenerate density matrices and states in the manifold $\mathbb{M}$. It is then straightforward to replace the density matrices by states in the expressions obtained in the first part of this section.

11 The relative modular operator

Araki [35] introduces the relative modular operator $\Delta_{\Phi,\Psi}$ for any pair of vectors Φ and Ψ in the natural positive cone $\mathcal{P}$.

Assume that Φ and Ψ are vectors in $\mathcal{P}$ which are separating for the algebra $\mathfrak{M}$. Then, a conjugate–linear operator is defined by

$$x\Psi \mapsto x^*\Phi, \qquad x \in \mathfrak{M}.$$

It is well-defined because by assumption, $x\Psi = 0$ implies that $x = 0$ so that also $x^*\Phi = 0$. It is a closable operator. Indeed, assume the sequence $x_n\Psi$ converges to 0. Then, one has for any y in the commutant $\mathfrak{M}'$ that

$$(x_n\Psi, y\Phi) = (y\Psi, x_n^*\Phi)$$

converges to 0. By assumption, Ψ is separating for $\mathfrak{M}$ so that it is cyclic for the commutant $\mathfrak{M}'$. Hence, if the sequence $x_n^*\Phi$ converges, then it converges to 0. This shows the closability of the operator.

Let $S_{\Phi,\Psi}$ denote the closure of this operator. It satisfies

$$S_{\Phi,\Psi}x\Psi = x^*\Phi, \qquad x \in \mathfrak{M}.$$

Its inverse equals $S_{\Psi,\Phi}$.

The relative modular operator $\Delta_{\Phi,\Psi}$ is defined by

$$\Delta_{\Phi,\Psi} = S_{\Phi,\Psi}^* S_{\Phi,\Psi}.$$

Important properties of the relative modular operator are

$$\Delta_{\Phi,\Phi} = \Delta_\Phi \quad \text{and} \quad S_{\Phi,\Psi} = J\Delta_{\Phi,\Psi}^{1/2}.$$

where J is the modular conjugation operator for the vector Φ.

12 Araki's relative entropy

> Araki [22, 23] *uses* the relative modular operator $\Delta_{\Phi,\Psi}$ to define the relative entropy/divergence $D\,(\phi\|\psi)$ of the corresponding states $\phi = \omega_\Phi$ and $\psi = \omega_\Psi$ by
> $D(\phi\|\psi) = ((\log\Delta_{\Phi,\Psi})\Phi, \Phi)$.

Proposition 12: The divergence $D\,(\phi\|\psi)$ satisfies $D\,(\phi\|\psi) \geq 0$ with equality if and only if $\phi = \psi$.

Proof:

Let

$$\Delta_{\Phi,\Psi} = \int \lambda \mathrm{d}E_\lambda$$

denote the spectral decomposition of the operator $\Delta_{\Phi,\Psi}$. From the concavity of the logarithmic function, it follows that

$$\begin{aligned} D(\phi\|\psi) &= -\left(\left(\log \Delta_{\Phi,\Psi}^{-1}\right)\Phi, \Phi\right) \\ &= -\int \log\lambda^{-1} \mathrm{d}(E_\lambda\Phi, \Phi) \\ &\geq -\log \int \lambda^{-1} \mathrm{d}(E_\lambda\Phi, \Phi) \\ &= -\log\left(\Delta_{\Phi,\Psi}^{-1/2}\Phi, \Delta_{\Phi,\Psi}^{-1/2}\Phi\right) \\ &= -\log(\Psi, \Psi) \\ &= 0. \end{aligned}$$

This shows that the divergence cannot be negative.

If $\phi = \psi$, then one has

$$D(\phi\|\phi) = ((\log\Delta_\Phi)\Phi, \Phi) = 0$$

because $\Delta_\phi\Phi = \Phi$.

Finally, $D\,(\phi\|\psi) = 0$ implies that Φ is in the domain of $\log\,\Delta_{\Phi,\Psi}$ and that $\log\,\Delta_{\Phi,\Psi}\Phi = 0$. The latter implies that

$$\Psi = \Delta_{\Phi,\Psi}^{-1}\Phi = \Phi.$$

This shows that $D\,(\phi\|\psi) = 0$ vanishes only when $\Phi = \Psi$.

