Toy hypocoercivity

Sébastien Gadat and Laurent Miclo

Institut de Mathématiques de Toulouse

Plan of the talk

1. Introduction and results
2. Spectral decomposition of the first model
3. Computations of the operator norms
4. Gaussian case
5. Hopes...

Warming up

Consider a Markov generator L admitting a reversible probability μ. Let $(P_t)_{t\geqslant 0}$ be the associated semi-group and denote by $\|\!|\cdot|\!\|$ the operator norm in $\mathbb{L}^2(\mu)$. For any time $t \geqslant 0$, we have

$$\|\!|P_t - \mu|\!\| \quad = \quad \exp(-\lambda t) \tag{1}$$

where $\lambda \geqslant 0$ is the spectral gap of $-L$.

Goal of this talk: to investigate what may happen when μ is only assumed to be an invariant probability of L.

Hypocoercivity

Traditionally, corresponds to a hypoelliptic diffusion generator L, with nevertheless an exponentially fast convergence to equilibrium: at least for nice functions $f \in \mathbb{L}^2(\mu)$, with $\mu[f] = 0$,

$$\forall\ t \geqslant 0, \qquad \|P_t f\| \leqslant C(f) \exp(-ct)$$

where $c > 0$ is independent from f.

Kinetic evolution equations

Simple example:

$$L = y\partial_x - U'(x)\partial_y + \partial_y^2 - y\partial_y$$

where $x \in \mathbb{T}$ (position) and $y \in \mathbb{R}$ (speed) and where $U : \mathbb{R} \to \mathbb{R}$ is a regular potential.
Probabilistic description:

$$\begin{cases} dX_t = Y_t\, dt \\ dY_t = -U'(X_t)\, dt + \sqrt{2}\, dB_t - Y_t\, dt \end{cases}$$

where $(B_t)_{t \geqslant 0}$ is a Brownian motion, not enabling to explore "at once" the whole space space $\mathbb{T} \times \mathbb{R}$.

Invariant measure

$$\mu(dx, dy) \quad = \quad \frac{\exp(-U(x))}{Z}\, dx \otimes \gamma(dy)$$

where Z is the normalizing constant and γ is the standard Gaussian distribution.

Not reversible: $\gamma(dy)$ is reversible for the Ornstein-Uhlenbeck generator $\partial_y^2 - y\partial_y$, but the vector field $y\partial_x - U'(x)\partial_y$ is turning around the level sets of the energy $U(x) + y^2/2$.

References

Two approaches:

Analytical (hypoellipticity, pseudo-differential operators ...):
Eckmann and Hairer [00], Desvillettes and Villani [01], Rey-Bellet and Thomas [02], Hérau and Nier [04], Hérau [07], Villani [09], Dolbeault, Mouhot and Schmeiser [10], Ottobre, Pavliotis and Pravda-Starov [11].

Probabilistic (Liapounov functions):
Bakry, Cattiaux and Guillin [08].

Motivation

These works give no clue about how $\|P_t - \mu\|$ is decreasing for small times and even the asymptotical rate c is never optimal.

For global optimization, we would like to resort to time-inhomogeneous stochastic algorithms which are instantaneously hypocoercive. This leads to the question of seeing how instantaneously these processes starts their convergence. The usual changes of norms (from $\mathbb{L}^2$ to appropriately weighted $\mathbb{H}^1$) are not convenient in this direction.

Hope: toy models will help the understanding.

First toy model

Corresponds to $U = 0$: $(X_t)_{t \geqslant 0}$ is the integral of a Ornstein-Uhlenbeck rounded around the circle. Let us add a parameter $a > 0$:

$$L_a \quad = \quad y\partial_x + a\partial_y^2 - y\partial_y$$

It is possible to get a formula for the associated semi-group $P_t^{(a)}$, $t \geqslant 0$, but it does not help to compute $\| P_t^{(a)} - \mu_a \|$, where the invariant measure is $\mu_a = \lambda \otimes \gamma_a$, with λ the normalized Lebesgue measure on $\mathbb{T}$ and γ_a the centered Gaussian distribution of variance a.

First result

Theorem

For any $a > 0$ and $t \geqslant 0$, we have

$$\|P_t^{(a)} - \mu_a\| = \max\left(\exp(-t), \exp\left[-a\left(t - 2\frac{1-\exp(-t)}{1+\exp(-t)}\right)\right]\right)$$

The fact that the rhs has not the form of an exponential function reflects that the functions where the norm is attained depend on time.

For small times $t > 0$,

$$\ln\left(\|P_t^{(a)} - \mu_a\|\right) \sim -\frac{a}{12}t^3$$

Initially, the convergence is quite slow with respect to that of reversible evolutions (1). The power 3 should be interpreted as a degree of hypocoercivity, it indicates how far is the evolution from an "immediate exploration".

Hypocoercive behavior

As t goes to $+\infty$,

$$-\ln\left(\left\|P_t^{(a)} - \mu_a\right\|\right) = \begin{cases} a(t-2+\mathcal{O}(e^{-t})) & \text{, if } a \leqslant 1 \\ t & \text{, if } a > 1 \end{cases}$$

This kind of hypocoercive bound seems to be new, the asymptotical rate was not obtained in the literature.

