Sturm's law of large numbers for the L^1-Karcher mean of positive operators

Miklós Pálfia

Department of Mathematics
Sungkyunkwan University
&
Bolyai Institute, Interdisciplinary Excellence Centre
University of Szeged

July 8, 2019

palfia.miklos@aut.bme.hu
joint work with Yongdo Lim

Introduction

In this talk, E will denote a Hilbert space

- $\mathbb{S}(E)$ denote the real Banach space of bounded, linear, self-adjoint operators over E
- $\mathbb{P}(E) \subseteq \mathbb{S}(E)$ denotes the cone of invertible positive definite and $\hat{\mathbb{P}}(E)$ the cone of positive (semi-definite) operators

$\mathbb{S}$ and $\mathbb{P}$ are partially ordered cones with the positive definite, also called Loewner order:

$$A \leq B \text{ iff } B - A \text{ is positive semidefinite, i.e.}$$

$$v^*(B - A)v \geq 0 \text{ for all } v \in E$$

The Riemannian geometry of positive definite matrices

Assume that E is finite dimensional. Then $\mathbb{P}(E)$ has a Riemannian structure with tangent space $\mathbb{S}(E)$:

$$\begin{aligned} \langle X, Y\rangle_A &= Tr\left\{A^{-1}XA^{-1}Y\right\}, \\ d^2(A,B) &= \langle \log_A(B), \log_A(B)\rangle_A = Tr\left\{\log^2(A^{-1/2}BA^{-1/2})\right\}, \\ \exp_A(X) &= A^{1/2}\exp(A^{-1/2}XA^{-1/2})A^{1/2}, \\ \log_A(B) &= A^{1/2}\log(A^{-1/2}BA^{-1/2})A^{1/2} \end{aligned}$$

for $X \in \mathbb{S}(E)$, $A, B \in \mathbb{P}(E)$.

Theorem

The space $(\mathbb{P}(E), d)$ is an NPC or CAT(0)-space, i.e. for each pair $A, B \in \mathbb{P}(E)$ there exists a unique $Z \in \mathbb{P}(E)$, s.t. for all $X \in \mathbb{P}(E)$

$$d^2(X,Z) \leq \frac{1}{2}d^2(X,A) + \frac{1}{2}d^2(X,B) - \frac{1}{4}d^2(A,B).$$

Moreover in the above $Z = A\#B$, i.e. the geometric mean is the unique midpoint. Thus distance minimizing curves (geodesics) $\gamma : [0,1] \mapsto \mathbb{P}(E)$ are unique between any pair $A, B \in \mathbb{P}(E)$ and the NPC-inequality extends to the whole curve

$$d^2(X, \gamma(t)) \leq (1-t)d^2(X, A) + td^2(X, B) - t(1-t)d^2(A, B)$$

where $\gamma(t) = A\#_t B := A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^t A^{1/2}$ for $t \in [0,1]$. The map $A\#_t B$ is called the *weighted geometric mean*. Thus points of the curve $A\#_t B$ admit the characterization

$$A\#_t B = \arg\min_{X \in \mathbb{P}} (1-t)d^2(X, A) + td^2(X, B).$$

By Kubo-Ando $A\#_t B$ is monotone in the $A, B \in \hat{\mathbb{P}}(E)$ variables.

The multivariable geometric or Karcher mean

$\omega = (w_1, \ldots, w_n) \in \Delta_n$ probability vector, $w_i > 0$, $\sum_{i=1}^n w_i = 1$,
$\mathbb{A} = (A_1, \ldots, A_n)$, $A_i \in \mathbb{P}$

$$\Lambda(\omega; \mathbb{A}) := \arg\min_{X \in \mathbb{P}} \sum_{i=1}^{n} w_i d^2(X, A_i)$$

Moakher 2005, SIMAX:

- $X \mapsto d^2(X, A)$ is a strictly geodesically convex function by the NPC property
- $\Lambda(\omega; \mathbb{A})$ is the unique solution of the (Riemannian) gradient equation called the *Karcher equation*:

$$\nabla C(X) = -2 \sum_{i=1}^{n} w_i \log_X(A_i) = 0$$

where $C(X) = \sum_{i=1}^n w_i d^2(X, A_i)$.

