Gradient flow and functional inequalities for quantum Markov semigroups, I

Eric A. Carlen

Rutgers University

MPIM Bonn, September 25, 2017

This talk is based on joint work with Jan Maas.

E. A. Carlen and J. Maas, *An analog of the 2-Wasserstein metric in non-commutative probability under which the fermionic Fokker-Planck equation is gradient flow for the entropy*, Comm. Math. Phys. **331**, (2014), 887–926.

E. A. Carlen and J. Maas, *Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance*, Jour. Func. Analysis, **273**, no. 5, (2017) 1810-1869

Related work Chen, Georgiou, Tannenbaum and also Mielke, Mittnenzweig, in Jour Stat. Phys. (All papers are on arXiv).

The classical starting point

Let $\sigma(x)$ be a smooth strictly positive probability density on R^n. If $\rho(x)$ as any other such density, the relative entropy of ρ with respect to σ is

$$D(\rho||\sigma) = \int_{\mathrm{R}^n} \rho[\log \rho(x) - \log \sigma(x)]\mathrm{d}x \ .$$

The evolution equation

$$\frac{\partial}{\partial t}\rho(x,t) = \nabla \cdot \rho(x,t)[\nabla(\log \rho(x,t) - \log \sigma(x))]$$

is the Kolmogorov forward equation for a diffusion process.

By the chain rule,

$$\rho(x,t)\nabla \log \rho(x,t) = \nabla\rho(x,t)$$

and therefore

$$\nabla \cdot \rho(x,t)[\nabla(\log \rho(x,t) - \log \sigma(x))] = \Delta\rho(x,t) - \nabla \cdot (\rho(x,t)\nabla \log \sigma(x)) \ .$$

However,

$$\nabla(\log \rho - \log \sigma) = \nabla \frac{\delta}{\delta\rho} D(\rho||\sigma)$$

relates the equation to gradient flow for the 2-Wasserstein metric, as Felix Otto observed and exploited.

Since his original work in 2000, this perspective has been found to be very useful, especially for establishing functional inequalities related to the rate of relative entropy dissipation.

Our goal is to extend this approach to the quantum Markov semigroup setting. We will do so, and shall prove a sharp entropy production inequality that had been conjectured by Huber, König and Vershynina, as well as other new inequalities.

The quantum counterpart

Let $\mathcal{A}$ be a finite-dimensional C^*-algebra with unit $\mathbf{1}$. If you like, take $\mathcal{A} = M_n(\mathbb{C})$, the $n \times n$ matrices over $\mathbb{C}$. Let $\mathfrak{S}_+(\mathcal{A})$ denote the set of faithful states of $\mathcal{A}$: In the matricial case this is the set of all positive $n \times n$ matrices ρ with unit trace, and the *state* corresponding to ρ is the linear functional $A \mapsto \mathrm{Tr}[\rho A]$.

A *Quantum Markov Semigroup* (QMS) is a continuous one-parameter semigroup of linear transformations

$$(\mathscr{P}_t)_{t \geq 0}$$

on $\mathcal{A}$ such that for each $t \geq 0$, $\mathscr{P}_t$ is *completely positive* and $\mathscr{P}_t \mathbf{1} = \mathbf{1}$.

Complete positivity means the following: Any linear transformation $\mathscr{K} : \mathcal{A} \to \mathcal{A}$ induces the linear transformation $\mathscr{K} \otimes \mathbf{1}_{M_m(\mathbb{C})}$ from $\mathcal{A} \otimes M_m(\mathbb{C})$ to $\mathcal{A} \otimes M_m(\mathbb{C})$:

$$\mathscr{K} \otimes \mathbf{1}_{M_m(\mathbb{C})} \left(\sum_{i,j=1}^{n} A_{i,j} \otimes E_{i,j} \right) = \sum_{i,j=1}^{m} \mathscr{K}(A_{i,j}) \otimes E_{i,j} \ ,$$

where $E_{i,j}$ is the element of $M_m(\mathbb{C})$ whose i,j entry is one, with all other entries being zero. The map $\mathscr{K}$ is completely positive in case for each $m \in \mathbb{N}$, $\mathscr{K} \otimes \mathbf{1}_{M_m(\mathbb{C})}$ is positivity preserving.