Scale changes

Up to scalings in time and in the speed variable and to a change of direction in position:

Corollary

For any $a, c > 0$ and $b \in \mathbb{R}\backslash\{0\}$, consider

$$L_{a,b,c} \quad := \quad by\partial_x + a\partial_y^2 - cy\partial_y$$

which admits $\mu_{a/c}$ as invariant probability. For the corresponding semi-group $(P_t^{(a,b,c)})_{t\geqslant 0}$,

$$\forall\ t \geqslant 0, \qquad \| P_t^{(a,b,c)} - \mu_{a/c} \|_{\mathbb{L}^2(\mu_{a/c})\circlearrowleft}$$
$$= \quad \max\left(\exp(-ct), \exp\left[-\frac{ab^2}{c^3}\left(ct - 2\frac{1-\exp(-ct)}{1+\exp(-ct)}\right)\right]\right)$$

Comparison with reversible MCMC

It is instructive to compare with the heat semi-group generated by the operator $K_a := a\partial_x^2$, injecting the same amount a of randomness per unit of time as the generators $L_{a,b,c}$. The probability λ is reversible for $-K_a$ whose spectral gap is a, so

$$\forall\ t \geqslant 0, \qquad \| \exp(tK_a) - \lambda \|_{\mathbb{L}^2(\lambda)\circlearrowleft} \quad = \quad \exp(-at).$$

Thus to sample MCMCly according to λ, it is asymptotically advantageous to tune $c > a$ and $b > c$ and to use the first coordinate generated by $L_{a,b,c}$. This is another (dubious) illustration of the paradigm that to go fast to equilibrium, it is better to resort to non-reversible Markov processes.

Second toy model

For position space, replace $\mathbb{T}$ by $\mathbb{R}$ and consider the confining potential $U(x) = ax^2/2$, where $a > 0$:

$$\widetilde{L}_a \quad := \quad y\partial_x - ax\partial_y + \partial_y^2 - y\partial_y.$$

The invariant probability measure is $\widetilde{\mu}_a := \gamma_{1/a} \otimes \gamma_1$, denote $(\widetilde{P}_t^{(a)})_{t \geqslant 0}$ the associated semi-group on $\mathbb{L}^2(\widetilde{\mu}_a)$.
The value $1/4$ is critical for the diagonalization of $\widetilde{L}_a$: for $a \in (0, 1/4)$, $\widetilde{L}_a$ is diagonalizable in $\mathbb{L}^2(\widetilde{\mu}_a)$ and its spectrum is real, while for $a \in (1/4, +\infty)$, $\widetilde{L}_a$ is still diagonalizable in $\mathbb{L}^2(\widetilde{\mu}_a)$ (complexified) but most of its eigenvalues are not real. In the critical case $a = 1/4$, $\widetilde{L}_a$ is not diagonalizable in $\mathbb{L}^2(\widetilde{\mu}_a)$ and contains Jordan blocks of all orders.

Second result (1)

Theorem

For any $a > 0$ and $t \geqslant 0$,

$$\|\widetilde{P}_t^{(a)} - \widetilde{\mu}_a\| \quad = \quad C_a(t) \exp\left(-\frac{1-\sqrt{(1-4a)_+}}{2} t\right)$$

where:

- *if $a \in (0, 1/4)$, let $\theta := \sqrt{1-4a}$ and define*

$$C_a(t) \quad := \quad \frac{\sqrt{e^{-\theta t} + \frac{1-\theta^2}{2\theta^2}\left(1-e^{-\theta t}\right)^2 + \frac{1-e^{-2\theta t}}{2}\left(1 + \frac{1}{\theta}\sqrt{1 + (\theta^{-2}-1)\left(\frac{e^{\theta t}-1}{e^{\theta t}+1}\right)^2}\right)}}{}$$

Second result (2)

Theorem (continued)

- *If* $a = 1/4$, *define*

$$C_a(t) \quad := \quad \sqrt{1 + \frac{t^2}{2} + t\sqrt{1 + \left(\frac{t}{2}\right)^2}}$$

- *If* $a \in (1/4, +\infty)$, *let* $\theta := \sqrt{4a-1}\,i$ *and define*

$$C_a(t) \quad := \quad \sqrt{1 + \frac{|e^{\theta t} - 1|}{2|\theta|^2}\left(|e^{\theta t} - 1| + \sqrt{|e^{\theta t} - 1|^2 + 4|\theta|^2}\right)}$$

Again for small times $t > 0$:

$$\ln\left(\|\widetilde{P}_t^{(a)} - \widetilde{\mu}_a\|\right) \quad \sim \quad -\left(\frac{a}{6} + \frac{(1-4a)_+}{2}\left(1 - \sqrt{(1-4a)_+}\right)\right) t^3$$

For large times

As t goes to infinity, different behaviors occur:

• if $a \in (0, 1/4)$,

$$\|\widetilde{P}_t^{(a)} - \widetilde{\mu}_a\| \quad \sim \quad \frac{1}{\theta} \exp\left(-\frac{1-\sqrt{1-4a}}{2} t\right)$$

• if $a = 1/4$, the pre-exponential factor explodes linearly:

$$\|\widetilde{P}_t^{(1/4)} - \widetilde{\mu}_{1/4}\| \quad \sim \quad t \exp\left(-\frac{t}{2}\right)$$

• if $a > 1/4$, the factor $C_a(t)$ is oscillating between the values 1 and $\sqrt{1 + 2(1 + 2\sqrt{a})(4a-1)^{-1}}$, with period $T_a := 2\pi/\sqrt{4a-1}$. These oscillations are sufficiently moderate so that $\mathbb{R}_+ \ni t \mapsto C_a(t) \exp(-t/2)$ is non-increasing, as it should be.

