Quantum entropy inequalities and reversible quantum Markov semigroups as gradient flow for quantum relative entropy

Eric Carlen

Rutgers University

Harvard, November 1, 2022

Quantum Markov semigroups – the physical origins

In 1974, Brian Davies showed how quantum Markov semigroups naturally arise in physics.

Consider a physical system on Hilbert space $\mathcal{H}$ with the Hamiltonian H that is in contact with a heat bath at inverse temperature β. The heat bath is taken to be an infinite free fermion system. Let $\mathcal{H}_B$ and H_B be the heat bath Hilbert space and Hamiltonian. The combined system lives on the Hilbert space

$$\mathcal{K} := \mathcal{H} \otimes \mathcal{H}_B$$

and for $\lambda > 0$, we take the Hamiltonian to be

$$K^\lambda := H \otimes \mathbb{1} + \mathbb{1} \otimes H_B + \lambda V$$

where V is an interaction term on which certain reasonable restrictions are imposed.

Let $\mathfrak{S}$ denote the density matrices on $\mathcal{H}$, and let $\mathfrak{S}_{\mathcal{K}}$ denote the density matrices on $\mathcal{K}$. Let $|\Omega\rangle$ be the vacuum state for the free fermi system, and imbed $\mathfrak{S}$ into $\mathfrak{S}_{\mathcal{K}}$ using the map

$$\rho \mapsto e^{-i(t/\lambda^2)K^0} \left(e^{i(t/\lambda^2)K^\lambda} \rho \otimes |\Omega\rangle\langle\Omega| e^{-i(t/\lambda^2)K^\lambda} \right) e^{i(t/\lambda^2)K^0} =: \rho_{\mathcal{K},\lambda,t} \ .$$

We recover a density matrix $\rho_{t,\lambda}$ on $\mathcal{H}$ by taking the partial trace over $\mathcal{K}$. That is, for a self-adjoint operator A on $\mathcal{H}$,

$$\mathrm{Tr}[\rho_{t,\lambda}A] = \mathrm{Tr}\left[\rho_{\mathcal{K},\lambda,t}(A \otimes \mathbb{1})\right] \ .$$

Davies proved that under reasonable conditions,

$$\lim_{\lambda \downarrow 0} \rho_{t,\lambda} = \mathscr{P}_t^\dagger \rho \quad \text{where} \quad \mathscr{P}_t^\dagger = e^{t\mathscr{L}^\dagger}$$

and he provided an explicit formula for the generator $\mathscr{L}^\dagger$.

Quantum operations

Let $\Phi : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ be a linear transformation. Equip $M_n(\mathbb{C})$ with the Hilbert-Schmidt inner product, making it a Hilbert space. We write $\Phi^\dagger$ to denote the adjoint with respect to the Hilbert-Schmidt inner product.

The map Φ is **unital** if it takes the identity to the identity; i.e.,

$$\Phi(\mathbb{1}_m) = \mathbb{1}_n \ .$$

The map Φ is **trace preserving** if

$$\mathrm{Tr}[\Phi(X)] = \mathrm{Tr}[X] \ .$$

for all X.
It is easy to check that Φ is unital if and only if $\Phi^\dagger$ is trace preserving.

Φ is *positive* when $\Phi(A) \geq 0$ for all $A \geq 0$. Φ is 2-positive if the block matrix $\begin{bmatrix} \Phi(A) & \Phi(B) \\ \Phi(C) & \Phi(D) \end{bmatrix} \geq 0$ whenever $\begin{bmatrix} A & B \\ C & D \end{bmatrix} \geq 0$. For each integer $k > 2$, the condition of k-positivity is defined in the analogous manner, and Φ is *completely* positive if it is k positive for all k.

Example: For any $m \times n$ matrix V, consider the map $\Phi : X \mapsto V^* X V$.

$$\begin{bmatrix} \Phi(A) & \Phi(B) \\ \Phi(C) & \Phi(D) \end{bmatrix} = \begin{bmatrix} V & 0 \\ 0 & V \end{bmatrix}^* \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & V \end{bmatrix}$$

Hence Φ is completely positive, and it follows that for any $\{V_1, \dots, V_\ell\}$ so is

$$\Phi : X \mapsto \sum_{j=1}^{\ell} V_j^* X V_j \ .$$

If Φ is completely positive and unital, its adjoint $\Phi^\dagger$ is a completely positive trace preserving map, and so evidently it takes density matrices to density matrices. Such maps $\Phi^\dagger$ are known as **quantum operations**.

Example: Let $m, n \in \mathbb{N}$. We may think of matrices in $M_{mn}(\mathbb{C})$ as $m \times m$ block matrices with entries in $M_n(\mathbb{C})$. Define

$$\Xi_m(X) = \begin{bmatrix} X & & \\ & \ddots & \\ & & X \end{bmatrix} ,$$

where the matrix on the right is the $m \times m$ block diagonal matrix each of whose diagonal entries is X. This is completely positive and unital. Its adjoint, $\Xi_m^\dagger$, is

$$\Xi_m^\dagger \left(\begin{bmatrix} X_{1,1} & \cdots & X_{1,m} \\ \vdots & \ddots & \vdots \\ X_{m,1} & \cdots & X_{m,m} \end{bmatrix} \right) = \sum_{j=1}^m X_{j,j} ,$$

is the *partial trace*.