Theorem 2.4 of [35] shows that

$$\log\Delta_{\Phi,\Psi} + J\log\Delta_{\Psi,\Phi}J = 0.$$

Because Φ belongs, by assumption, to the natural positive cone $\mathcal{P}$, it satisfies $\Phi = J\Phi$. Hence, one has also

$$D(\phi\|\psi) = -((\log\Delta_{\Psi,\Phi})\Phi, \Phi).$$

13 A theorem

> *Each self-adjoint element h of the von Neumann algebra $\mathfrak{M}$ defines an exponential arc with a generator equal to the energy function defined by h.*

[21] constructs for each self-adjoint operator h in $\mathfrak{M}$ a vector Φ_h in the natural positive cone $\mathcal{P}$ and calls h the relative Hamiltonian. Inspection of the explicit expression used in [21] shows that

$$\Phi_h = \Omega + Xh\Omega + O\left(h^2\right) \tag{15}$$

with operator X given by

$$X = \int_0^{1/2} \mathrm{d}u \Delta_\Omega^u.$$

The vector Φ_h defines a state ϕ_h by

$$\phi_h(x) = e^{-\xi(h)}(x\Phi_h, \Phi_h), \qquad x \in \mathfrak{M}.$$

Here, $\xi(h)$ is the normalization

$$\xi(h) = \log(\Phi_h, \Phi_h).$$

Theorem 3.10 of [35] implies that the state ϕ_h obtained in this way satisfies for all ψ in $\mathbb{M}$

$$D(\psi\|\phi_h) = D(\psi\|\omega) - \psi(h) + \xi(h). \tag{16}$$

Take $\psi = \phi_h$ and $\psi = \omega$ to find that the normalization $\xi(h)$ is given by

$$\xi(h) = \phi_h(h) - D(\phi_h\|\omega) = \omega(h) + D(\omega\|\phi_h).$$

Consider now the path γ defined by $\gamma_t = \phi_{th}$. Then, (16) becomes

$$D(\psi\|\gamma_t) = D(\psi\|\omega) - t\psi(h) + \zeta(t). \tag{17}$$

with

$$\zeta(t) = t\gamma_t(h) - D(\gamma_t\|\omega) = t\omega(h) + D(\omega\|\gamma_t) = \xi(th).$$

From this last expression, one obtains

$$0 \leq D(\gamma_t\|\omega) + D(\omega\|\gamma_t) = t[\gamma_t(h) - \omega(h)].$$

From (15), we infer that γ_t converges to ω as $t \downarrow 0$. Hence, $D\,(\gamma_t\|\omega)$ and $D\,(\omega\|\gamma_t)$ converge to 0 faster than t. This implies that the derivative $\dot{\zeta}(0)$ exists and equals $\omega(h)$. This also implies that

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} D(\psi\|\gamma_t) = \omega(h) - \psi(h). \tag{18}$$

Elimination of $\zeta(t)$ from (17) yields

$$D(\psi\|\gamma_t) = D(\psi\|\omega) + D(\omega\|\gamma_t) + t(\omega(h) - \psi(h)).$$

This shows that γ is an exponential arc connecting γ_1 to $\gamma_0 = \omega$.

Proposition 13: One has

$$\dot{\gamma}_0(x) = (T_\Omega h\Omega, T_\Omega[x-\omega(x)]^*\Omega) \tag{19}$$

with the operator T_Ω given by

$$T_\Omega = \left(\frac{\Delta_\Omega - 1}{\log\Delta_\Omega}\right)^{1/2}. \tag{20}$$

It should be noted that this operator T_Ω was introduced in [36].

Proof:

From (15), one obtains

$$\dot{\gamma}_0(x) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\gamma_t(x) = (xXh\Omega,\Omega) + (x\Omega, Xh\Omega) - \dot{\zeta}(0)\omega(x). \tag{21}$$

Write

$$\begin{aligned}(xXh\Omega,\Omega) &= \int_0^{1/2}\mathrm{d}u\,(x\Delta_\Omega^u h\Omega,\Omega)\\ &= \int_0^{1/2}\mathrm{d}u\,(\Delta_\Omega^{u/2}h\Omega, \Delta_\Omega^{u/2}x^*\Omega)\end{aligned}$$

and

$$\begin{aligned}(x\Omega, Xh\Omega) &= \int_0^{1/2}\mathrm{d}u\,(x\Omega,\Delta_\Omega^u h\Omega)\\ &= \int_0^{1/2}\mathrm{d}u\,(\Delta_\Omega^{u/2}J\Delta_\Omega^{1/2}x^*\Omega, \Delta_\Omega^{u/2}J\Delta_\Omega^{1/2}h\Omega)\\ &= \int_0^{1/2}\mathrm{d}u\left(J\Delta_\Omega^{(1-u)/2}x^*\Omega, J\Delta_\Omega^{(1-u)/2}h\Omega\right)\\ &= \int_{1/2}^{1}\mathrm{d}u\,(J\Delta_\Omega^{u/2}x^*\Omega, J\Delta_\Omega^{u/2}h\Omega)\\ &= \int_{1/2}^{1}\mathrm{d}u\,(\Delta_\Omega^{u/2}h\Omega, \Delta_\Omega^{u/2}x^*\Omega).\end{aligned}$$

The two contributions to (21) can now be taken together. One obtains

$$\begin{aligned}\dot{\gamma}_0(x) &= \int_0^1\mathrm{d}u\,(\Delta_\Omega^{u/2}h\Omega, \Delta_\Omega^{u/2}x^*\Omega) - \dot{\zeta}(0)\omega(x)\\ &= (T_\Omega h\Omega, T_\Omega x^*\Omega) - \dot{\zeta}(0)\omega(x).\end{aligned}$$

Take $x = 1$ to see that

$$\dot{\zeta}(0) = (T_\Omega h\Omega, T_\Omega\Omega) = \omega(h)$$

so that it follows (19).