Periodic hesitations toward equilibrium

For $a > 1/4$: from the small time behavior,

$$\frac{d}{dt}\|\widetilde{P}_t^{(a)} - \widetilde{\mu}_a\|\Big|_{t=0} = 0$$

and from periodicity, $\forall\ k \in \mathbb{Z}_+$, $\forall\ t \geqslant 0$,

$$C_a(kT_a + t)\exp(-(kT_a + t)) = \exp(-kT_a)C_a(t)\exp(-t)$$

As a consequence, $\forall\ k \in \mathbb{Z}_+$,

$$\frac{d}{dt}\|\widetilde{P}_t^{(a)} - \widetilde{\mu}_a\|\Big|_{t=kT_a} = 0$$

Scalings

For any $c, d > 0$ and $a, b \in \mathbb{R}$ with $ab > 0$, let

$$\widetilde{L}_{a,b,c,d} \quad := \quad by\partial_x - ax\partial_y + c\partial_y^2 - dy\partial_y$$

whose invariant probability is $\widetilde{\mu}_{a,b,c,d} := \gamma_{bc/(ad)} \otimes \gamma_{c/d}$. For the corresponding semi-group $(\widetilde{P}_t^{(a,b,c,d)})_{t \geqslant 0}$, $\forall\ t \geqslant 0$,

$$\|\widetilde{P}_t^{(a,b,c,d)} - \widetilde{\mu}_{a,b,c,d}\| \quad = \quad C_{ab/d^2}(dt) \exp\left(-\frac{1 - \sqrt{(1 - 4abd^{-2})_+}}{2} dt\right)$$

Similar remarks as above for the comparison with the reversible Ornstein-Uhlenbeck $c\partial_x^2 - \frac{da}{b} x\partial_x$ are valid.

Natural functions

First task: to find orthogonal subspaces generating $\mathbb{L}^2(\mu_a)$ and left stable by L_a. It is tempting to look at the action of L_a on functions of the form $\varphi_p \otimes h_{q,a}$ and $\psi_p \otimes h_{q,a}$, where for any $p \in \mathbb{Z}_+$ and any $x \in \mathbb{T}$,

$$\begin{aligned} \varphi_p(x) &:= \frac{2^p p!}{\sqrt{(2p)!}} \cos(px) \\ \psi_p(x) &:= \frac{2^p p!}{\sqrt{(2p)!}} \sin(px) \end{aligned}$$

and for any $q \in \mathbb{Z}_+$ and any $y \in \mathbb{R}$,

$$\begin{aligned} h_q(y) &:= \frac{(-1)^q}{\sqrt{q!}} \exp(y^2/2) \frac{d^q}{dy^q} \exp(-y^2/2) \\ h_{q,a}(y) &:= h_q(y/\sqrt{a}) \end{aligned}$$

The family $(\varphi_p \otimes h_{q,a}, \psi_{p+1} \otimes h_{q,a})_{p,q \in \mathbb{Z}_+}$ is an orthonormal basis of $\mathbb{L}^2(\mu_a)$.

Orthogonal stable subspaces

It is obvious that for $q \in \mathbb{Z}_+$, $\varphi_0 \otimes h_{q,a}$ is an eigenfunction of L_a associated to the eigenvalue $-q$, denote $\mathcal{U}_q := \mathrm{Vect}(\varphi_0 \otimes h_{q,a})$. On the other hand, one computes that for $p \in \mathbb{N}$, the following vector subspaces

$$\begin{aligned} \mathcal{V}_p &:= \overline{\mathrm{Vect}(\varphi_p \otimes h_{q,a}, \psi_p \otimes h_{q+1,a} \,:\, q \in 2\mathbb{Z}_+)} \\ \mathcal{W}_p &:= \overline{\mathrm{Vect}(\psi_p \otimes h_{q,a}, \varphi_p \otimes h_{q+1,a} \,:\, q \in 2\mathbb{Z}_+)} \end{aligned}$$

are stable by L_a. So

$$\mathbb{L}^2(\mu_a) = \bigoplus_{q \in \mathbb{Z}_+} \mathcal{U}_q \bigoplus_{p \in \mathbb{N}} \mathcal{V}_p \bigoplus_{p' \in \mathbb{N}} \mathcal{W}_{p'}$$

where the components are orthogonal and stable.
It is sufficient to study the restrictions of L_a to the $\mathcal{V}_p$, $p \in \mathbb{N}$, since they are isometrically conjugate to the restrictions of L_a to the $\mathcal{W}_p$, $p \in \mathbb{N}$.