The structure of quantum operations

Every quantum operation $\Phi^\dagger$ on $M_n(\mathbb{C})$ has a **Stinespring factorization**:

Define $\Phi_m : M_n(\mathbb{C}) \to M_{mn}(\mathbb{C})$ by

$$\Psi_m(A) = \begin{bmatrix} A & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} = A \otimes |1\rangle\langle 1| \ .$$

Then for some unitary U on $M_{mn}(\mathbb{C})$, and some m,

$$\begin{aligned} \Phi^\dagger(X) &= \Xi_m^\dagger(U^*\Psi_m(X)U) \\ &= \sum_{j=1}^m U_{1,j}^* X U_{1,j} = \sum_{j=1}^m V_j^* X V_j \ . \end{aligned}$$

Definition

A Quantum Markov Semigroup (QMS) is a semigroup $\{\mathscr{P}_t\}_{t\geq 0}$ of completely positive unital maps on $M_n(\mathbb{C})$. Its dual $\{\mathscr{P}_t^\dagger\}_{t\geq 0}$ is a semigroup of completely positive trace preserving maps. Such semigroups are called *Quantum Dynamical Semigroups*.

For any QMS $\{\mathscr{P}_t\}_{t\geq 0}$, and any $t > 0$ since $P_t\mathbb{1} = \mathbb{1}$, 1 is an eigenvalue of each $\mathscr{P}_t$, and hence of each $\mathscr{P}_t^\dagger$. By a simple Perron-Frobenius argument, there will be invariant density matrices σ for the dynamical semigroup $\{\mathscr{P}_t\}_{t\geq 0}$; i.e., density matrices σ such that $\mathscr{P}_t\sigma = \sigma$ for all $t \geq 0$.

We will generally be interested in the **ergodic case** in which there is exactly one such state σ, as in the cases treated by Davies, where it is

$$\sigma_\beta := \frac{1}{Z_\beta} e^{-\beta H} \ , \quad Z_\beta := \mathrm{Tr}[e^{-\beta H}] \ .$$

In this finite dimensional setting in the ergodic case with unique invariant state σ, soft arguments provide a spectral gap, from which it follows that for any $\rho \in \mathfrak{S}$,

$$\lim_{t\to\infty} \mathscr{P}_t^\dagger \rho = \sigma$$

with exponentially fast convergence in the Hilbert-Schmidt norm.

We would like to quantify the exponential rate of convergence as precisely as possible and in the more meaningful **trace norm**

$$\|\mathscr{P}_t^\dagger \rho - \sigma\|_1 = \mathrm{Tr}[|\mathscr{P}_t^\dagger \rho - \sigma|] \ .$$

We are mainly interested in methods for doing this when the system $(\mathcal{H}, H)$ that is in contact with the heat bath consists of N interacting particles so that $\mathcal{H}$, while finite dimensional, has a dimension that is exponential in N. It is natural to make use of **quantum relative entropy**.

Quantum relative entropy

Given two density matrices ρ and σ, the **quantum relative entropy of ρ with respect to σ** is the quantity

$$D(\rho||\sigma) := \mathrm{Tr}[\rho(\log \rho - \log \sigma)] \ .$$

It has a direct physical meaning. Loosely speaking the number of experiments required to distinguish the state ρ from σ with an error probability of size ϵ is $-\log \epsilon / D(\rho||\sigma)$, in a complete analogy with "hypothesis testing" in classical probability.

In this sense, the relative entropy is a good measure of the "divergence"' of ρ from σ. While the relative entropy is not a metric – $D(\rho||\sigma)$ is not symmetric in ρ and σ – it does dominate the square of a natural metric, as expressed by the quantum **Pinsker inequality**

$$D(\rho||\sigma) \geq \frac{1}{2}\mathrm{Tr}[|\rho - \sigma|]^2 \ .$$

Some fundamental entropy inequalities

In 1973, Elliott Lieb proved two convexity inequalities that are fundamental for what we discuss here, but also in many other settings.

Theorem (Lieb 1973, Lieb Concavity Theorem)

For $0 \leq t \leq 1$, and any fixed $K \in M_n(\mathbb{C})$, the function

$$(X, Y) \mapsto \mathrm{Tr}[K^* Y^t K X^{1-t}]$$

is jointly concave on $M_n^+(\mathbb{C}) \times M_n^+(\mathbb{C})$.

Theorem (Lieb 1973)

$$(X, Y, K) \mapsto \mathrm{Tr}\left[\int_0^\infty K^* \frac{1}{sI + Y} K \frac{1}{sI + X} \mathrm{d}s\right]$$

is jointly convex on $M_n^{++}(\mathbb{C}) \times M_n^{++}(\mathbb{C}) \times M_n(\mathbb{C})$.

The first of these theorems in known as the **Lieb Concavity Theorem (LCT)**. The second does not have a name, and its direct significance was overlooked for many years.

In 1975, Lindblad proved, as a consequence of the LCT, the **Data Processing Inequality (DPI)**

Theorem (Lindblad 1975, Data Processing Inequality)

For any two $\rho, \sigma \in \mathfrak{S}$, and any quantum operation $\Phi^\dagger$ on $\mathfrak{S}$,

$$D(\Phi^\dagger \rho || \Phi^\dagger \sigma) \leq D(\rho || \sigma) \ .$$

Roughly speaking, performing any quantum operation on ρ and σ can only make then harder to distinguish. This inequality is one of the cornerstones of quantum information theory.