In summary, one can infer

Theorem 1: Let ω in $\mathbb{M}$ be a vector state with cyclic and separating vector Ω. Choose the divergence function equal to the relative entropy of Araki as defined by (15). For each self-adjoint element h in $\mathfrak{M}$, an energy function $\mathfrak{h}$ is defined by $\mathfrak{h}(\phi) = \phi(h)$ and there exists an exponential arc γ with generator $\mathfrak{h}$ connecting some state γ_1 of $\mathbb{M}$ to the state $\gamma_0 = \omega$. For any state ψ in $\mathbb{M}$, the derivative of $D(\psi\|\gamma_t)$ at $t = 0$ exists and is given by $\omega(h) - \psi(h)$. The derivative of the exponential arc at $t = 0$ satisfies (19).

Further properties hold for the exponential arc of the above theorem.

Proposition 14: For any exponential arc γ constructed in Theorem 1, the derivative $\dot{\gamma}_0$ is a Fréchet derivative.

Proof:

Let $\Xi(h)$ denote the remainder of order h^2 in (15), i.e.,

$$\Phi_h = \Omega + Xh\Omega + \Xi(h).$$

Then one can use (19) for

$$\begin{aligned}\gamma_t(x) - \gamma_0(x) - t\dot{\gamma}_0(x) &= e^{-\zeta(t)}(x\Phi_{th},\Phi_{th}) - \omega(x) - t(xXh\Omega,\Omega)\\ &\quad - t(x\Omega, Xh\Omega) + t\omega(h)\omega(x)\\ &= \left[e^{-\zeta(t)} - 1 + t\omega(h)\right]\omega(x) + t\left[e^{-\zeta(t)} - 1\right]\\ &\quad\times[(xXh\Omega,\Omega) + (x\Omega, Xh\Omega)]\\ &\quad + e^{-\zeta(t)}[(x\Xi(th),\Omega) + (x\Omega,\Xi(th))\\ &\quad + t^2(xXh\Omega, Xh\Omega)\\ &\quad + t(x\Xi(th), Xh\Omega)\\ &\quad + t(xXh\Omega,\Xi(th))\\ &\quad + (x\Xi(th),\Xi(th))].\end{aligned}$$

This yields

$$\begin{aligned}\|\gamma_t - \gamma_0 - t\dot{\gamma}_0\| \le\ & |e^{-\zeta(t)} - 1 + t\omega(h)|\\ & + 2t|e^{-\zeta(t)} - 1|\|Xh\Omega\|\\ & + 2e^{-\zeta(t)}\|\Xi(th)\|\\ & + e^{-\zeta(t)}[\|\Xi(th)\| + t\|Xh\Omega\|]^2.\end{aligned}$$

Each of the terms in the right-hand side of this expression is of order less than t as t tends to 0. Hence, $\dot{\gamma}_0$ is a Fréchet derivative.

Proposition 15 (Additivity of generators): If the state ϕ is connected to the state ω by the exponential arc with generator h and ψ is connected to ϕ by the exponential arc with generator k, then ψ is connected to ω by the exponential arc with generator $h + k$ and ω is connected to ψ by the exponential arc with generator $-h$.

For the proof, see Proposition 4.5 of [21].

14 The metric

Eguchi [37] introduced the technique of deriving the metric of the tangent space by taking two derivatives of the divergence. Application here yields the metric which is used in the Kubo–Mori theory of linear response [38, 39].

Consider two exponential arcs $t\mapsto\gamma_t$ and $s\mapsto\eta_s$ with respective generators h and k. They connect the states γ_1 and η_1 to the reference state ω. The tangent vectors at $s = t = 0$ are $\dot{\gamma}_0$ and $\dot{\eta}_0$. They belong to the tangent space $T_\omega\mathbb{M}$. The scalar product between them is by definition given by

$$(\dot{\eta}_0,\dot{\gamma}_0)_\omega = -\frac{\partial}{\partial s}\frac{\partial}{\partial t}\Big|_{s=t=0} D(\eta_s\|\gamma_t).$$

Assume now that these exponential arcs are those constructed in Theorem 1. Then, one has

$$\begin{aligned}(\dot{\eta}_0,\dot{\gamma}_0)_\omega &= \frac{\partial}{\partial s}\Big|_{s=t=0}(\omega(h) - \eta_s(h))\\ &= \dot{\eta}_0(h)\\ &= (T_\Omega k\Omega, T_\Omega(h-\omega(h))\Omega)\\ &= (T_\Omega(k-\omega(k))\Omega, T_\Omega(h-\omega(h))\Omega),\end{aligned} \tag{22}$$

with the operator T_Ω defined by (20). It should be noted that in most applications, one assumes that the expectations $\omega(h)$ of the generator h and $\omega(k)$ of the generator k vanish. Then, the result obtained here coincides with that used in [36]. In what follows, a non-vanishing expectation of the generators is taken into account.