Back to $l^2(\mathbb{Z}_+)$

Consider the orthonormal basis $e_0 := \varphi_p \otimes h_{0,a}$, $e_1 := \psi_p \otimes h_{1,a}$, $e_2 := \varphi_p \otimes h_{2,a}$ etc. Identifying $\mathcal{V}_p$ with $l^2(\mathbb{Z}_+)$, the restriction of L_a to $\mathcal{V}_p$ is given by the infinite tridiagonal matrix M

$$M := \begin{pmatrix} 0 & \sqrt{a}p & 0 & 0 & \cdots \\ -\sqrt{a}p & -1 & -\sqrt{2}\sqrt{a}p & 0 & \cdots \\ 0 & \sqrt{2}\sqrt{a}p & -2 & \sqrt{3}\sqrt{a}p & \ddots \\ 0 & 0 & -\sqrt{3}\sqrt{a}p & -3 & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

This object is only parametrized by $c := \sqrt{a}p$, $M = D + cS - cS^*$, where D, cS and $-cS^*$ are the diagonal, the upper-diagonal and the lower-diagonal of M. Denote $\mathcal{S}$, the subspace of $(z(q))_{q\in\mathbb{Z}_+} \in l^2(\mathbb{Z}_+)$ such that for any $r \geqslant 0$, $\sum_{q\in\mathbb{Z}_+} q^r z^2(q) < +\infty$.

Spectral decomposition of M

Theorem

Define $\xi_0 := (\xi_0(q))_{q\in\mathbb{Z}_+} \in \mathcal{S}$ *by*

$$\xi_0(q) \quad := \quad (-1)^{\lfloor \frac{q+1}{2} \rfloor} \frac{c^q}{\sqrt{q!}} \exp(-c^2/2)$$

for any $q \in \mathbb{Z}_+$. *Consider the elements of* $\mathcal{S}$ *given by*

$$\forall\ n \in \mathbb{Z}_+, \qquad \xi_n \quad = \quad (cI - S^*)^n \xi_0,$$

(I *is the identity operator). Then for any* $n \in \mathbb{Z}_+$, ξ_n *is an eigenvector of* M *associated to the eigenvalue* $-c^2 - n$. *Furthermore* $(\xi_n)_{n\in\mathbb{Z}_+}$ *is a (Hilbert) basis of* $l^2(\mathbb{Z}_+)$.

The solving Lie algebra

Consider V the vector space generated by D, S, S^* and I, it is a 4-dimensional Lie algebra, because one computes that

$$\begin{aligned} [S, S^*] &= I, \\ [D, S] &= S, \\ [D, S^*] &= -S^* \end{aligned}$$

One is then led to consider the adjoint operator of M

$$\mathrm{ad}_M : V \ni X \mapsto [M, X] \in V$$

which is easy to diagonalize:

Lemma

The kernel of the operator ad_M is generated by I and M. There are two other eigenvalues, 1 and -1, whose corresponding eigenspaces are respectively generated by $J_+ := cI + S$ and $J_- := cI - S^*$.

Sketch of the decomposition of M

As a consequence, if z is a eigenvector of M associated to l, either $J_\pm(z)$ is an eigenvector of M associated to $l \pm 1$, either $J_\pm(z) = 0$. Indeed, for instance for $\pm = +$,

$$\begin{aligned} MJ_+(z) &= J_+M(z) + J_+(z) \\ &= (l+1)J_+(z) \end{aligned}$$

Since M comes from a Markovian operator, its eigenvalues have a non-positive real part and the above procedure cannot be repeated ad libidum. This suggests that $\mathrm{Ker}(J_+) \neq \{0\}$ and one computes that $\mathrm{Ker}(J_+) = \mathrm{Vect}(\xi_0)$ and that ξ_0 is an eigenvector of M associated to $-c^2$. Starting from ξ_0, the other eigenvectors are obtained by applying J_-, whose kernel is reduced to $\{0\}$.

An isometry

To get that $(\xi_n)_{n\in\mathbb{N}}$ is an Hilbert basis is more technical, since these vectors are not orthogonal.

To a vector $z = \sum_{n\in\mathbb{Z}_+} f(n)\xi_n$ associate the holomorphic function $F(X) := \sum_{n\in\mathbb{Z}_+} f(n)X^n$ and consider

$$\forall\ n \in \mathbb{Z}_+, \qquad G(n) \ := \ \left(1 + \frac{1}{c}\frac{d}{dX}\right)^n F(X)\Big|_{X=c}$$

The mapping $\mathcal{Q} \ni z \mapsto G$ is an isometry between $l^2(\mathbb{Z}_+)$ and $\mathbb{L}^2(\mathcal{P}(c^2))$, where $\mathcal{P}(c^2)$ stands for the Poisson distribution of parameter c^2. The basic ingredient is the following formula.

A crucial technical result

Lemma

For any $n \in \mathbb{Z}_+$, we have

$$(S^*)^n(\xi_0) = \frac{1}{c^n} D(D+1)(D+2)\cdots(D+n-1)(\xi_0) \quad (2)$$

In particular for $n = 1$, it gives $M(\xi_0) = -c^2\xi_0$ (recalling that $J_+(\xi_0) = 0$, so $S(\xi_0) = -c\xi_0$).