The DPI is a prototypical **monotonicity theorem**. The LCT is a prototypical **convexity theorem**. The passage back and forth between monotonicity and convexity is crucial to our subject, and therefore we briefly explain how LCT implies DPI.

In the special case $K = 1$, the LCT asserts the convexity of

$$\frac{\mathrm{Tr}[X] - \mathrm{Tr}[Y^t X^{1-t}]}{t}$$

for all $t > 0$. Taking the limit $t \downarrow 0$ yields the joint convexity of

$$(X, Y) \mapsto D(X||Y) \ .$$

Next, recall the Stinespring factorization

$$\Phi^\dagger(X) = \Xi_m^\dagger(U^*\Psi_m(X)U) \ .$$

It turns out Ξ_m can be written as an average over unitary conjugations, and this allows the convexity to be applied. Evidently unitary conjugations have no effect on the relative entropy, and neither does the initial embedding Ψ_m.

In exactly the same way, one can prove monotonicity versions of the two theorems of Lieb introduced above, and these are:

Theorem (Uhlmann 1977)

For all $0 \leq t \leq 1$, all $m, n \in \mathbb{N}$, all $X, Y \in M_m^+(\mathbb{C})$, all $K \in M_n(\mathbb{C})$, and all completely positive unital maps $\Phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$,

$$\mathrm{Tr}[\Phi(K^*)Y^{1-t}\Phi(K)X^t] \leq \mathrm{Tr}[K^*\Phi^\dagger(Y)^{1-t}K\Phi^\dagger(X)^t] \ ,$$

Theorem (Petz 1996)

For all $0 \leq t \leq 1$, all $m, n \in \mathbb{N}$, all $X, Y \in M_m^{++}(\mathbb{C})$, all $K \in M_m(\mathbb{C})$ and all completely positive unital maps $\Phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$,

$$\mathrm{Tr}\left[\int_0^\infty \Phi^\dagger(K^*)\frac{1}{sI + \Phi^\dagger(Y)}\Phi^\dagger(K)\frac{1}{sI + \Phi^\dagger(X)}\mathrm{d}s\right] \leq \mathrm{Tr}\left[\int_0^\infty K^*\frac{1}{sI + Y}K\frac{1}{sI + X}\mathrm{d}s\right] \ .$$

In the form stated here, Uhlmann's Theorem is equivalent to LCT, but Uhlmann actually proved more: He showed that is monotonicity inequality is true for all unital maps Φ that satisfy the Schwarz inequality

$$\Phi(X^*X) \geq \Phi(X^*)\Phi(X) \ .$$

This class of maps strictly includes the class of unital completely positive maps. This class of maps is the best possible.

Other monotonicity theorems, some now and some extended, may be found in Carlen 2022, Carlen and Müller-Hermes 2022, and Carlen and Zhang 2022. However, in this talk we need only the two inequalities discussed here, and only for quantum operations and not more general classes of positive maps.

Monotonic decrease of the relative entropy

Consider a quantum dynamical semigroup $\{\mathscr{P}_t^\dagger\}_{t\geq 0}$ with invariant state σ. Then by the invariance of σ and the DPI, for any density matrix ρ.

$$D(\mathscr{P}_t^\dagger\rho||\sigma) = D(\mathscr{P}_t^\dagger\rho||\mathscr{P}_t^\dagger\sigma) \leq D(\rho||\sigma)\ .$$

That is, in complete generality, the function

$$t \mapsto D(\mathscr{P}_t^\dagger\rho||\sigma)$$

is monotone decreasing. The research discussed in the rest of this lecture is motivated by the following questions:

(1) Under what circumstances can we write the evolution $\rho \mapsto \mathscr{P}_t^\dagger \rho$ as gradient for for the relative entropy with respect to some metric on the space $\mathfrak{S}$?

The gradient flow equation, in $\mathbb{R}^n$ say, is $\dot{x}(t) = -\nabla F(x(t))$, hence

$$\frac{\mathrm{d}}{\mathrm{d}t} F(x(t)) = \nabla F(x(t)) \cdot \dot{x}(t) = -|\nabla F(x(t))|^2 < 0 \ .$$

(2) When it is possible to write the evolution $\rho \mapsto \mathscr{P}_t^\dagger \rho$ as gradient for the relative entropy with respect to some metric on the space $\mathfrak{S}$, under what circumstances can we explicitly find such a metric on $\mathfrak{S}$ and relate the rate at which $\lim_{t\to\infty} D(\mathscr{P}_t^\dagger \rho || \sigma) = 0$ to geometric properties of the metric? Is the rate of convergence exponential?