Let us now discuss some technical issues. The scalar product is well-defined by (22). This follows from

Lemma 2: If two exponential arcs with initial point ω with generators h, respectively k, both in $\mathfrak{M}$, have *the* same initial tangent vector, then one has

$$T_\Omega(h-\omega(h))\Omega = T_\Omega(k-\omega(k))\Omega.$$

Proof:

Let γ and η be two exponential arcs with generators h, respectively k in $\mathfrak{M}$, such that $\gamma_0 = \eta_0 = \omega$. Without restriction, assume that $\omega(h) = \omega(k) = 0$ and $\dot{\gamma}_0 = \dot{\eta}_0$. Then, (19) implies that

$$(T_\Omega(h-k)\Omega, T_\Omega x^*\Omega) = 0, \qquad x \in \mathfrak{M}.$$

Take $x = h - k$. Then, it follows that $T_\Omega(h-k)\Omega = 0$.

This lemma shows that the map

$$\dot{\gamma}_0 \mapsto T_\Omega(h - \omega(h))\Omega \tag{23}$$

is one-to-one and identifies the tangent vector $\dot{\gamma}_0$ with the vector $T_\Omega h\Omega$ in the Hilbert space $\mathcal{H}$.

Expression (22) defines a bilinear form. This follows from.

Lemma 3: The map (23) is linear.

Proof:

Let γ be an exponential arc with generator h in $\mathfrak{M}$. Then, $t \mapsto \gamma_{\epsilon t}$ is an exponential arc with generator ϵh for any ϵ in $[-1, 1]$ and the tangent vector is $\epsilon\dot{\gamma}_0$. Hence, (23) maps $\epsilon\dot{\gamma}_0$ onto $\epsilon T_\Omega h\Omega$.

Next, consider a pair of exponential arcs γ and η with generators k, and h, respectively, in $\mathfrak{M}$ and with $\gamma_0 = \eta_0 = \omega$. Let θ denote the exponential arc with generator $h + k$. It exists by Theorem 1. The state θ_t can then be written as

$$\theta_t(x) = (x\Phi_{th+tk}, \Phi_{th+tk})$$

with Φ_{th+tk} being the unique element in the natural positive cone representing the state θ_t. Now, use (15) to write

$$\theta_t(x) = \omega(x) + \frac{t}{2}\int_0^1 \mathrm{d}u\,(x\Delta_\Omega^{u/2}(h+k)\Omega, \Omega)$$
$$+ \frac{t}{2}\int_0^1 \mathrm{d}u\,(x\Omega, \Delta_\Omega^{u/2}(h+k)\Omega) + \mathrm{o}(t).$$

This implies

$$\dot{\theta}_0 = \dot{\gamma}_0 + \dot{\eta}_0.$$

Both observations together prove the linearity of map (23).

Proposition 16: Expression (22) defines a non-degenerate scalar product on the space of tangent vectors of the form $\dot{\gamma}_0$ with γ an exponential arc as constructed in Theorem 1.

Proof:

The two lemmas show that (22) is a well-defined bilinear form. Positivity of the form is clear. The symmetry follows from (22). It remains to be shown that it is non-degenerate.

Assume that $(\dot{\gamma}_0, \dot{\gamma}_0) = 0$. This implies

$$T_\Omega(h - \omega(h))\Omega = 0,$$

with h the generator of γ. The operator T_Ω is invertible—see the proof of Lemma II.2 of [36]. Hence, it follows that

$$(h - \omega(h))\Omega = 0.$$

Because O is separating for $\mathfrak{M}$, it follows that h is a multiple of the identity. The latter implies that $\dot{\gamma} = 0$.

15 Dual geometries

The geodesics of the e-connection are the exponential arcs. In the m-connection, the geodesics are made up by convex combinations of a pair of states. The m- and e-connections are each others' dual with respect to the metric of Section 14.