Sturm's law of large numbers

Theorem (Sturm 2002, Annals of Prob.)

Let (X, d) be a $CAT(0)$-space and let $\mathcal{P}^2(X)$ denote the set of all probability measures μ s.t. $\int_X d^2(x,a)d\mu(a) < \infty$. Let $a\#_t b$ denote the unique geodesic between $a, b \in X$. Then for $\mu \in \mathcal{P}^2(X)$

$$\Lambda(\mu) := \arg\min_{x \in X} \int_X d^2(x,a)d\mu(a)$$

exists and is unique. Moreover consider an i.i.d. sequence of random variables $\{Y_i\}_{i\in\mathbb{N}}$ with law μ and define

$$S_1 := Y_1,$$
$$S_{k+1} := S_k \#_{\frac{1}{k+1}} Y_{k+1}.$$

Then S_k converges to $\Lambda(\mu)$ almost surely, if $\mathrm{supp}(\mu)$ is bounded.

Theorem (Wasserstein contraction property, Sturm 2002)

Moreover

$$d(\Lambda(\mu),\Lambda(\nu)) \leq W_1(\mu,\nu) := \inf_{\gamma\in\Pi(\mu,\nu)} \int_{X\times X} d(a,b)\,d\gamma(a,b)$$

where $\Pi(\mu,\nu)$ denotes the set of all Borel probability measures on $X\times X$ with marginals μ and ν.

Since the geometric mean $A\#_t B$ is monotone in (A,B), Sturm's result implies:

Theorem (Lawson-Lim 2011; Bhatia-Karandikar 2012, Math. Ann.)

For a fixed $\omega=(w_1,\dots,w_n)\in\Delta_n$ the map $\Lambda(\omega,\cdot)$ is monotone, i.e. for $A_i, B_i\in\mathbb{P}(E), 1\leq i\leq n$ with $A_i\leq B_i$ we have

$$\Lambda(\omega,\mathbb{A})\leq\Lambda(\omega,\mathbb{B}).$$

Nodice theorem for the Karcher mean

There is a deterministic version of Sturm's law of large numbers, first established for $\mathbb{P}(E)$ by (Holbrook 2012, J. Ramanujan M. S.):

Theorem (Lim-Pálfia 2014, Bull. LMS)

Let (X, d) be a CAT(0)-space and let $\mu := \sum_{i=0}^{n-1} \frac{1}{n}\delta_{a_i}$ with $a_i \in X$. Consider the deterministic sequence $\{S_k\}_{k\in\mathbb{N}}$ defined as the inductive sequence of geometric means

$$S_1 := a_0,$$
$$S_{k+1} := S_k \#_{\frac{1}{k+1}} a_{\overline{k}}$$

where $\overline{k} := k \mod (n)$. Then $S_k \to \Lambda(\mu)$ with rate $d(S_k, \Lambda(\mu)) = O(1/k)$.

The above along with Sturm's slln even generalizes to CAT(κ) spaces (Ohta-Pálfia 2015, Yokota 2018, Calc. Var. PDE).

The infinite dimensional case of positive operators

In the case of $\dim(E) = +\infty$, we no longer have Riemannian structure, so the manifold $\mathbb{P}(E)$ cannot be a $CAT(0)$-space. Thus the Karcher mean Λ is not known to admit a definition as a unique solution of an optimization problem. However we still have the corresponding gradient equation called the Karcher equation:

$$\sum_{i=1}^{n} w_i \log_X(A_i) = 0$$

for a probability vector $\omega \in \Delta_n$ and operators $A_i \in \mathbb{P}(E)$ for $1 \leq i \leq n$.

Thompson's part metric

For $A, B \in \mathbb{P}(E)$ let $d_\infty(A, B) := \| \log(A^{-1/2}BA^{-1/2})\|$.