The above formula enables to compute the coefficients $\xi_n(q)$, $q \in \mathbb{Z}_+$ and next to come back to the functions $(x, y) \mapsto \xi_n(x, y)$. They are linear combinations of terms $y^m \cos(p(x+y))$ and $y^m \sin(p(x+y))$, for $m \in [\![0, n]\!]$.

Scalar products

More important for our purposes: the scalar products of the eigenvectors.

Lemma

We have for all $n, m \in \mathbb{Z}_+$,

$$\langle \xi_n, \xi_m \rangle = (2c)^{n+m} \exp((4c^2)^{-1}) \mathbb{E}[n^{(N_{1/(4c^2)})} m^{(N_{1/(4c^2)})}]$$

where $N_{1/(4c^2)}$ is a Poisson random variable of parameter $1/(4c^2)$ and where $n^{(N)} = n(n-1)\cdots(n-N+1)$.

It follows that $\mathcal{V}_p$ (and similarly for $\mathcal{W}_p$) cannot be decomposed into orthogonal and stable non-trivial subspaces.

Convenient vectors

As a consequence, the scalar product of certain vectors is easy to compute: for $\widetilde{\rho}$, $\widehat{\rho} \in \mathbb{R}$, consider

$$\widetilde{z} := \sum_{n \in \mathbb{Z}_+} \frac{\widetilde{\rho}^n}{n!} \xi_n \qquad \text{and} \qquad \widehat{z} := \sum_{n \in \mathbb{Z}_+} \frac{\widehat{\rho}^n}{n!} \xi_n$$

Then we have

$$\langle \widetilde{z}, \widehat{z} \rangle \quad = \quad \exp(\widetilde{\rho}\widehat{\rho} + 2c(\widetilde{\rho} + \widehat{\rho}))$$

In particular, for $z = \sum_{n \in \mathbb{Z}_+} \frac{\rho^n}{n!} \xi_n$, with $\rho \in \mathbb{R}$, we get

$$\|z\|^2 \quad = \quad \exp\left(\rho^2 + 4c\rho\right).$$

Lower bound on operator norms

The semi-group $(\exp(tM))_{t\geqslant 0}$ acts simply on this kind of vectors:

$$\forall\ t \geqslant 0, \qquad \exp(tM)z \quad = \quad \exp(-c^2 t) \sum_{n\in\mathbb{Z}_+} \exp(-nt)\frac{\rho^n}{n!}\xi_n,$$

so that for any $t \geqslant 0$,

$$\|\exp(tM)z\|^2 \quad = \quad \exp(-2c^2 t)\exp\left(\exp(-2t)\rho^2 + 4\exp(-t)c\rho\right).$$

We deduce a lower bound on the operator norm $|\!|\!|\exp(tM)|\!|\!|$ in $l^2(\mathbb{Z}_+)$, by optimizing over $\rho \in \mathbb{R}$: for any $t \geqslant 0$, we have

$$|\!|\!|\exp(tM)|\!|\!| \quad \geqslant \quad \exp\left(-c^2\left(t - 2\frac{1-\exp(-t)}{1+\exp(-t)}\right)\right)$$

Upper bound (1)

To get the corresponding upper bound, consider $\mathcal{Z}$ the space of vectors of the form

$$z \quad = \quad \sum_{n \in \mathbb{Z}_+} \sum_{l \in [\![r]\!]} \nu_l \frac{\rho_l^n}{n!} \xi_n,$$

where $r \in \mathbb{N}$ and ν_l, ρ_l are real numbers, for $l \in [\![r]\!]$.
One begins by checking that $\mathcal{Z}$ is dense in $l^2(\mathbb{Z}_+)$, so that

$$\forall \ t \geqslant 0, \qquad \| \exp(tM) \| \quad = \quad \sup_{z \in \mathcal{Z} \setminus \{0\}} \frac{\|\exp(tM) z\|}{\|z\|}$$

For z as above, denote ν and ρ the vectors of coordinates $(\nu_l)_{l \in [\![r]\!]}$ and $(\rho_l)_{l \in [\![r]\!]}$.

Upper bound (2)

We get:

$$\|z\|^2 \quad = \quad \nu' A(\rho)\nu,$$

where $A(\rho)$ is the $r \times r$-matrix given by

$$\forall\ k, l \in [\![r]\!], \qquad A_{k,l}(\rho) \quad := \quad \exp\left(\rho_k \rho_l + 2c(\rho_k + \rho_l)\right)$$

Similarly

$$\|\exp(tM)z\|^2 \quad = \quad \exp(-2c^2 t)\nu' A(\exp(-t)\rho)\nu$$

So next result enables to conclude:

Lemma

For any $t \geqslant 0$, any $r \in \mathbb{N}$ and any $\nu = (\nu_k)_{k \in [\![r]\!]}, \rho = (\rho_k)_{k \in [\![r]\!]} \in \mathbb{R}^r$, we have

$$\nu' A(\exp(-t)\rho)\nu \quad \leqslant \quad \exp\left(-4c^2 \frac{1 - \exp(-t)}{1 + \exp(-t)}\right) \nu' A(\rho)\nu$$

Optimizing functions

It is possible to find the functions where $\|P_t^{(a)} - \mu_a\|$ is attained.