Consider two states ω and ϕ in the manifold $\mathbb{M}$. The tangent vector

$$\dot{\gamma}_t = \frac{\mathrm{d}}{\mathrm{d}t}\gamma_t = \phi - \omega, \qquad 0 < t < 1,$$

is independent of t. Hence, it is a geodesic for the connection in which all parallel transport operators are taken equal to the identity operator. It should be noted that the tangent space $T_\omega\mathbb{M}$ coincides with the space of σ-weakly continuous linear functionals χ, satisfying $\chi(1) = 0$ and hence it is the same everywhere . This connection is by definition the m-connection.

For t in (0, 1), the tangent vector $\dot{\gamma}_t$ belongs to the subspace $\mathcal{T}_{\gamma_t}$ of the tangent space $T_\omega\mathbb{M}$ which is introduced in Section 6. Conversely, every vector χ in $\mathcal{T}_{\gamma_t}$ is the tangent vector of an m-geodesic passing through the point γ_t. However, this does not imply that through parallel transport $\Pi(\gamma_t \mapsto \gamma_s)$, the space $\mathcal{T}_{\gamma_t}$ maps onto the space $\mathcal{T}_{\gamma_s}$.

The transport operators Π^* of the dual geometry are defined by

$$(\Pi(\phi \mapsto \omega)V, \Pi^*(\phi \mapsto \omega)W)_\omega = (V, W)_\phi.$$

In this expression, V and W are vector fields and $(\cdot,\cdot)_\omega$ is the scalar product defined in the previous section and evaluated at the point ω of the manifold $\mathbb{M}$.

It can be shown that any exponential arc γ is a geodesic for this dual geometry. To do so, we have to show that

$$\Pi^*(\gamma_s \mapsto \gamma_t)\dot{\gamma}_s = \dot{\gamma}_t.$$

The tangent vector $\dot{\gamma}_t$ at $t = 0$ is given by (19). Its value for arbitrary t is given by the following proposition.

Proposition 17: Let γ denote an exponential arc γ with generator h belonging to $\mathfrak{M}$. Let Φ_t be the normalized vector in the natural positive cone $\mathcal{P}$ representing the state γ_t. The derivative $\dot{\gamma}_t$ is given by

$$\dot{\gamma}_t(x) = (T_{\Phi_t}h\Phi_t, T_{\Phi_t}[x - \gamma_t(x)]^*)\Phi_t), \qquad x \in \mathfrak{M}. \tag{24}$$

Proof:

The state γ_1 is connected to ω by the exponential arc with generator h and γ_t is connected to ω by the exponential arc with generator th. Let

$$\psi_s = \gamma_{(1-s)t+s}.$$

It follows from Proposition 8 that $s \mapsto \psi_s$ is an exponential arc with generator $(1-t)\,h$ connecting γ_t to γ_1. Application of (19) to the latter arc gives

$$\dot{\psi}_0(x) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0}\psi_s(x) = (1-t)\big(T_\Psi h\Psi, T_\psi(x - \omega(x))^*\Psi\big) \tag{25}$$

with $\Psi = \Phi_t$. This implies (24) because $\dot{\psi}_0 = (1-t)\dot{\gamma}_t$.

Theorem 2: Any exponential arc γ with generator h in $\mathfrak{M}$ is a geodesic for the dual of the m-connection with respect to the metric introduced in Section 14.

Proof:

Let $t \mapsto \phi_t$ be an exponential arc with generator k in $\mathfrak{M}$ such that $\phi_0 = \gamma_t$. Fix t in $[0, 1]$ and let Φ_t denote the normalized element of the natural positive cone $\mathcal{P}$ representing the state γ_t. Let η be an exponential arc with generator k starting at γ_t, i.e., $\eta_0 = \gamma_t$. Because $\Pi(\gamma_s \mapsto \gamma_t)$ is the identity, the definition of the dual transport operator yields

$$\begin{aligned}(\dot{\eta}_0, \Pi^*(\gamma_s \mapsto \gamma_t)\dot{\gamma}_s)_{\gamma_t} &= (\dot{\eta}_0, \dot{\gamma}_s)_{\gamma_s}\\ &= (T_{\Phi_t}(k - \gamma_t(k))\Phi_t, T_{\Phi_t}(l - \gamma_t(l))\Phi_t),\end{aligned}$$

with l the generator of the arc $s \mapsto \gamma_{(1-s)t+s}$. It equals $l = (1 - t)h$. This last expression equals

$$= (\dot{\eta}_0, \dot{\gamma}_t)_{\gamma_t}.$$

By proposition 16, the scalar product $(\cdot, \cdot)_{\gamma_t}$ is non-degenerate. Therefore, one can conclude that

$$\Pi^*(\gamma_s \mapsto \gamma_t)\dot{\gamma}_s = \dot{\gamma}_t.$$

This shows that the exponential arc γ is a geodesic for the dual of the m-connection.

