Gradient flow and functional inequalities for quantum Markov semigroups, III

Eric A. Carlen

Rutgers University

MPIM Bonn, September 29, 2017

Geodesic convexity

We now develop the advantages of having written the evolution equation $\frac{\partial}{\partial t}\rho = \mathscr{L}^\dagger \rho$ as gradient flow for the relative entropy. We draw on work of Otto and Westdickenberg and also of Daneri and Savaré.

Let $(\mathcal{M}, g)$ be any smooth Riemannian manifold. The Riemannian distance $d_g(x,y)$ between x and y is given by

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

where

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

If F is a smooth function on $\mathcal{M}$, let $\mathrm{grad}_g F$ denote its Riemannian gradient. Consider the semigroup S_t of transformations on $\mathcal{M}$ given by solving $\dot{\gamma}(t) = -\mathrm{grad}_g F(\gamma(t))$; assume that nice global solutions exist. The semigroup S_t, $t \geq 0$, is *gradient flow for* F.

For $\lambda \in \mathbb{R}$, the function F is λ-convex in case whenever $\gamma : [0,1] \to \mathcal{M}$ 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 $x, y \in \mathcal{M}$ 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 contraction as its starting point.

They use the gradient flow transformation S_t to define a one-parameter family of paths $\gamma^t : [0,1] \to \mathcal{M}$, $t \geq 0$ defined by

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

From their work and that of Daneri and Savaré, we have that if for all smooth curves $\gamma : [0,1] \to \mathcal{M}$,

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

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

We now return the QMS setting.

Let $\rho : [0,1] \to \mathfrak{S}_+$ be a smooth path in $\mathfrak{S}_+$, and define the one-parameter family of paths, $\rho^t(s)$, $(s,t) \in [0,1] \times [0,\infty)$ by

$$\rho^t(s) = \mathscr{P}_t^\dagger \rho(s) \ .$$

By what has been explained above, if we can prove that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\left\| \frac{\mathrm{d}}{\mathrm{d}s}\rho^t(s) \right\|^2_{g(\rho^t(s))} \right) \Big|_{0+} \leq -2\lambda \left\| \frac{\mathrm{d}}{\mathrm{d}s}\rho^0(s) \right\|^2_{g(\rho^0(s))}$$

for all smooth $\rho : [0,1] \to \mathcal{M}$ and all $s \in (0,1)$, we will have proved the geodesic convexity of the relative entropy functional, and consequently, we shall have proved

$$D(\mathscr{P}_t^\dagger \rho || \sigma) \leq e^{-2\lambda t} D(\rho || \sigma) \ .$$

A path forward is indicated by a proof of Ledoux of the Gaussian logarithmic Sobolev inequality.
Recall Mehler's formula: (γ_β is the centered gaussian probability density with covariance $\beta\mathbf{1}$.)

$$P_t f(x) = \int_{\mathbb{R}^n} f(e^{-t}x + (1-e^{-2t})^{1/2}y)\gamma_\beta(y)\,\mathrm{d}y\ .$$

Hence $\nabla P_t f(x) = e^{-t}P_t \nabla f(x)$. Since $(x,t) \mapsto |x|^2/t$ is jointly convex,

$$\frac{|\nabla P_t f(x)|^2}{P_t f(x)} = e^{-2t}\frac{|P_t \nabla f(x)|^2}{P_t f(x)} \leq e^{-2t} P_t \frac{|\nabla f(x)|^2}{f(x)}\ .$$

Now integrate. There are two key ingredients: (1) An intertwining relation. (2) A convexity property of the action.

Intertwining for QMS

Lemma 0.1. *Suppose that for some numbers a_j, $j \in \mathcal{J}$,*

$$[\partial_j, \mathscr{L}] = -a_j \partial_j$$

for each $j \in \mathcal{J}$. Then defining $\vec{\mathscr{P}}_t$ on $\bigoplus^{|\mathcal{J}|}\mathcal{A}$ by

$$\vec{\mathscr{P}}_t(A_1, \ldots, A_{|\mathcal{J}|}) = (e^{-ta_1}\mathscr{P}_t A_1, \ldots, e^{-ta_{|\mathcal{J}|}}\mathscr{P}_t A_{|\mathcal{J}|}) ,$$

we have the intertwining relation $\partial_j \mathscr{P}_t = \vec{\mathscr{P}}_t \partial_j$ on $\mathcal{A}$.