• If $\|P_t^{(a)} - \mu_a\| = \exp(-t)$, the elements of $\mathcal{U}_1 \backslash \{0\}$ are maximizing, for instance the mapping $\mathbb{T} \times \mathbb{R} \ni (x, y) \mapsto y$.

• If $\|P_t^{(a)} - \mu_a\| > \exp(-t)$, the maximizing functions belong to $\mathcal{V}_1 \oplus \mathcal{W}_1$ and correspond to

$$\mathbb{T} \times \mathbb{R} \ni (x, y) \quad \mapsto \quad \exp\left(-\frac{2iy}{1 + \exp(-t)} + i(x + y)\right)$$

Stable subspaces

We follow the same approach, but it will be disturbed by the critical value $a = 1/4$ for the spectrum to be real.
Change of notation for the Hermite polynomial, due to the invariant measure $\widetilde{\mu}_a = \gamma_{1/a} \otimes \gamma$:

$$\forall\ p \in \mathbb{N},\ \forall\ x \in \mathbb{R}, \qquad h_{p,a}(y) \ := \ h_p(\sqrt{a}x)$$

Looking at the action of $\widetilde{L}_a$ on $h_{p,a} \otimes h_q$, it appears that for $n \in \mathbb{Z}_+$, the space

$$H_n \ := \ \mathrm{Vect}(h_{p,a} \otimes h_{n-p},\ p \in [\![0, n]\!])$$

is stable and clearly we have

$$\mathbb{L}^2(\widetilde{\mu}_a) \ = \ \bigoplus_{n \in \mathbb{Z}_+} H_n$$

where the components are orthogonal.

Reduction to matrices

In the orthonormal basis $(h_{p,a} \otimes h_{n-p})_{p \in [\![0,n]\!]}$, the matrix of the restriction of $\widetilde{L}_a$ to H_n is given by

$$\widetilde{M}_n = \begin{pmatrix} -n & \sqrt{an} & 0 & \cdots & 0 \\ -\sqrt{an} & -(n-1) & \sqrt{a2(n-1)} & & \vdots \\ 0 & -\sqrt{a2(n-1)} & -(n-2) & \ddots & 0 \\ & & \ddots & \ddots & \sqrt{an} \\ 0 & \cdots & 0 & -\sqrt{an} & 0 \end{pmatrix}$$

Again we decompose it into its diagonal, supdiagonal and subdiagonal parts, $\widetilde{M}_n = \widetilde{D}_n + \sqrt{a} S_n - \sqrt{a} S_n^*$. It will be more convenient to consider

$$\begin{aligned} D_n &:= \widetilde{D}_n + \frac{n}{2} I_n \\ M_n &:= \widetilde{M}_n + \frac{n}{2} I_n, \end{aligned}$$

The Lie algebra $\mathfrak{sl}(2,\mathbb{R})$

Let V_n be the vector space generated by the three matrices D_n, S_n and S_n^*, it is a 3-dimensional Lie algebra, since

$$\begin{aligned}
[S_n, S_n^*] &= -2D_n \\
[S_n, D_n] &= S_n \\
[S_n^*, D_n] &= -S_n^*.
\end{aligned}$$

One recognizes the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$. Indeed, for $n = 1$, $-D_1, S_1/\sqrt{2}$ and $S_1^*/\sqrt{2}$ form the usual basis of $\mathfrak{sl}(2,\mathbb{R})$:

$$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Curiously, $(V_n)_{n\in\mathbb{N}}$ is the family of all irreducible representations of $\mathfrak{sl}(2,\mathbb{C})$.
We are led to consider $\mathrm{ad}_{M_n} : V_n \ni X \mapsto [M_n, X] \in V_n$.

Diagonalization of ad_{M_n}

Lemma

Let $n \in \mathbb{N}\backslash\{1\}$ be fixed. The kernel of the operator ad_{M_n} is generated by M_n. For $a \neq 1/4$, there are two other eigenvalues, θ and $-\theta$ where

$$\theta \quad := \quad \begin{cases} \sqrt{1-4a} & , \text{ if } a \in [0, 1/4) \\ \sqrt{4a-1}\,i & , \text{ if } a > 1/4 \end{cases}$$

The corresponding eigenspaces are respectively generated by

$$\begin{aligned} J_+ &= 4\sqrt{a}D_n + (1-\theta)S_n - (1+\theta)S_n^* \\ J_- &= 4\sqrt{a}D_n + (1+\theta)S_n - (1-\theta)S_n^* \end{aligned}$$

For $a = 1/4$, the operator ad_{M_n} is not diagonalizable and its matrix is equal to the 3×3 Jordan block $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ associated to the eigenvalue 0, in the basis $(M_n, D_n - 2\sqrt{a}S_n^*, -2\sqrt{a}S_n^*)$.