Theorem (Thompson 1963, Proc. AMS)

The space $(\mathbb{P}(E), d_\infty)$ *is a complete metric space and* $d_\infty(A, B) = \log \max\{M(A \setminus B), M(B \setminus A)\}$ *where* $M(A \setminus B) := \inf\{\beta > 0 : B \leq \beta A\}$.

Lemma (Lim-Pálfia 2012, JFA)

The map $f : \mathbb{P}(E) \mapsto \mathbb{P}(E)$ *defined as*

$$f(X) := \sum_{i=1}^{n} w_i X\#_t A_i$$

is a strict contraction in $(\mathbb{P}(E), d_\infty)$ *for* $t \in (0, 1]$ *and* $\omega \in \Delta_n$ *and operators* $A_i \in \mathbb{P}(E)$ *for* $1 \leq i \leq n$. *That is* $d_\infty(f(X), f(Y)) \leq (1 - t)d_\infty(X, Y)$.

The power means

By Banach fixed point theorem and some calculation:

Theorem (Lim-Pálfia 2012, JFA and Lawson-Lim 2012, Trans. AMS)

For $t \in [-1, 1]$, $t \neq 0$ and $\omega \in \Delta_n$ and operators $A_i \in \mathbb{P}(E)$ for $1 \leq i \leq n$, the nonlinear operator equation

$$\sum_{i=1}^{n} w_i X \#_t A_i = X$$

has a unique solution in $\mathbb{P}(E)$ called the t-power mean $P_t(\omega; \mathbb{A})$ which is monotone in $\mathbb{A}$. Moreover P_t is also monotone in t, thus $\{P_t(\omega; \mathbb{A})\}_{t\in(0,1]}$ is a decreasing net, it has a greatest lower bound $\Lambda(\omega, \mathbb{A})$, and the strong operator $\lim_{t\to 0+} P_t(\omega; \mathbb{A}) = \Lambda(\omega, \mathbb{A})$.

The power means

The implicit function theorem with some trick implies:

Theorem (Lim-Pálfia 2012, JFA and Lawson-Lim 2012, Trans. AMS)

The strong operator $\lim_{t\to 0+} P_t(\omega;\mathbb{A}) = \Lambda(\omega,\mathbb{A})$ *is the unique solution of the Karcher equation*

$$\sum_{i=1}^{n} w_i \log_X(A_i) = 0$$

for $X \in \mathbb{P}(E)$ *and it is monotone in the second variable, i.e. if* $A_i \leq B_i$ *for* $1 \leq i \leq n$ *and* $A_i, B_i \in \mathbb{P}(E)$*, then* $\Lambda(\omega,\mathbb{A}) \leq \Lambda(\omega,\mathbb{B})$*.*

If all A_i mutually commute for $1 \leq i \leq n$, then

$$P_t(\omega;\mathbb{A}) = \left(\sum_{i=1}^{n} w_i A_i^t\right)^{1/t}.$$

The infinite dimensional L^1-Karcher mean

Let $\mathcal{P}^1(\mathbb{P}(E))$ denote the τ-additive Borel probability measures on $(\mathbb{P}(E), d_\infty)$ s.t. $\int_{\mathbb{P}(E)} d_\infty(X, A) d\mu(A) < \infty$ for $X \in \mathbb{P}(E)$. Note that τ-additive is equivalent to $\mu(\mathrm{supp}(\mu)) = 1$. Using the W_1-density of finitely supported probability measures in $\mathcal{P}^1(\mathbb{P}(E))$:

Theorem (Lim-Pálfia 2017, Lawson 2018)

For all $\mu \in \mathcal{P}^1(\mathbb{P}(E))$ there exists a unique solution of

$$\int_{\mathbb{P}(E)} \log_X(A) d\mu(A) = 0$$

denoted by $\Lambda(\mu)$, which satisfies

$$d_\infty(\Lambda(\mu), \Lambda(\nu)) \leq W_1(\mu, \nu)$$

for all $\nu \in \mathcal{P}^1(\mathbb{P}(E))$.