Note that

$$\mathscr{P}_t^\dagger \operatorname{div} \mathbf{A}(s) = \operatorname{div} \vec{\mathscr{P}}_t^\dagger \mathbf{A}(s) .$$

Now consider any smooth path $\rho : [0,1] \to \mathfrak{S}_+$, and for each $s \in (0,1)$ write

$$\dot{\rho}(s) = \mathrm{div}\,\mathbf{A}(s)$$

where $\mathbf{A}(s)$ is the solution of $\dot{\rho}(s) = \mathrm{div}\,\mathbf{A}(s)$ that minimizes $\langle \mathbf{A}, [\rho]_{\vec{\omega}}^{-1}\mathbf{A}\rangle_{\mathscr{L},\rho}$ so that

$$g_{\sigma,\rho}(\dot{\rho}(s), \dot{\rho}(s)) = \sum_{j\in\mathcal{J}} \langle A_j(s), [\rho(s)]_{\omega_j}^{-1} A_j(s)\rangle_{\mathfrak{H}_{\mathcal{A}}} \ .$$

Set $\rho^t(s) := \mathscr{P}_t^\dagger \rho(s)$, and suppose that the semigroup $\vec{\mathscr{P}}_t$ given by

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

intertwines with $\mathscr{P}_t$. It follows that

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

Consequently,

$$\left\|\frac{\mathrm{d}}{\mathrm{d}s}\rho^t(s)\right\|^2_{g(\rho^t(s))} \leq e^{-2\lambda t}\sum_{j\in\mathcal{J}} \langle \mathscr{P}_t^\dagger A_j(s), [\mathscr{P}_t^\dagger \rho(s)]^{-1}_{\omega_j} \mathscr{P}_t^\dagger A_j(s)\rangle_{\mathfrak{H}_\mathcal{A}}$$

Monotone metrics

Our work is done if we can show that

$$\sum_{j\in\mathcal{J}} \langle \mathscr{P}_t^\dagger A_j(s), [\mathscr{P}_t^\dagger \rho(s)]_{\omega_j}^{-1} \mathscr{P}_t^\dagger A_j(s)\rangle_{\mathfrak{H}_\mathcal{A}} \leq$$

$$\sum_{j\in\mathcal{J}} \langle A_j(s), [\rho(s)]_{\omega_j}^{-1} A_j(s)\rangle_{\mathfrak{H}_\mathcal{A}} \ .$$

This is where the convexity enters, and the problem is solved by the theory of “monotone metrics” developed by Chentsov and Morozova in the classical case, and by Petz in the non-commutative case.

The map

$$(\rho, A) \mapsto \langle A, [\rho]_\omega^{-1} A\rangle_{\mathfrak{H}_\mathcal{A}} = \\ \mathrm{Tr}\left[\int_0^\infty (t\mathbf{1} + e^{-\omega/2}\rho)^{-1} A^* (t\mathbf{1} + e^{\omega/2}\rho)^{-1} A\,\mathrm{d}t\right]$$

is jointly convex on $\mathfrak{S}_+ \times \mathcal{A}$. If ρ and A are scalars, the right-hand side reduces to A^2/ρ. The non-commutative convexity result ultimately derives from Lieb's concavity Theorem. Since $\mathscr{P}_t^\dagger$ is completely positive,

$$\langle \mathscr{P}_t^\dagger A, [\mathscr{P}_t^\dagger \rho]_\omega^{-1} \mathscr{P}_t^\dagger A\rangle_{\mathfrak{H}_\mathcal{A}} \leq \langle A, [\rho]_\omega^{-1} A\rangle_{\mathfrak{H}_\mathcal{A}} \ .$$

Summarizing, whenever we can show that the semigroup $\vec{\mathscr{P}}_t$ given by

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

intertwines with $\mathscr{P}_t$, we have λ-convexity of the relative entropy, and we have the entropy dissipation inequality

$$D(\mathscr{P}_t^\dagger \rho || \sigma) \leq e^{-2\lambda t} D(\rho || \sigma) \ .$$

The problem now is to verify the intertwining property. In Ledoux's proof, this was done using an explicitly formula for the action of the semigroup $(\mathscr{P}_t)_{t \geq 0}$. In our case it will be easier to work with the generator $\mathscr{L}$.

Lemma 0.2. *Suppose that for some numbers a_j, $j \in \mathcal{J}$,*

$$[\partial_j, \mathscr{L}] = -a_j \partial_j$$

for each $j \in \mathcal{J}$. Then defining $\vec{\mathscr{P}}_t$ on $\mathfrak{H}_{\mathcal{A},\mathcal{J}}$ by

$$\vec{\mathscr{P}}_t(A_1, \ldots, A_{|\mathcal{J}|}) = (e^{-ta_1}\mathscr{P}_t A_1, \ldots, e^{-ta_{|\mathcal{J}|}}\mathscr{P}_t A_{|\mathcal{J}|}) \ ,$$

we have the intertwining relation $\partial_j \mathscr{P}_t = \vec{\mathscr{P}}_t \partial_j$ on $\mathcal{A}$.