Let $\mathcal{A} = M_n(\mathbb{C})$, and $\mathscr{K} A = A^T$, which is positive.

$$\mathscr{K} \otimes \mathbf{1}_{M_2(\mathbb{C})} : \begin{bmatrix} E_{1,1} & E_{1,2} \\ E_{2,1} & E_{2,2} \end{bmatrix} = \begin{bmatrix} E_{1,1}^T & E_{1,2}^T \\ E_{2,1}^T & E_{2,2}^T \end{bmatrix} .$$

$$\begin{bmatrix} E_{1,1} & E_{1,2} \\ E_{2,1} & E_{2,2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} .$$

Choi 1974: For $\mathcal{A} = M_n(\mathbb{C})$, $\mathscr{K}$ is completely positive iff

$$\mathscr{K} \otimes \mathbf{1}_{M_n(\mathbb{C})} \left(\sum_{i,j=1}^{n} E_{i,j} \otimes E_{i,j} \right) \geq 0 .$$

This notion was introduced by Stinespring and is the meaningful generalization of positivity preservation to the non-commutative case for both physical and mathematical reasons.

Being a stronger condition than positivity preservation, there is much more that one can say about the structure of completely positive operators than one can say about operators that merely preserve positivity. And since the notion arises naturally in quantum mechanics in the consideration of coupled systems, what one can say is physically meaningful.

Let $(\mathscr{P}_t)_{t\geq 0}$ be a QMS acting on $\mathcal{A}$. The algebra $\mathcal{A}$ is the algebra of *observables*. Let $\mathfrak{H}_{\mathcal{A}}$ be the Hilbert space obtained by equipping $\mathcal{A}$ with the Hilbert-Schmidt inner product. A dagger † always denotes the adjoint with respect to the inner product in $\mathfrak{H}_{\mathcal{A}}$.
For all A, B,

$$\mathrm{Tr}[A^* \mathscr{P}_t B] = \mathrm{Tr}[(\mathscr{P}_t^\dagger A)^* B] \ .$$

The semigroup $\mathscr{P}_t$ gives the *Heisenberg picture* of the evolution. The dual semigroup $(\mathscr{P}_t^\dagger)_{t\geq 0}$ acting on $\mathfrak{S}_+(\mathcal{A})$ gives the *Schödinger picture* of the evolution.

The QMS $\mathscr{P}_t$ is *ergodic* in case $\mathbf{1}$ spans the eigenspace of $\mathscr{P}_t$ for the eigenvalue 1. In that case, there is a unique invariant state σ. While σ need not be faithful, a natural projection operation allows us to assume, effectively without loss of generality, that $\sigma \in \mathfrak{S}_+(\mathcal{A})$.

We consider a class of ergodic QMS that satisfy a quantum *detailed balance condition* with respect to their unique invariant state σ.

We show that all such semigroups (in the Schrödinger picture) are *gradient flow for the relative entropy with respect to a natural analog of the* 2*-Wasserstein metric*, and we use this to prove new functional inequalities, one of which proves a recent conjecture of Huber, König and Vershynina.

How QMS arise in practice

One might wonder how QMS arise in practice. After all the basic evolution equation in quantum mechanics is the Schrödinger equation which describes unitary, reversible, non-dissipative time evolution. So one might wonder QMS arise naturally even in quantum mechanics, but also in other fields.

The answer is “yes” on both counts. We discuss two ways they come up in quantum mechanics, focusing in work of Nelson, Gross and Davies, and also one way that they come up in the study of classical heat kernel estimates on graphs, focusing on a paper of Davies.

Let μ be the unit Gaussian measure on R^n. Consider the Dirichlet form

$$\mathcal{E}(\varphi,\varphi) = \int_{\mathrm{R}^n} |\nabla\varphi|^2 \mathrm{d}\mu$$

Then with $\langle\cdot,\cdot\rangle$ denoting the corresponding L^2 inner product,

$$\mathcal{E}(\varphi,\varphi) = \langle\varphi, \mathscr{N}\varphi\rangle$$

where $\mathscr{N}$ is the *number operator*

$$\mathscr{N} = -(\nabla - x)\cdot\nabla\ .$$

The Schödiner operator $\mathscr{H} := \mathscr{N} + \mathscr{V}$ where $\mathscr{V}$ is multiplication by the real function V arises in quantum field theory, and one would like a lower bound on the bottom of the spectrum.

Theorem 0.1 (Nelson's Theorem)**.** *Let V be a measurable function on R^n such that the negative part of V, V_-, is exponentially integrable. Let $H = N + V$, the sum of the number operator and multiplication by V. Then for all $\varphi \in L^2(\gamma)$,*

$$\langle \varphi, \mathscr{H} \varphi \rangle \geq -\frac{1}{2} \log \left(\int_{\mathrm{R}^n} e^{2V_-} \, \mathrm{d}\mu \right) \|\varphi\|_2 \; .$$

The theorem is dimension-free, and extends directly to the infinite dimensional case.
Nelson's proof relied on the *hypercontractivity* of the Mehler semigroup $(\mathscr{P})_{t \geq 0}$ where $\mathscr{P}_t = e^{-t\mathscr{N}}$.