To clarify the last question, again consider gradient flow in $\mathbb{R}^n$ specified by $\dot{x}(t) = -\nabla F(x(t))$ where F is a twice differentiable function such that

$$D^2 F(x) \geq 2\lambda \mathbb{1}$$

for all x. Then for any solution,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla F(x(t))\|^2 = 2\nabla F(x(t)) \cdot D^2 F(x(t)) \nabla F(x(t)) \leq -2\lambda \|\nabla F(x(t))\|^2 \ .$$

Hence $\|\nabla F(x(t))\|^2 \leq e^{-2\lambda t}\|\nabla F(x(0))\|^2$ and then assuming the minimum value of F is 0, and $x(0) = x$,

$$F(x) = -\int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t} F(x(t))\mathrm{d}t \leq \frac{1}{2\lambda}\|\nabla F(x)\|^2$$

and

$$F(x(t)) \leq e^{-2\lambda t} F(x(0)) \ .$$

It turns out that a QMS does not always describe gradient flow for relative entropy with respect to a Riemannian metric on $\mathfrak{S}$. The following theorem, proved by myself and Jan Maas, refers to a particular self-adjointness condition that will be explained next.

Theorem (Carlen-Mass 2020)

Let $(\mathscr{P}_t)_{t\geq 0}$ be an ergodic QMS with generator $\mathscr{L}$ and invariant state $\sigma \in \mathfrak{S}$. If there exists a continuously differentiable Riemannian metric g_ρ on $\mathfrak{S}$ such that the quantum master equation $\frac{\partial}{\partial t}\rho = \mathscr{L}^\dagger \rho$ is the gradient flow equation for $D(\rho||\sigma)$ with respect to g_ρ, then each $\mathscr{P}_t$ is self-adjoint with respect to the BKM inner product associated to σ.

We now introduce a family of inner products on $M_n(\mathbb{C})$ including the BKM inner product.

Inner products on $M_n(\mathbb{C})$ associated to $\sigma \in \mathfrak{S}$

Let $\mathcal{P}[0,1]$ denote the set of probability measures on the interval $[0,1]$. Notice that for each $s \in [0,1]$,

$$\mathrm{Tr}[B^*\sigma^{1-s}A\sigma^s] = \mathrm{Tr}[(\sigma^{(1-s)/2}B\sigma^{s/2})^*\sigma^{(1-s)/2}A\sigma^{s/2}] \ ,$$

and the right hand side is strictly positive when $B = A \neq 0$.

Definition

For each $m \in \mathcal{P}[0,1]$, $\langle \cdot, \cdot \rangle_m$ denotes the inner product on $M_n(\mathbb{C})$ given by

$$\langle B, A\rangle_m = \mathrm{Tr}[B^*\mathscr{M}_m(A)] \quad \text{where} \quad \mathscr{M}_m(A) = \int_0^1 \sigma^s A \sigma^{1-s}\, \mathrm{d}m(s) \ .$$

The *Gelfand-Naimark-Segal* (GNS) inner product corresponds to $m = \delta_0$, the point mass at $s = 0$.

Other cases are known by name. Taking $m = \delta_{1/2}$ yields the *Kubo-Martin-Schwinger* (KMS) inner product

$$\langle B, A\rangle_{KMS} = \mathrm{Tr}[B^*\sigma^{1/2}A\sigma^{1/2}] \ .$$

Taking m to be uniform on $[0,1]$ yields the *Bogoliubov-Kubo-Mori* (BKM) inner product. The BKM inner product is defined by

$$\langle B, A\rangle_{BKM} = \int_0^1 \mathrm{Tr}[B^*\sigma^s A\sigma^{1-s}]\mathrm{d}s \ .$$

The condition that an operator $\mathscr{L}$ on $M_n(\mathbb{C})$ be self adjoint with respect to the GNS inner product is is quite restrictive, so that such an operator is automatically self-adjoint with respect to any of the other inner products $\langle\cdot,\cdot\rangle_m$. Such a QMS generator is said to satisfy the **detailed balance condition**.

Alicki proved the following structure theorem:

Theorem (Alicki 1976, Structure of QMS with detailed balance)

Let $\mathscr{P}_t = e^{t\mathscr{L}}$ be a quantum Markov semigroup on $M_n(\mathbb{C})$ satisfying detailed balance with respect to $\sigma \in \mathfrak{S}$. Then the generator $\mathscr{L}$ and its adjoint $\mathscr{L}^\dagger$ have the form

$$\mathscr{L} = \sum_{j\in\mathcal{J}} e^{-\omega_j/2}\mathscr{L}_j\ , \qquad \mathscr{L}_j(A) = V_j^*[A, V_j] + [V_j^*, A]V_j\ ,$$

$$\mathscr{L}^\dagger = \sum_{j\in\mathcal{J}} e^{-\omega_j/2}\mathscr{L}_j^\dagger\ , \qquad \mathscr{L}_j^\dagger(\rho) = [V_j, \rho V_j^*] + [V_j\rho, V_j^*]\ ,$$

where $\mathcal{J}$ is a finite index set, the operators $V_j \in B(\mathscr{H})$ satisfy $\{V_j\}_{j\in\mathcal{J}} = \{V_j^*\}_{j\in\mathcal{J}}$, and $\omega_j \in \mathbb{R}$ satisfies

$$\sigma V_j \sigma^{-1} = e^{-\omega_j} V_j \quad \text{for all } j \in \mathcal{J}\ .$$

Gradient flow on $\mathfrak{S}$

$\mathfrak{S}$ is a relatively open subset of $\{A \in M_n^+(\mathbb{C}) \ : \mathrm{Tr}[A] = 1\}$. Identify the tangent space $T_\rho\mathfrak{S}$ at $\rho \in \mathfrak{S}$ with

$$\mathcal{V} := \{A \in M_n(\mathbb{C}) : A = A^* \quad \text{and} \quad \mathrm{Tr}[A] = 0\} \ .$$

The cotangent space $T_\rho^\dagger\mathfrak{S}$ may also be identified with $\mathcal{V}$ through the duality pairing $\langle A, B\rangle = \mathrm{Tr}[AB]$ for $A, B \in \mathcal{A}_0$.