Onwards to examples!

The Fermi O-U Semigroup

Let $\mathcal{A}$ be the Clifford algebra $\mathfrak{C}^n$ of dimension $n = 2m$ for some $m \in \mathbb{N}$. Consider a set of generators

$$\{Q_1, \ldots, Q_m, P_1, \ldots, P_m\} \ ,$$

where

$$Q_j Q_k + Q_k Q_j = P_j P_k + P_k P_j = 2\delta_{j,k}\mathbf{1}$$

and

$$Q_j P_k + P_k Q_j = 0 \quad \text{for all } 1 \le j,k \le m \ .$$

Think of $\mathcal{A}$ as the full set of *phase space observables*, the subalgebra generated by $\{Q_1, \ldots, Q_m\}$ as the algebra of *configuration space observables*, and the subalgebra generated by $\{P_1, \ldots, P_m\}$ as the algebra of *momentum space observables*.

Form the operators

$$Z_j = \frac{1}{\sqrt{2}}(Q_j + iP_j) \quad \text{so that} \quad Z_j^* = \frac{1}{\sqrt{2}}(Q_j - iP_j)\ .$$

It is easy to check that

$$Z_j Z_k + Z_k Z_j = 0 \quad \textsf{and} \quad Z_j Z_k^* + Z_k^* Z_j = 2\delta_{j,k}\mathbf{1}$$

for all j, k. The formulas

$$N_j = \frac{1}{2} Z_j^* Z_j \quad \text{and} \quad N_j^\perp = \frac{1}{2} Z_j Z_j^* \quad \text{for all } \ 1 \le j \le m \ \ .$$

define m pairs of complementary orthogonal projections.

For any set of m real numbers $\{e_1, \ldots, e_m\}$, and any parameter $\beta \in (0, \infty)$, to be interpreted as the *inverse temperature*, define the *free Hamiltonian* h and the *Gibbs state* σ_β by

$$h = \sum_{j=1}^{m} e_j N_j \quad \text{and} \quad \sigma_\beta = \frac{1}{\tau[e^{-\beta h}]} e^{-\beta h} \ .$$

where τ is the normalized trace.

Let $W := i^m \prod_{j=1}^{m} Q_j P_j$ so that W is self-adjoint and unitary, and for all $A \in \mathcal{A}$, let $\Gamma(A) = WAW$.

Since W commutes with every even element of $\mathcal{A}$, simple computations show that

$$\Delta_{\sigma_\beta}(WZ_j) = e^{\beta e_j} WZ_j \quad \text{and} \quad \Delta_{\sigma_\beta}(Z_j^* W) = e^{-\beta e_j} Z_j^* W \ .$$

Define the operators

$$V_j := WZ_j, \qquad 1 \le j \le m \ ,$$

so that $\frac{1}{2} V_j^* V_j = N_j$ and $\frac{1}{2} V_j V_j^* = N_j^\perp$.

The operator $\mathscr{L}_\beta$

$$\mathscr{L}_\beta A := \frac{1}{4}\sum_{j=1}^m e^{\beta e_j/2}\Big(V_j^*[A,V_j] + [V_j^*,A]V_j\Big) +$$

$$\frac{1}{4}\sum_{j=1}^m e^{-\beta e_j/2}\Big(V_j[A,V_j^*] + [V_j,A]V_j^*\Big)$$

is the generator of a QMS $\mathscr{P}_t = e^{t\mathscr{L}_\beta}$ that satisfies the σ_β-DBC. This is the *Fermi Ornstein-Uhlenbeck semigroup at inverse temperature* β.

It is a simple matter to diagonalize $\mathscr{L}_\beta$: For each $1 \leq j \leq m$, define the four operators

$$K_{j,(0,0)} = \mathbf{1}\ , \quad K_{j,(1,0)} = Z_j$$
$$K_{j,(0,1)} = Z_j^* \quad \text{and} \quad K_{j,(1,1)} = e^{\beta e_j/2} N_j - e^{-\beta e_j/2} N_j^\perp\ .$$

One readily checks that this set of four operators is orthonormal in any of the inner products $\langle \cdot, \cdot \rangle_s$ based on σ_β.