16 Finite-dimensional submanifolds

A finite set of linearly independent generators is shown to define a finite-dimensional submanifold in which all states are connected to the reference state by an exponential arc. The submanifold defined in this way is a dually flat quantum statistical manifold.

Let ω be the reference state of $\mathbb{M}$. It is a vector state with a cyclic and separating vector Ω. Choose an independent set of self-adjoint operators $h_1, \ldots, h_n$ in $\mathfrak{M}$. By Theorem 1, there exists an exponential arc γ with generator $h = \theta^i h_i$ connecting some state γ_1 in $\mathbb{M}$ to the state $\gamma_0 = \omega$. A parameterized family of states ω_θ, $\theta \in \mathbb{R}^n$ is now defined by putting $\omega_\theta = \gamma_1$. These states form a submanifold of $\mathbb{M}$.

From the definition of an exponential arc, one obtains immediately that for any ψ in $\mathbb{M}$

$$D(\psi\|\omega_\theta) = D(\psi\|\omega) + D(\omega\|\omega_\theta) + \theta^i(\omega(h_i) - \psi(h_i)). \tag{26}$$

Take $\psi = \omega_\theta$ in this expression to find

$$\theta^i\omega(h_i) \le D(\omega\|\omega_\theta) + \theta^i\omega(h_i) = \theta^i\omega_\theta(h_i) - D(\omega_\theta\|\omega) \le \theta^i\omega_\theta(h_i). \tag{27}$$

Hence, the quantity $\theta^i\omega(h_i)$ is maximal if and only if ω_θ equals the reference state ω.

Proposition 18: Dual coordinates η_i are defined by

$$\eta_i = \omega_\theta(h_i).$$

They satisfy

$$\frac{\partial\eta_i}{\partial\theta^j} = (\partial_i, \partial_j)_\theta$$

with $(\cdot,\cdot)_\theta$ equal to the scalar product $(\cdot,\cdot)_{\omega_\theta}$ introduced in Section 14 and with basis vectors ∂_i equal to $\partial\omega_\theta/\partial\theta^i$.

Proof:

Introduce the path $\gamma^{(i)}$ defined by

$$\gamma^{(i)}: t \mapsto \omega_{\theta+tg_i}.$$

It satisfies

$$\frac{\partial}{\partial\theta^i}\omega_\theta = \partial_i = \dot{\gamma}^{(i)}(0).$$

By definition, $\omega_{\theta+g_i}$ is the end point of the exponential arc with generator $(\theta^j + g_i^j)h_i$. From Proposition 15, it then follows that $\gamma^{(i)}$ is an exponential arc with generator h_i connecting $\omega_{\theta+g_i}$ to ω_θ. These arcs $\gamma^{(i)}$ are used in the calculation that follows.

The definition of the scalar product at the beginning of Section 14 gives

$$\begin{aligned}(\partial_i, \partial_j)_{\omega_\theta} &= \left(\dot{\gamma}^{(i)}(0), \dot{\gamma}^{(j)}(0)\right)_{\omega_\theta}\\ &= -\frac{\partial}{\partial s}\frac{\partial}{\partial t}\bigg|_{s=t=0} D\left(\gamma_s^{(i)}\|\gamma_t^{(j)}\right)\\ &= -\frac{\partial}{\partial s}\bigg|_{s=0}\frac{\partial}{\partial\theta^j}D(\gamma_s^{(i)}\|\omega_\theta)\\ &= \frac{\partial}{\partial s}\bigg|_{s=0}\gamma_s^{(i)}(h_j)\\ &= \partial_i(h_j)\\ &= \frac{\partial\eta_j}{\partial\theta^i}.\end{aligned}$$

Corollary 3: There exists a potential $\Phi(\theta)$ such that

$$\eta_i = \frac{\partial\Phi}{\partial\theta^i}. \tag{28}$$

This follows because the scalar product is symmetric so that

$$\frac{\partial\eta_j}{\partial\theta^i} = (\partial_i, \partial_j)_\theta = (\partial_j, \partial_i)_\theta = \frac{\partial\eta_i}{\partial\theta^j}.$$

This symmetry is a sufficient condition for the potential $\Phi(\theta)$ to exist.

Consider the following generalization of the potential introduced in Section 9.

$$\Phi(\theta) = D(\omega\|\omega_\theta) + \theta^i\omega(h_i). \tag{29}$$

Apply (18) to the exponential arc $\gamma^{(i)}$ which connects $\omega_{\theta+g_i}$ to ω_θ to find

$$\frac{\partial}{\partial\theta^i}D(\omega\|\omega_\theta) = \omega_\theta(h_i) - \omega(h_i).$$

This implies that $\Phi(\theta)$ satisfies (28).