Let g_ρ be a smooth Riemannian metric on $\mathfrak{S}$. Then g_ρ determines an operator $\mathscr{G}_\rho : T_\rho\mathfrak{S} \to T_\rho^\dagger\mathfrak{S}$ defined by

$$\langle A, \mathscr{G}_\rho B\rangle_{\mathfrak{H}} = g_\rho(A, B)$$

for $A, B \in T_\rho\mathfrak{S}$. Clearly, $\mathscr{G}_\rho$ is invertible and self-adjoint with respect to the Hilbert–Schmidt inner product on $\mathfrak{H}$. Define $\mathcal{K}_\rho : T_\rho^\dagger\mathfrak{S} \to T_\rho\mathfrak{S}$ by $\mathcal{K}_\rho = (\mathscr{G}_\rho)^{-1}$.

For a smooth functional $F : \mathfrak{S} \to \mathbb{R}$ and $\rho \in \mathfrak{S}$, its **differential** $\mathrm{d}\mathcal{F}(\rho) \in T_\rho^\dagger \mathfrak{S}$ is defined by $\lim_{\varepsilon \to 0} \varepsilon^{-1}(F(\rho + \varepsilon A) - F(\rho)) = \langle A, \mathrm{d}F(\rho)\rangle$ for $A \in T_\rho \mathfrak{S}$. The **gradient** $\nabla_g F(\rho) \in T_\rho \mathfrak{S}$ depends on the Riemannian metric through the duality formula $g_\rho(A, \nabla_g F(\rho)) = \langle A, \mathrm{d}\mathcal{F}(\rho)\rangle$ for $A \in T_\rho \mathfrak{S}$. That is

$$\nabla_g F(\rho) = \mathcal{K}_\rho \mathrm{d}F(\rho) \ .$$

For $F(\rho) = D(\rho||\sigma)$, $dF(\rho) = \log \rho - \log \sigma$. If $\mathscr{L}$ is a QMS generator and $\sigma \in \mathfrak{S}$ satisfies $\mathscr{L}^\dagger \sigma = 0$, we seek a metric, specified by some $\mathcal{K}_\rho$, such that

$$\mathscr{L}^\dagger \rho = \mathcal{K}_\rho(\log \rho - \log \sigma) \ .$$

A classical model

Jordan, Kinderlehrer and Otto showed that

$$\begin{aligned}\frac{\partial}{\partial t}\rho(x,t) &= \Delta\rho(x,t) - \nabla\cdot(\rho(x,t)\nabla\log\sigma(x))\\ &= \nabla\cdot(\rho(x,t)(\nabla\log\rho(x) - \nabla\log\sigma(x)))\end{aligned}$$

describes gradient flow for the classical relative entropy $\int_{\mathbb{R}^n}\rho(x)\log\frac{\rho(x)}{\sigma(x)}\,\mathrm{d}x$ with respect to the 2-Wasserstein metric.

This structure is formally given in terms of the operator K_ρ defined by

$$K_\rho\psi = -\nabla\cdot(\rho\nabla\psi)\ ,$$

for probability densities ρ on $\mathbb{R}^n$. We get a **linear equation** since

$$\rho\nabla\log\rho = \nabla\rho\ .$$

Non commutative differential structure

Consider a QMS generator $\mathscr{L}$, satisfying detailed balance, as specified by σ and $\{V_1, \ldots, V_{|\mathcal{J}|}\}$. Define

$$\partial_j A = [V_j, A] \ .$$

Define the Hilbert space $\mathfrak{H}_{\mathcal{J}}$ by $\mathfrak{H}_{\mathcal{J}} = \bigoplus_{j \in \mathcal{J}} \mathfrak{H}_j$, where each $\mathfrak{H}_j$ is a copy of $\mathfrak{H}$. For $\mathbf{A} \in \mathfrak{H}_{\mathcal{J}}$ and $j \in \mathcal{J}$, let A_j denote the component of $\mathbf{A}$ in $\mathfrak{H}_j$. Pick some linear ordering of $\mathcal{J}$, and write

$$\mathbf{A} = (A_1, \ldots, A_{|\mathcal{J}|}) \ .$$

Equip $\mathfrak{H}_{\mathcal{J}}$ with the inner product $\langle \mathbf{A}, \mathbf{B} \rangle_{\mathfrak{H}_{\mathcal{J}}} = \sum_{j \in \mathcal{J}} \langle A_j, B_j \rangle_{\mathfrak{H}_j}$.