Using the fact that for each j, V_j and V_j^* commute with P_k and Q_k for all $k \neq j$, and using the identities $V_j K_{j,(1,1)} = e^{\beta e_j/2} V_j$ and $K_{j,(1,1)} V_j = -e^{-\beta e_j/2} V_j$, we readily compute that

$$\mathscr{L}_\beta Z_j = -\cosh(\beta e_j/2) Z_j$$

$$\mathscr{L}_\beta K_{j,(1,1)} = -2\cosh(\beta e_j/2) K_{j,(1,1)} \ .$$

Therefore, for all $0 \leq k, \ell \leq 1$,

$$\mathscr{L}_\beta K_{j,(k,\ell)} = -(k+\ell)\cosh(\beta e_j/2) K_{j,(k,\ell)} \ .$$

Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)$ denote a generic element of the index set $\{\{0,1\} \times \{0,1\}\}^m$, and for $\alpha = (k, \ell) \in \{0,1\} \times \{0,1\}$, define $|\alpha| = k + \ell$. Then the functions

$$K_{\boldsymbol{\alpha}} := K_{1,\alpha_1} K_{2,\alpha_2} \cdots K_{m,\alpha_m}$$

are an orthogonal (but not normalized) basis for $\mathfrak{H}_{\mathcal{A}}$ consisting of eigenvectors of $\mathscr{L}_\beta$:

$$\mathscr{L}_\beta K_{\boldsymbol{\alpha}} = -\left(\sum_{j=1}^{m} |\alpha_j| \cosh(\beta e_j/2) \right) K_{\boldsymbol{\alpha}} \ .$$

Intertwining - with a twist

There is another differential calculus, more closely adapted to $\mathscr{L}_\beta$: For $1 \leq j \leq m$, define

$$\check{\partial}_j A = \frac{1}{2}(Z_j A - \Gamma(A) Z_j) = \frac{1}{2} W[V_j, A]$$

$$\overline{\check{\partial}}_j A = \frac{1}{2}(Z_j^* A - \Gamma(A) Z_j^*) = -\frac{1}{2} W[V_j^*, A]$$

$$\check{\partial}_j K_{j,(0,0)} = \check{\partial}_j K_{j,(1,0)} = 0 \ , \quad \check{\partial}_j K_{j,(0,1)} = K_{j,(0,0)}$$
$$\text{and} \quad \check{\partial}_j K_{j,1,1} = \cosh(\beta e_j/2) K_{j,(1,0)} \ ,$$

$$\overline{\check{\partial}}_j K_{j,(0,0)} = \overline{\check{\partial}}_j K_{j,(0,1)} = 0 \ , \quad \overline{\check{\partial}}_j K_{j,(1,0)} = K_{j,(0,0)}$$
$$\text{and} \quad \overline{\check{\partial}}_j K_{j,(1,1)} = -\cosh(\beta e_j/2) K_{j,(0,1)} \ .$$

Using the fact that for each j, V_j and V_j^* commute with P_k and Q_k for all $k \neq j$, one determines the effect of $\check{\partial}_j$ and $\overline{\check{\partial}}_j$ on all of $\mathcal{A}$. The orthonormal basis $\{K_{\boldsymbol{\alpha}}\}$ may be viewed as consisting of analogs of multivariate Krawtchouck polynomials – the discrete analogs of the Hermite polynomials.

$\check{\partial}_j$ and $\overline{\check{\partial}}_j$, which are skew derivations, have the advantage that they always lower the "degree" of any $K_{\boldsymbol{\alpha}}$ by one, as one would expect. The operators $\partial_j A$ and $\overline{\partial}_j A$ do not do this.

Using this, one readily deduces the identities,

$$\check{\partial}_j \mathscr{L}_\beta K_{\boldsymbol{\alpha}} - \mathscr{L}_\beta \check{\partial}_j K_{\boldsymbol{\alpha}} = -\cosh(\beta e_j/2)\check{\partial}_j K_{\boldsymbol{\alpha}} \tag{1}$$

and

$$\overline{\check{\partial}}_j \mathscr{L}_\beta K_{\boldsymbol{\alpha}} - \mathscr{L}_\beta \overline{\check{\partial}}_j K_{\boldsymbol{\alpha}} = -\cosh(\beta e_j/2)\overline{\check{\partial}}_j K_{\boldsymbol{\alpha}} \ . \tag{2}$$

Theorem 0.3. *For $\beta \geq 0$, 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 $\lambda_\beta = \min\{\cosh(\beta e_j/2) \ : \ j = 1, \dots, m\}$. Moreover, the relative entropy functional $\rho \mapsto D(\rho || \sigma_\beta)$ is geodesically λ_β convex in the Riemannain metric associated to $\mathscr{L}_\beta$.