One can conclude that the selection of an independent set of self-adjoint operators $h_1, \ldots, h_n$ in $\mathfrak{M}$ defines a parameterized statistical model $\theta \mapsto \omega_\theta$ of states on the von Neumann algebra $\mathfrak{M}$. An obvious basis in the tangent plane $T_{\omega_\theta}\mathbb{M}$ is formed by the derivative operators ∂_i. The scalar product $(\partial_i, \partial_j)_{\omega_\theta}$ introduced in Section 14 starting from the relative entropy of Araki defines a Hessian metric on the tangent planes. Exponential arcs are geodesics for the e-connection which is the dual of the m-connection.

17 Discussion

- The manifold $\mathbb{M}$ under consideration consists of vector states on a sigma-finite von Neumann algebra $\mathfrak{M}$ in its standard

representation. Such a manifold has nice properties described by the Tomita–Takesaki Theory and hence is an obvious study object when exploring quantum statistical manifolds in an infinite-dimensional setting. Particular attention is given in the present work on the definition of the tangent planes. This is also a point of concern in the commutative context of manifolds of probability measures. See, for instance, the approach of [14]. A convenient choice for the tangent space $T_\omega\mathbb{M}$ at the state ω in the manifold $\mathbb{M}$ is to take it equal to the space of all σ-weakly continuous Hermitian linear functionals χ on $\mathfrak{M}$ vanishing on the identity operator $\mathbb{I}$. However, it is well-possible that the equivalence class of smooth curves through ω with initial tangent equal to a given χ is empty. Approximate tangent planes are considered an alternative in Section 6. They form a subspace of $T_\omega\mathbb{M}$ as defined previously. Nevertheless, the initial tangent vectors of Fréchet-differentiable paths starting at ω belong to the approximate tangent space. It is not clear whether the initial tangents of exponential arcs are dense in the approximate tangent space with respect to the inner product of Section 14. Further research is needed at this point.
- A new definition of exponential arcs is given. It depends on the choice of a divergence function/relative entropy defined on pairs of points in the manifold and on the choice of a generator which is a linear functional defined on a domain in the manifold. It is general enough to cover different approaches that one can follow to solve the non-uniqueness problem of the Radon–Nikodym derivative in the context of non-commutative probability. Nevertheless, one can prove in full generality nice properties such as uniqueness of the generator, existence of scalar potential, and Pythagorean relations. The additivity of generators when composing exponential arcs is shown in the specific context of Araki's relative entropy. See Proposition 15.
- The second half of the paper focuses on the relative entropy of Araki. Only exponential arcs with bounded generators belonging to the von Neumann algebra are considered. This suffices to reach the goal of replacing the existing approach based on density matrices and Umegaki's relative entropy. However, the solution of the problem mentioned previously regarding the extent of the tangent spaces most likely requires the handling of unbounded generators.
- The scalar product of Bogoliubov presented in Section 14 is used extensively in Linear Response Theory, also known as Kubo–Mori theory. Its link with the KMS condition of Section 2 is not highlighted in the present text. It is tradition in the Kubo–Mori theory and more generally in statistical mechanics to focus on a small number of variables. It is shown in Section 16 that the selection of a finite number of variables defines a quantum statistical manifold supporting Amari's dually flat geometry.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material: further inquiries can be directed to the corresponding author.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

The publication charges are covered by the Universiteit Antwerpen.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Dixmier J. *Les C*-algèbres et leurs représentations*. Paris: Gauthier-Villars (1964).

2. Dixmier J. *Les algèbres d'operateurs dans l'espace Hilbertien*. Paris: Gauthier-Villars (1969).

3. Ruelle D. *Statistical mechanics, Rigorous results*. New York: W.A. Benjamin, Inc. (1969).

4. Emch G. *Algebraic methods in statistical mechanics and quantum field theory*. New Jersey, United States: John Wiley & Sons (1972).

5. Bratteli O, Robinson D. *Operator algebras and quantum statistical mechanics*. New York, Berlin: Springer (1979).

6. Haag R, Hugenholtz NM, Winnink M. On the equilibrium states in quantum statistical mechanics. *Commun Math Phys* (1967) 5:215–36. doi:10.1007/bf01646342

7. Takesaki M. Tomita's theory of modular Hilbert algebras and its applications. In: *Lecture notes in mathematics*. Berlin: Springer-Verlag (1970).

8. Chentsov NN. Statistical decision rules and optimal inference. In: *Transl. Math. Monographs*. Rhode Island, United States: American Mathematical Society (1972).

9. Efron B. Defining the curvature of a statistical problem. *Ann Stat* (1975) 3:1189–242.

10. Amari S. Differential-geometrical methods in statistics. In: *Lecture notes in statistics*. New York, Berlin: Springer (1985).

11. Amari S, Nagaoka H. Methods of information geometry. In: *Translations of mathematical monographs*. Oxford, United Kingdom: Oxford University Press (2000).