Define an operator $\nabla : \mathfrak{H} \to \mathfrak{H}_{\mathcal{J}}$ by

$$\nabla A = (\partial_1, \dots, \partial_{|\mathcal{J}|} A) \ .$$

Thinking of elements of $\mathfrak{H}$ as non-commutative analogs of functions on a manifold, we may think of $\mathbf{A} = (A_1, \dots, A_{|\mathcal{J}|})$ as a vector field. We define the operator div : $\mathfrak{H}_{\mathcal{J}} \to \mathfrak{H}$ by

$$\operatorname{div} \mathbf{A} = -\sum_{j \in \mathcal{J}} \partial_j^\dagger A_j = \sum_{j \in \mathcal{J}} [A_j, V_j^*] \ .$$

Note that div is minus the adjoint of the map $\nabla : \mathfrak{H} \to \mathfrak{H}_{\mathcal{J}}$, so that $\mathscr{L}_0$ is negative semi-definite. We call ∇ and div the **non-commutative gradient** and the **non-commutative divergence** associated to $\mathscr{L}$,

The chain rule

Define an operator $\mathbb{M}_{\rho,j}$ by

$$\mathbb{M}_{\rho,j}(A) = \int_0^1 \left(e^{\omega_j/2}\rho\right)^{1-s} A \left(e^{-\omega_j/2}\rho\right)^s \,\mathrm{d}s \ .$$

so that

$$\mathbb{M}_{\rho,j}^{-1}(A) = \int_0^\infty \frac{1}{\lambda + e^{-\omega/2}\rho} A \frac{1}{\lambda + e^{\omega/2}\rho} \,\mathrm{d}\lambda \ .$$

Note that if ρ and A commute, and $\omega_j = 0$,

$$\mathbb{M}_{\rho,j}(A) = \rho A \ .$$

Finally,

$$\mathbb{M}_\rho(\mathbf{A}) := (\mathbb{M}_{\rho,1}(A_1), \dots, \mathbb{M}_{\rho,|\mathcal{J}|}(A_{|\mathcal{J}|})$$

Lemma (Chain rule for the logarithm)

For all $\rho \in \mathfrak{S}$ and $j \in \mathcal{J}$ we have

$$e^{-\omega_j/2} V_j \rho - e^{\omega_j/2} \rho V_j = \mathbb{M}_{\rho,j} \left(\partial_j (\log \rho - \log \sigma) \right) \ .$$

We then define

$$\mathcal{K}_\rho A := \sum_{j \in \mathcal{J}} \partial_j^\dagger \left(\mathbb{M}_{\rho,j} (\partial_j A) \right) = \mathrm{div} \left(\mathbb{M}_{\rho,j} \nabla A \right) \ .$$

Theorem (Carlen-Maas 2017)

For $\rho \in \mathfrak{S}$ we have the identity

$$\mathscr{L}^\dagger \rho = -\mathcal{K}_\rho \mathrm{d} D(\rho || \sigma) \ ,$$

hence the gradient flow equation of $D(\rho||\sigma)$ with respect to the Riemannian metric induced by $\mathcal{K}_\rho$ is the master equation $\partial_t \rho = \mathscr{L}^\dagger \rho$.

Also in 2017, Mielke and Mittenzwerg constructed a metric, by different means, for which the master equation $\frac{\partial}{\partial t}\rho = \mathscr{L}^\dagger \rho$, again under the detailed balance condition, is gradient flow for the relative entropy.

The metric is not unique, and Carlen-Maas 2020 gives a more flexible construction yielding a range of such metrics.

We now explain how the explicit form of our metric allows one to prove entropy-entropy production inequalities. To do this, we write our metric in Brenier-Benamou form:

Connection with the Brenier-Benamou formula

The **Brenier-Benamou formula** for the 2-Wasserstein distance on is:

$$
\begin{aligned}
&W_2(\rho_0,\rho_1)^2 \\
&= \inf\left\{ \int_0^1 |\nabla\psi_t(x)|^2\,\mathrm{d}\rho_t(x)\,\mathrm{d}t \ :\ \partial_t\rho_t + \mathrm{div}(\rho_t\nabla\psi_t) = 0\ ,\rho_t|_{t=0,1} = \rho_{0,1} \right\} \\
&= \inf\left\{ \int_0^1 \frac{|P_t(x)|^2}{\rho_t(x)}\,\mathrm{d}x\,\mathrm{d}t \ :\ \partial_t\rho_t + \mathrm{div}\,P_t = 0\ ,\rho_t|_{t=0,1} = \rho_{0,1} \right\}\ .
\end{aligned}
$$

In our case, for a smooth curve $\rho(t)$, we have that

$$
g_\rho(\dot\rho,\dot\rho) = \langle \dot\rho, \mathcal{K}_\rho^{-1}\dot\rho\rangle_{\mathfrak{H}} = \langle (\mathcal{K}_\rho^{-1}\dot\rho), \mathcal{K}_\rho(\mathcal{K}_\rho^{-1}\dot\rho)\rangle_{\mathfrak{H}}\ .
$$

That is, writing $\dot\rho(t) =: \mathcal{K}_{\rho(t)}B(t)$, $g_\rho(\dot\rho,\dot\rho) = \langle B, \mathcal{K}_\rho B\rangle$.