$$\|\mathscr{P}_t \varphi\|_{1+e^{2t}}^2 \leq \|\varphi\|_2^2 \; .$$

Federbrush gave a partially alternate proof: He differentiated the hypercontractivity inequality at $t = 0$ to obtain

$$\int_{\mathrm{R}^n} |\varphi|^2 \log |\varphi|^2 \mathrm{d}\mu - \|\varphi\|_2^2 \log \|\varphi\|_2^2 \leq 2 \int_{\mathrm{R}^n} |\nabla\varphi|^2 \mathrm{d}\mu \ .$$

Let ρ be a probability density with respect to μ. Then taking $\varphi = \sqrt{\rho}$, and defining

$$H(\rho) = \int_{\mathrm{R}^n} \rho \log \rho \mathrm{d}\mu \quad \text{and} \quad I(\rho) = \frac{1}{2} \int_{\mathrm{R}^n} \frac{|\nabla\rho|^2}{\rho} \mathrm{d}\mu,$$

this becomes

$$H(\rho) \leq \frac{1}{2} I(\rho) \ .$$

Both sides of the inequality are convex functions of ρ, the right hand side as a consequence of the joint convexity on $\mathrm{R}^n \times (0,\infty)$ of

$$(a,t) \mapsto \frac{|a|^2}{t} \ .$$

Hence, for any bounded, continuous real function W,

$$\sup_\rho \left\{ \int_{\mathrm{R}^n} W\rho \mathrm{d}\mu - H(\rho) \right\} \geq \sup_\rho \left\{ \int_{\mathrm{R}^n} W\rho \mathrm{d}\mu - \tfrac{1}{2} I(\rho) \right\} \ .$$

An easy computation shows that

$$\sup_\rho \left\{ \int_{\mathbb{R}^n} W\rho \mathrm{d}\mu - H(\rho) \right\} = \log \left(\int_{\mathbb{R}^n} e^W \mathrm{d}\mu \right)$$

and

$$\sup_\rho \left\{ \int_{\mathbb{R}^n} W\rho \mathrm{d}\mu - \tfrac{1}{2} I(\rho) \right\} = -\lambda_{\min}(2\mathscr{N} - \mathscr{W}) \quad .$$

Hence

$$\lambda_{\min}(\mathscr{N} - \tfrac{1}{2}\mathscr{W}) \geq -\frac{1}{2} \log \left(\int_{\mathbb{R}^n} e^W \mathrm{d}\mu \right) \quad .$$

Take $V = -2W$.

In these last few slides we have been discussing a *classical* Markov semigroup $(\mathscr{P}_t)_{t\geq 0}$ – the Mehler semigroup. The operator $\mathscr{N}$ in this case is the boson number operator. It can be written as

$$\mathscr{N} = \sum_{j=1}^{n} a_j^* a_j$$

where

$$a_j = \nabla_j \quad \text{and} \quad a_j^* = -\nabla_j + x_j \ .$$

These satisfy the *canonical commutation relations*:

$$a_j a_k^* - a_k^* a_j = \delta_{j,k}\mathbf{1} \ .$$

For systems of fermions, the physically natural operators are expressed in terms of operators satisfying *canonical anti-commutation relations*:

$$a_j a_k^* + a_k^* a_j = \delta_{j,k} \mathbf{1} \ .$$

The operators a_j may be realized as skew-derivations ∇_j. on a Clifford algebra $\mathcal{A}$ with N generators, $Q_1, \ldots, Q_N$, and the normalized trace τ is an analog of the Gaussian measure, as Segal had emphasized. The Clifford algebra $\mathcal{A}$ comes with a privileged involutive $*$-automorphism Γ, and to say that ∇_j is a skew derivation means that for all $A, B \in \mathcal{A}$,

$$\nabla_j(AB) = (\nabla_j A)B + \Gamma(A)\nabla_j B \ .$$

Let $\langle \cdot, \cdot \rangle$ denote the inner product $\langle A, B \rangle = \tau[A^* B]$. Define *Gross's Clifford Dirichlet form* $\mathcal{E}(A, A)$ by

$$\mathcal{E}(A, A) = \sum_{j=1}^{n} \tau[(\nabla_j A)^* (\nabla_j A)] \ .$$

Define

$$\langle A, \mathscr{N} A \rangle = \mathcal{E}(A, A) \quad \text{and} \quad \mathscr{P}_t = e^{-t\mathscr{N}} \ .$$

Define L^p norms by $\|A\|_p = \tau[(A^* A)^{p/2}]^{2/p}$. Gross conjectured that the exact analog of Nelson's inequality:

$$\|\mathscr{P}_t \varphi\|_{1+e^{2t}}^2 \leq \|\varphi\|_2^2 \ .$$

This was proved by Elliott Lieb and myself in 1993. Gross had proved a slightly weaker result, showing that such in inequality is is true if one replaces $p(t) := 1 + e^{2t}$ by a function of t that grows somewhat more slowly.