The embedded walk on the cube

Each of the vectors $K_{\boldsymbol{\alpha}}$ is an eigenvector of Δ_{σ_β}. Moreover, if $\{e_1, \ldots, e_m\}$ is linearly independent over the integers, then $\Delta_{\sigma_\beta} K_{\boldsymbol{\alpha}} = K_{\boldsymbol{\alpha}}$ if and only if for each k, $|\alpha_k| \neq 1$. The span of the set of such $K_{\boldsymbol{\alpha}}$ is the same as the span of

$$\{N_1, N_1^\perp, \ldots, N_m, N_m^\perp\} \ .$$

Hence in this case, the *modular algebra* $\mathcal{A}_{\sigma_\beta}$ is the algebra generated by the commuting projections listed above. Denote this algebra by $\mathcal{B}$. While it need not be the modular algebra when $\{e_1, \ldots, e_m\}$ is not linearly independent over the integers, it is easy to see (by continuity or computation) that it is always invariant under $\mathscr{P}_t$.

The projections $\{N_1, N_1^\perp, \ldots, N_m, N_m^\perp\}$ are not minimal in $\mathcal{B}$, but the set of the 2^m distinct non-zero products one can form from them is a full set of minimal projections. Identify this set with the discrete hypercube $\mathscr{Q}^m = \{0,1\}^m$: Set $\mathcal{J} = \{1, \ldots, m\}$, and let $s_j : \mathscr{Q}^m \to \mathscr{Q}^m$ define the j-th coordinate swap defined by $s_j(x_1, \ldots, x_m) = (x_1, \ldots, -x_j, \ldots, x_m)$. Let $\mathbf{x}$ denote a generic point of $\mathscr{Q}^m$. Define $E_{\mathbf{x}} = \prod_{j=1}^m N_j^{x_1}(N_j^\perp)^{1-x_1}$. The restriction $\widetilde{\mathscr{P}_t}$ of $\mathscr{P}_t$ to $\mathcal{B}$ is a nearest neighbor random walk on $\mathscr{Q}^m$.

For a standard representation in which the elements of $\mathcal{A}$ operate on $\mathbb{C}^{2^m}$, and τ is the normalized trace, each $E_{\mathbf{x}}$ is rank one, so that the transition rate matrix D for the walk is simply $D_{\mathbf{x},\mathbf{x}'} = \mathrm{Tr}[E_{\mathbf{x}}, \mathscr{L} E_{\mathbf{x}'}]$. One readily computes that $D_{\mathbf{x},\mathbf{x}'} = 0$ unless $\mathbf{x}' = s_j(\mathbf{x})$ for some j, and in that case

$$D_{\mathbf{x},\mathbf{x}'} = \begin{cases} \dfrac{2\cosh(\beta e_j)}{1+e^{-\beta e_j}} & x_j = 1 \\ \dfrac{2\cosh(\beta e_j)}{1+e^{\beta e_j}} & x_j = 0 \ , \end{cases}$$

and this gives the jump rates along the edges of $\mathscr{Q}^m$ for the classical Markov chain corresponding to $\widetilde{\mathscr{P}_t}$.

The transition matrix D is the one, for this model, arising in the general equation for the evolution of occupation probability for a reversible Markov chain:

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_\ell(t) = \sum_{k=1}^{m} \left(\rho_k(t)D_{k,\ell} - \rho_\ell(t)D_{\ell,k}\right) \ .$$

$$\sigma_k D_{k,\ell} = \sigma_\ell D_{\ell,k}$$

The Bose O.U. Semigroup

Let Z and Z^* be Bose annihilation and creation operators: $[Z, Z^*] = \mathbf{1}$. Define

$$\sigma_\beta = \left(\mathrm{Tr}\left[e^{-\beta h}\right]\right)^{-1} e^{-\beta h} \ .$$

Theorem 0.4. *Let $\mathscr{P}_t$ be the Bose Ornstein-Uhlenbeck semigroup with generator $\mathscr{L}_\beta$ given by*

$$\mathscr{L}_\beta = e^{\beta/2}\left(Z^*AZ - \tfrac{1}{2}\{ZZ^*, A\}\right) + e^{-\beta/2}\left(ZAZ^* - \tfrac{1}{2}\{Z^*Z, A\}\right) \ ,$$

and let σ_β be its invariant state. Then for all $\rho \in \mathfrak{S}_+$,

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