Diagonalization of M_n (1)

Consider the case $a \neq 1/4$, the matrices J_+ and J_- play the same role as before. Due to the fact that they are finite tridiagonal matrices, it is easy to deduce that their kernels are at most one-dimensional. They are indeed one dimensional, otherwise one would be able to construct an infinity of eigenvalues for M_n. We deduce that the matrix M_n is diagonalizable and all its eigenvalues have multiplicity 1. More precisely if λ is an eigenvalue of M_n such that $\lambda + \theta$ is not an eigenvalue of M_n, then the spectrum of M_n is the set $\{\lambda - k\theta \,:\, k \in [\![0, n]\!]\}$. Furthermore, for $k \in [\![1, n]\!]$, J_+ (respectively J_-) transforms the spectral line associated to $\lambda - k\theta$ (resp. $\lambda - (k-1)\theta$) into the spectral line associated to $\lambda - (k-1)\theta$ (resp. $\lambda - k\theta$).

Diagonalization of M_n (2)

To end the determination of the spectrum of M_n, note that M_n is skew-centrosymmetric, i.e. $\mathcal{T}(M_n) = -M_n$, where for any $(n+1) \times (n+1)$ matrix $M = (M_{k,l})_{k,l \in [\![0,n]\!]}$,

$$\forall\ k, l \in [\![0, n]\!], \qquad (\mathcal{T}(M))_{k,l} \ := \ M_{n-k,n-l}.$$

We deduce that the spectrum of M_n is symmetrical with respect to zero. The first part of next result follows.

Proposition

For $a \neq 1/4$, the spectrum of M_n is $\{(k - n/2)\theta \ : \ k \in [\![0, n]\!]\}$. For $a = 1/4$, M_n is similar to the Jordan block of size $n+1$ associated to the eigenvalue 0 (in particular M_n is not diagonalizable for $n \geqslant 1$).

This result was already obtained by Risken [89], by a slightly different approach.

Back to finite dimensional hypocoercivity

As before, we don't want to stop with the spectral decomposition of $\widetilde{L}_a$, but to compute the operator norms of the associated semi-group. Note that it is sufficient to work on the H_n: for any $t \geqslant 0$,

$$\begin{aligned} \|\widetilde{P}_t^{(a)} - \widetilde{\mu}_a\|^2_{\mathbb{L}^2(\widetilde{\mu}_a)\circlearrowleft} &= \sup_{n\in\mathbb{N}} \|\widetilde{P}_t^{(a)}\|^2_{H_n\circlearrowleft} \\ &= \sup_{n\in\mathbb{N}} \exp\left(-\frac{n}{2}t\right) \|\exp(tM_n)\| \end{aligned}$$

First we investigate the case $a \in (0, 1/4)$, so that $\theta = \sqrt{1-4a} \in \mathbb{R}$. Let ξ_0 be a normalized vector generating the kernel of J_- and for all $p \in [\![1, n]\!]$, $\xi_p := \left(-\frac{\sqrt{a}}{2\theta^2}\right)^p J_+^p \xi_0$, so that $(\xi_p)_{p\in[\![0,n]\!]}$ is a family of eigenvectors of M_n associated to the eigenvalues $((p-n/2)\theta)_{p\in[\![0,n]\!]}$.

On ξ_0

Again $\xi_0^2 = (\xi_0^2(p))_{p\in[\![0,n]\!]}$ corresponds to a well-known law, the binomial distribution $\beta_{(1-\theta)/2}$ of parameter $(1-\theta)/2$:

$$\forall\ p \in [\![0,n]\!], \qquad \xi_0^2(p) \quad = \quad \binom{n}{p}\left(\frac{1-\theta}{2}\right)^p\left(\frac{1+\theta}{2}\right)^{n-p}$$

To be able to compute conveniently, we look for a formula similar to (2). It is more involved, but we end up with

$$\forall\ p \in [\![0,n]\!], \qquad \xi_p \quad = \quad P_p\left(\widetilde{D}_n + nI_n\right)\xi_0$$

where

$$P_p(X) \quad := \quad \sum_{k\in[\![0,p]\!]} \frac{p^{(p-k)}(n-k)^{(p-k)}}{(p-k)!}\left(\frac{1}{2}\left(\frac{1}{\theta}-1\right)\right)^{p-k}\Pi_k(X)$$

$$\Pi_p(X) \quad := \quad \prod_{k\in[\![0,p-1]\!]}(X-k)$$

Convenient vectors

The interest of the previous formula is that

$$\forall\ p, q \in [\![0, n]\!], \qquad \langle \xi_p, \xi_q \rangle \ = \ \beta_{(1-\theta)/2}[P_p P_q]$$

As a consequence, if for any $\rho \in \mathbb{R}$, we consider $z(\rho) := \sum_{p \in [\![0,n]\!]} \frac{\rho^p}{p!} \xi_p$, then for any $\widetilde{\rho}, \widehat{\rho} \in \mathbb{R}$, we have

$$\langle z(\widetilde{\rho}), z(\widehat{\rho}) \rangle \ = \ \left(1 + \frac{1-\theta^2}{2\theta}(\widetilde{\rho} + \widehat{\rho}) + \frac{1-\theta^2}{4\theta^2}\widetilde{\rho}\widehat{\rho}\right)^n$$

Then working in a similar spirit as in the first toy model and after quite fastidious computations, we get the result announced in the introduction.