Gross used the fact that each $\mathscr{P}_t$ is positivity preserving to prove a uniqueness theorem for ground states, making use of hypercontractivity and a variant of the Perron-Frobenius Theorem.

In all of these investigations, the semigroup $(\mathscr{P}_t)_{t\geq 0}$ enters the discussion because one is trying to study the properties of a Schrödinger operator $\mathscr{H} = \mathscr{N} + \mathscr{V}$. It does not have an direct dynamical significance; the parameter t is not the physical time that appears in the corresponding Schrödinger equation.

QMS do also arise in quantum dynamics: Brian Davies showed in the 1970's that if one coupled a finite quantum system to an infinite fermion "heat bath" and then considers the evolution over long times but with weak coupling, a QMS arises. Specifically, let H_0 be the Hamiltonian of the finite quantum system. Let H_1 be the Hamiltonian for the heat bath. Let λK be an interaction energy. Define

$$U_t^{(\lambda)} = e^{i(t/\lambda)(H_0+H_1+\lambda K)} \ .$$

Then

$$\lim_{\lambda\to 0} U^{(0)}_{t/\lambda}(U^{(\lambda)}_{-t/\lambda} A U^{(\lambda)}_{t/\lambda}) U^{(0)}_{-t/\lambda} = \mathscr{P}_t(A)$$

and $(\mathscr{P}_t)$ is a QMS – with additional properties.

Quantum Markov Semigroups also arise natural in the context of *classical* Markov chains; in particular when one tries to get heat kernel bounds in a discrete setting.

We recall that a family of logarithmic Sobolev inequalities

$$\int_M |\varphi|^2 \log |\varphi|^2 \mathrm{d}\mu - \|\varphi\|_2^2 \log \|\varphi\|_2^2 \leq \epsilon \int_M |\nabla\varphi|^2 \mathrm{d}\mu + (C_0 - \frac{n}{2}\log \epsilon)\|\varphi\|_2^2$$

imply uniform heat kernel bounds

$$K_t(x,y) \leq C_1 e^{C_2 t} t^{-n/2} \ .$$

To get Gaussian bounds, let φ be a Lipschitz function and introduce the modified semigroup with Kernel

Detailed balance

Let $P_{i,j}$ be the Markov transition matrix for a Markov chain on a finite state space $S = \{x_1, \ldots, x_n\}$. Suppose that σ is a probability density on S with

$$\sigma_j = \sum_{i=1}^{n} \sigma_i P_{i,j} \ .$$

The transition matrix satisfies the *detailed balance condition with respect to* σ in case

$$\sigma_i P_{i,j} = \sigma_j P_{j,i} \qquad \text{for all } i, j \ .$$

The matrix P is self-adjoint on $\mathbb{C}^n$ equipped with the inner product

$$\langle f, g\rangle_\sigma = \sum_{k=1}^{n} \sigma_k \overline{f_k} g_k \ ,$$

if and only if the detailed balance condition is satisfied.

There are a number of different ways one might generalize this inner product to the quantum setting, and these give different notions of self-adjointness.

Definition 0.2 (Compatible inner product). *An inner product* $\langle \cdot, \cdot \rangle$ *is* compatible *with* $\sigma \in \mathfrak{S}_+(\mathcal{A})$ *in case for all* $A \in \mathcal{A}$, $\mathrm{Tr}[\sigma A] = \langle \mathbf{1}, A \rangle$.

If a quantum Markov semigroup $\mathscr{P}_t$ is self-adjoint with respect to an inner product $\langle \cdot, \cdot \rangle$ that is compatible with $\sigma \in \mathfrak{S}_+$, then for all $A \in \mathcal{A}$,

$$\mathrm{Tr}[\sigma A] = \langle \mathbf{1}, A \rangle = \langle \mathscr{P}_t \mathbf{1}, A \rangle = \langle \mathbf{1}, \mathscr{P}_t A \rangle = \mathrm{Tr}[\sigma \mathscr{P}_t A] \ ,$$

and thus σ is invariant under $\mathscr{P}_t^\dagger$.

Definition 0.3. *Let* $\sigma \in \mathfrak{S}_+$ *be a non-degenerate density matrix. For each* $s \in \mathbb{R}$, *and each* $A, B \in \mathcal{A}$, *define*

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