12. Ay N, Jost J, Vân Lê H, Schwachhöfer L. *Information geometry*. Berlin, Germany: Springer (2017).

13. Petz D. *Quantum information theory and quantum statistics*. (Berlin: Springer-Verlag) (2008).

14. Pistone G, Sempi C. An infinite-dimensional structure on the space of all the probability measures equivalent to a given one. *Ann Stat* (1995) 23:1543–61.

15. Grasselli MR, Streater RF. On the uniqueness of the chentsov metric in quantum information geometry. *Infin Dim Anal Quan Prob. Rel. Top.* (2001) 4:173–82. doi:10.1142/s0219025701000462

16. Streater RF. Duality in quantum information geometry. *Open Syst Inf Dyn* (2004) 11: 71–7. doi:10.1023/b:opsy.0000024757.25401.db

17. Streater RF. Quantum orlicz spaces in information geometry. *Open Syst Inf Dyn* (2004) 11:359–75. doi:10.1007/s11080-004-6626-2

18. Jenčová A. A construction of a nonparametric quantum information manifold. *J Funct Anal* (2006) 239:1–20. doi:10.1016/j.jfa.2006.02.007

19. Grasselli MR. Dual connections in nonparametric classical information geometry. *Ann Inst Stat Math* (2010) 62:873–96. doi:10.1007/s10463-008-0191-3

20. Umegaki H. Conditional expectation in an operator algebra. IV. Entropy and information. *Kodai Math Sem Rep* (1962) 14:59–85. doi:10.2996/kmj/1138844604

21. Araki H. Relative Hamiltonian for faithful normal states of a von Neumann algebra. *RIMS* (1973) 9:165–209.

22. Araki H. Some properties of modular conjugation operator of von Neumann algebras and a non-commutative Radon–Nikodym theorem with a chain rule. *Pac J Math* (1974) 50:309–54. doi:10.2140/pjm.1974.50.309

23. Araki H. Relative entropy of states of von Neumann algebras. *Publ RIMS Kyoto Univ* (1976) 11:809–33.

24. Pistone G, Cena A. Exponential statistical manifold. *AISM* (2007) 59:27–56. doi:10.1007/s10463-006-0096-y

25. Pistone G. Nonparametric information geometry. In: F Nielsen F Barbaresco, editors. *Geometric science of information*. Berlin, Germany: Springer (2013). p. 5–36.

26. Santacroce M, Siri P, Trivellato B. On mixture and exponential connection by open arcs. In: F Nielsen F Barbaresco, editors. *Geometric science of information*. Berlin, Germany: Springer (2017). p. 577–84.

27. Naudts J. Exponential arcs in the manifold of vector states on a σ-finite von Neumann algebra. *Inf Geom* (2022) 5:1–30. doi:10.1007/s41884-021-00064-4

28. Naudts J. Quantum statistical manifolds. *Entropy* (2018) 20:472. doi:10.3390/e20060472

29. Naudts J. Correction: Naudts, J. Quantum statistical manifolds. Entropy 2018, 20, 472. *Entropy* (2018) 20:796. doi:10.3390/e20100796

30. Ciaglia FM, Ibort A, Jost J, Marmo G. Manifolds of classical probability distributions and quantum density operators in infinite dimensions. *Inf Geom* (2019) 2:231–71. doi:10.1007/s41884-019-00022-1

31. Simon L. Lectures on geometric measure theory. In: *Proceedings of the centre for mathematical Analysis*. Australia: Australian National University (1983).

32. Niestegge G. Absolute continuity for linear forms on b^*-algebras and a Radon-Nikodym type theorem (quadratic version). *Rend Circ Mat Palermo* (1983) 32:358–76. doi:10.1007/bf02848539

33. Csiszár I. I-divergence geometry of probability distributions and minimization problems. *Ann Probab* (1975) 3:146–58.

34. Cziszár I, Matúš F. Generalized minimizers of convex integral functionals, Bregman distance, Pythagorean identities. *Kybernetika* (2012) 48:637–89.

35. Araki H. Relative entropy for states of von Neumann algebras II. *Publ Rims, Kyoto Univ* (1977) 13:173–92.

36. Naudts J, Verbeure A, Weder R. Linear response theory and the KMS condition. *Comm Math Phys* (1975) 44:87–99. doi:10.1007/bf01609060

37. Eguchi S. Information geometry and statistical pattern recognition. *Sugaku Expositions* (2006) 19:197–216.

38. Kubo R. Statistical-Mechanical theory of irreversible processes. I General theory and simple applications to magnetic and conduction problems. *J Phys Soc Jpn* (1957) 12:570–86. doi:10.1143/jpsj.12.570

39. Mori H. Transport, collective motion, and Brownian motion. *Progr Theor Phys* (1965) 33:423–55. doi:10.1143/ptp.33.423