Therefore

$$d_g^2(\rho_0, \rho_1) =$$
$$\inf \left\{ \int_0^1 \langle B_t, \mathcal{K}_{\rho_t} B_t \rangle \, \mathrm{d}t \ : \ \partial_t \rho_t = \mathcal{K}_{\rho_t} B_t \ , \rho_t|_{t=0} = \rho_0 \ , \rho_t|_{t=1} = \rho_1 \right\}$$

Making the change of variables,

$$\mathbf{A_t} := \mathbb{M}_{\rho(t)} \nabla B_t \ ,$$

we get, using a minimizing property of gradients, that

$$d_g^2(\rho_0, \rho_1) =$$
$$\inf \left\{ \int_0^1 \langle \mathbf{A}_t, \mathbb{M}_{\rho(t)}^{-1} \mathbf{A}_t \rangle \, \mathrm{d}t \ : \ \partial_t \rho_t + \mathrm{div}\, \mathbf{A}_t = 0 \ , \rho_t|_{t=0} = \rho_0 \ , \rho_t|_{t=1} = \rho_1 \right\}$$

which is our analog of the Brenier-Benamou formula.

Recall that

$$\langle \mathbf{A}_t, \mathbb{M}_{\rho(t)}^{-1}\mathbf{A}_t\rangle = \sum_{j\in\mathcal{J}}\int_0^\infty \mathrm{Tr}\left[(A_t)_j^* \frac{1}{\lambda + e^{-\omega_j/2}\rho}(A_t)_j \frac{1}{\lambda + e^{\omega_j/2}\rho}\right]\, \mathrm{d}\lambda\ ,$$

and the second monotonicity theorem, equivalent ot the second convexity theorem of Lieb, now applies to each summand. This will be the key to proving entropy-entropy production inequalities.

Geodesic convexity and relaxation to equilibrium

Let $(\mathscr{M}, g)$ be any smooth, finite-dimensional Riemannian manifold. For x, y in $\mathscr{M}$, the Riemannian distance $d_g(x,y)$ between x and y is given by minimizing an action integral of paths $\gamma : [0,1] \to \mathscr{M}$ running from x to y:

$$d_g^2(x,y) = \inf\left\{ \int_0^1 \|\dot\gamma(s)\|^2_{g(\gamma(s))}\,\mathrm{d}s \ :\ \gamma(0) = x,\ \gamma(1) = y \right\} ,$$

where

$$\|\dot\gamma(s)\|^2_{g(\gamma(s))} = g_{\gamma(s)}(\dot\gamma(s), \dot\gamma(s)) .$$

(If the infimum is achieved, any minimizer γ will be a geodesic.) If F is a smooth function on $\mathscr{M}$, let $\mathrm{grad}_g F$ denote its Riemannian gradient.

For $\lambda \in \mathbb{R}$, the function F is λ-*convex* in case whenever $\gamma : [0,1] \to \mathscr{M}$ is a distance minimizing geodesic, then for all $s \in (0,1)$,

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2} F(\gamma(s)) \geq \lambda g(\dot\gamma(s), \dot\gamma(s)) \ .$$

It is a standard result that whenever F is λ-convex, the gradient flow for F is λ-contracting in the sense that for all $x, y \in \mathscr{M}$ and $t > 0$,

$$\frac{\mathrm{d}}{\mathrm{d}t} d_g^2(S_t(x), S_t(y)) \leq -2\lambda d_g^2(S_t(x), S_t(y)) \ .$$

Otto and Westdickenberg developed an approach to geodesic convexity that takes this last conclusion as its starting point. Consider the semigroup S_t of transformations on $\mathscr{M}$ given by solving $\dot\gamma(t) = -\mathrm{grad}_g F(\gamma(t))$; we assume that nice global solutions exist, which will be the case in our application. The semigroup S_t, $t \geq 0$, is gradient flow for F.

Let $\{\gamma(s)\}_{s\in[0,1]}$ be any smooth path in $\mathscr{M}$ with $\gamma(0) = x$ and $\gamma(1) = y$. They use the gradient flow transformation S_t to define a one-parameter family of paths $\gamma^t : [0,1] \to \mathscr{M}$, $t \geq 0$ defined by

$$\gamma^t(s) = S_t\gamma(s) \ .$$

Since γ^t is admissible for the variational problem that defines $d_g(S_t(x), S_t(y))$, it is immediate that for each $t \geq 0$,

$$d_g^2(S_t(x), S_t(y)) \leq \int_0^1 \|\dot{\gamma}(s)\|^2_{g(\gamma(s))}\, \mathrm{d}s \ .$$

In the present smooth setting it is shown by Danieri and Savaré that if for all smooth curves $\gamma : [0,1] \to \mathscr{M}$,

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{0+} \left(\|\dot{\gamma}^t(s)\|^2_{g(\gamma^t(s))}\right) \leq -2\lambda\|\dot{\gamma}^0(s)\|^2_{g(\gamma^0(s))} \ ,$$

for all $s \in (0,1)$, then F is geodetically λ-convex.

If one wishes to check convexity directly, one needs to take two derivatives. A direct brute force computation is difficult at best, and nobody has succeeded for physically interesting cases. The problem is that one does not have good expressions for the geodesics of the metric.

However, the theorem of Otto and Westdickeberg reduces the problem to one of monotonicity – one only needs to check one derivative. Even this would be difficult to do by brute force in interesting cases, but fortunately this is not necessary.

The quantity $\|\dot\gamma^0(s)\|^2_{g(\gamma^0(s))}$ is closely related to the monotone metric discussed earlier that is provided by a convexity theorem of Lieb. Using this, one can readily prove the required monotonicity inequality in a number of important cases.