Other situations for a

- The case $a = 1/4$ is obtained through the limit $a \to 1/4_-$.
- For $a > 1/4$, the previous computations can be adapted. Some of the properties are purely algebraic and no modification is required, such as the definition of the polynomials $P_p(X)$, their coefficients are now complex. It appears that for any $p, q \in [\![0, n]\!]$,

$$\langle \xi_p, \xi_q \rangle \quad = \quad (1 + |\theta|^2)^{n/2} \beta_{1/2}[P_p \overline{P_q}]$$

We deduce that for any $\widetilde{\rho}, \widehat{\rho} \in \mathbb{C}$, we have

$$\langle z(\widetilde{\rho}), z(\widehat{\rho}) \rangle \quad = \quad \left(1 + |\theta|^2\right)^{-n/2} \left(\gamma + \left(1 + \frac{\widetilde{\rho}}{\delta}\right) \overline{\left(1 + \frac{\widehat{\rho}}{\delta}\right)}\right)^n$$

with $\delta := \frac{2\theta}{1+|\theta|^2}$. This enables to conclude after some more computations.

For any $a > 0$, the maximizing functions are linear, the situation seems different from the first model.

Global decompositions (1)

Decompositions of the form $D + cS - cS^*$ on the stable subspaces, which were the beginning of our developments, can be lifted up to the generators.
In the first model, one gets $L_a = K + R - R^*$, with

$$\begin{aligned} K &= a\partial_y^2 - y\partial_y \\ R &= y\partial_x - a\partial_x\partial_y \\ R^* &= -a\partial_x\partial_y \end{aligned}$$

satisfying

$$[K, R] = R \quad \text{and} \quad [R, R^*] = aJ$$

where $J = \partial_x^2$ has the missing coercitivity on $\mathbb{T}$.

Global decompositions (2)

Similarly for the second model, we can write $\widetilde{L}_a = K + R - R^*$, with

$$\begin{aligned} K &= \partial_y^2 - y\partial_y \\ R &= y\partial_x - \partial_x\partial_y \\ R^* &= ax\partial_y - \partial_x\partial_y \end{aligned}$$

satisfying

$$[K, R] = R \quad \text{and} \quad [R, R^*] = J - aK$$

where $J := \partial_x\partial_x^* = \partial_x^2 - ax\partial_x$ is "the" missing coercive Ornstein-Ulhenbeck operator on $\mathbb{R}$.

Global decompositions (3)

More generally, given a smooth potential $U : \mathbb{T} \to \mathbb{R}$, the kinetic operator

$$L \quad := \quad y\partial_x - U'(x)\partial_y + \partial_y^2 - y\partial_y$$

can de decomposed into $L = K + R - R^*$, where

$$\begin{aligned} K &= \partial_y^2 - y\partial_y \\ R &= y\partial_x - \partial_x\partial_y \\ R^* &= U'(x)\partial_y - \partial_x\partial_y \end{aligned}$$

satisfying

$$[K, R] = R \quad \text{and} \quad [R, R^*] = J - U''K$$

where $J := \partial_x^*\partial_x = \partial_x^2 - U'(x)\partial_x$ is the usual coercive Langevin operator associated to U on $\mathbb{T}$.

Toward an alternative approach? (1)

Note the difference with traditional approaches, where brackets of first order operators are preferred (Hörmander's conditions).

How to use the previous relations to get hypocoercive bounds? In view of the behavior in t^3 for small times $t > 0$, the first idea is to differentiate three times instead of once. More precisely, for $f \in \mathbb{L}^2(\mu)$ with $\mu[f] = 0$, denote for $t \geqslant 0$, $f_t := P_t[f]$ and $F_t := \mu[f_t^2]$. The usual method to show that this expression ends up converging exponentially fast toward 0 is to add terms to F_t (typically $\langle \partial_x f_t, \partial_y f_t \rangle = \langle f_t, R f_t \rangle$) to get a functional satisfying a Gronwall inequality. We would like to work with F_t only. So let us differentiate it:

Toward an alternative approach? (2)

$$\begin{aligned} F_t' &= 2\langle Kf_t, f_t\rangle, \\ F_t'' &= 4\left\langle K^2 f_t, f_t\right\rangle - 4\langle f_t, Rf_t\rangle, \\ F_t''' &= 8\left\langle K^3 f_t, f_t\right\rangle - 24\langle Kf_t, Rf_t\rangle - 12\langle f_t, Rf_t\rangle + 4\langle [R, R^*] f_t, f_t\rangle \end{aligned}$$

The term $[R, R^*]$ contains the missing coercivity. So for quite a long time, we tried to find $A, B, C > 0$ so that

$$AF_t + BF_t' + CF_t'' + F_t''' \leqslant 0$$

(it is sufficient to consider the time $t = 0$), before proving that it is not possible!

The interest of the first toy model is that it can serve as a prototype: up to a change of the constants $A, B, C > 0$, it would have been enough to obtain the above bound for this example.