We now present a simple sufficient condition for the monotonicity inequality that we shall be able to verify in a number of interesting examples.

Definition

A semigroup $\vec{\mathscr{P}}_t$ on $\mathfrak{H}_{\mathcal{J}}$ *intertwines with* a semigroup $\mathscr{P}_t$ on $\mathfrak{H}$ in case for all $t > 0$, and all $A \in \mathfrak{H}$,

$$\nabla \mathscr{P}_t A = \vec{\mathscr{P}}_t \nabla A \ .$$

By duality, the intertwining relation $\nabla \circ \mathscr{P}_t = \vec{\mathscr{P}}_t \circ \nabla$ implies the identity

$$\mathscr{P}_t^\dagger \operatorname{div}(\mathbf{A}) = \operatorname{div}({\vec{\mathscr{P}}_t}^\dagger \mathbf{A}) \ , \qquad \text{for } \mathbf{A} \in \mathfrak{H}_{\mathcal{J}} \ .$$

We will be particularly interested in cases in which for some $\lambda \in \mathbb{R}$,

$$\vec{\mathscr{P}}_t \mathbf{A} = (e^{-\lambda t} \mathscr{P}_t A_1, \ldots, e^{-\lambda t} \mathscr{P}_t A_{|\mathcal{J}|}) \ .$$

Suppose that the intertwining relation

$$\mathscr{P}_t^\dagger \operatorname{div}(\mathbf{A}) = \operatorname{div}(\vec{\mathscr{P}_t}^\dagger \mathbf{A}) , \qquad \text{for } \mathbf{A} \in \mathfrak{H}_{\mathcal{J}} .$$

with

$$\vec{\mathscr{P}}_t \mathbf{A} = (e^{-\lambda t} \mathscr{P}_t A_1, \ldots, e^{-\lambda t} \mathscr{P}_t A_{|\mathcal{J}|}) .$$

is valid.

To apply it, consider any smooth path $\rho : [0,1] \to \mathfrak{S}$ with $\rho(0) = \rho_0$ and $\rho(1) = \rho_1$. Recall the formula

$$d_g^2(\rho_0, \rho_1) =$$
$$\inf \left\{ \int_0^1 \langle \mathbf{A}_t, \mathbb{M}_{\rho(t)}^{-1} \mathbf{A}_t \rangle \,\mathrm{d}t \ : \ \partial_t \rho_t + \operatorname{div} \mathbf{A}_t = 0 \ , \rho_t\big|_{t=0} = \rho_0 \ , \rho_t\big|_{t=1} = \rho_1 \right\}$$

Set $\rho^t(s) := \mathscr{P}_t^\dagger \rho(s)$. Since we suppose that the semigroup $\vec{\mathscr{P}}_t$ intertwines with $\mathscr{P}_t$. It follows that

$$\frac{\mathrm{d}}{\mathrm{d}s}\rho^t(s) = \mathscr{P}_t^\dagger \,\mathrm{div}\, \mathbf{A}(s) = \mathrm{div}\, \vec{\mathscr{P}}_t^\dagger \mathbf{A}(s) \ .$$

Consequently,

$$\begin{aligned} \left\| \frac{\mathrm{d}}{\mathrm{d}s}\rho^t(s) \right\|_{g(\rho^t(s))}^2 &\leq e^{-2\lambda t} \langle \mathscr{P}_t^\dagger \mathbf{A}(s), \mathbb{M}^{-1}_{\mathscr{P}_t^\dagger \rho(s)} \mathscr{P}_t^\dagger \mathbf{A}(s) \rangle_{\mathfrak{H}} \\ &\leq e^{-2\lambda t} \langle \mathbf{A}(s), \mathbb{M}^{-1}_{\rho(s)} \mathbf{A}(s) \rangle_{\mathfrak{H}} = e^{-2\lambda t} \left\| \frac{\mathrm{d}}{\mathrm{d}s}\rho(s) \right\|_{g(\rho(s))}^2 , \end{aligned}$$

which is the desired monotonicity inequality.

Using this, Jan Maas and I proved:

Theorem

Let $\mathscr{P}_t$ be the Bose Ornstein-Uhlenbeck semigroup with generator $\mathscr{L}_\beta$, and let σ_β be its invariant state. Then for all $\rho \in \mathfrak{S}_+$,

$$D(\mathscr{P}_t\rho||\sigma_\beta) \leq e^{-2\sinh(\beta/2)t} D(\rho||\sigma_\beta) \ .$$

Theorem

For $\beta > 0$, let $\mathscr{P}_t$ be the Fermi Ornstein-Uhlenbeck semigroup with generator $\mathscr{L}_\beta$, and let σ_β be its invariant state. Then for all $\rho \in \mathfrak{S}$,

$$D(\mathscr{P}_t\rho||\sigma_\beta) \leq e^{-2\lambda_\beta t} D(\rho||\sigma_\beta)$$

where $\lambda_\beta = \min\{\cosh(\beta e_j/2) \ : \ j = 1, \dots, m\}$.

Michael Loss and I are now working on probing such entropy production theorems for quantum analogs of the Boltzmann equation, and have succeeded in a number of interesting cases. But much remains to be done!

Thank you for your Interest!