The stochastic partial order for $\mathcal{P}^1(\mathbb{P}(E))$

Definition (Kim-Lee 2015, LAA and Lawson 2017, JMAA)

A set $U \subseteq \mathbb{P}(E)$ is *upper* if for an $X \in \mathbb{P}(E)$ there exists an $Y \in U$ such that $Y \leq X$, then $X \in U$. Then the *stochastic order* for $\mu, \nu \in \mathcal{P}^1(\mathbb{P}(E))$ is defined as $\mu \leq \nu$ if $\mu(U) \leq \nu(U)$ for all upper sets $U \subseteq \mathbb{P}(E)$.

The stochastic order extends the usual positive definite order:

Proposition (Kim 2017, Taiwanese J. Math.)

Let $\mu := \sum_{i=1}^{n} \frac{1}{n}\delta_{A_i}, \nu := \sum_{i=1}^{n} \frac{1}{n}\delta_{B_i} \in \mathcal{P}^1(\mathbb{P}(E))$. *Then* $\mu \leq \nu$ *iff* $A_i \leq B_i$ *for all* $1 \leq i \leq n$.

Corollary

The L^1*-Karcher mean is monotone with respect to the stochastic order.*

Resolvent maps induced by the Karcher mean

Definition (Resolvent operator)

Given $\mu \in \mathcal{P}^1(\mathbb{P}(E))$ we define the resolvent operator for $\lambda > 0$ and $X \in \mathbb{P}(E)$ as

$$J_\lambda^\mu(X) := \Lambda\left(\frac{\lambda}{\lambda+1}\mu + \frac{1}{\lambda+1}\delta_X\right),$$

the solution of

$$\frac{\lambda}{\lambda+1}\int_{\mathbb{P}(E)} \log_Z A d\mu(A) + \frac{1}{\lambda+1}\log_Z(X) = 0$$

we obtained for $Z \in \mathbb{P}(E)$ in the previous theorem.

Note: if $\dim(E) < \infty$ then
$J_\lambda^\mu(X) = \arg\min_{Y\in\mathbb{P}(E)} \int_{\mathbb{P}(E)} d^2(A, Y)d\mu(A) + \frac{1}{\lambda}d^2(Y, X)$.

Proposition (Lim-Pálfia 2017)

Let $\mu \in \mathcal{P}^1(\mathbb{P})$, for $\tau > \lambda > 0$, $\lambda_i > 0$ and $X, Y \in \mathbb{P}$. Then we have

1. *(Resolvent contraction)*

$$d_\infty(J_\lambda^\mu(X), J_\lambda^\mu(Y)) \leq \frac{1}{1+\lambda} d_\infty(X, Y).$$

2. *(Resolvent identity)*

$$J_\tau^\mu(X) = J_\lambda^\mu \left(J_\tau^\mu(X) \#_{\frac{\lambda}{\tau}} X \right).$$

3. *(Whatever inequality)*

$$d_\infty \left(J_{\lambda_1}^\mu \circ \cdots \circ J_{\lambda_n}^\mu(X), X \right) \leq \sum_{i=1}^{n} \frac{\lambda_i}{1+\lambda_i} \int_{\mathbb{P}} d_\infty(X, A) d\mu(A).$$

A key metric inequality for resolvent iterates

Theorem (Lim-Pálfia 2017 & 2019)

Let $\{t_i\}_{i\in\mathbb{N}}, \{\hat{t}_j\}_{j\in\mathbb{N}}$ denote sequences such that $t_i, \hat{t}_j \in [0,\infty)$ and $t_{i+1} > t_i, \hat{t}_{j+1} > \hat{t}_j$. Let $X \in \mathbb{P}$ and $\mu \in \mathcal{P}^1(\mathbb{P})$ and $\tau_i := t_i - t_{i-1}$, $\hat{\tau}_i := \hat{t}_i - \hat{t}_{i-1}$. Let $X_0 = \hat{X}_0 = X$ and define $X_{i+1} := J^{\mu}_{\tau_{i+1}}(X_i)$ and $\hat{X}_{i+1} := J^{\mu}_{\hat{\tau}_{i+1}}(\hat{X}_i)$. Let $m, n \in \mathbb{N}$. Then

$$d_\infty(X_m, \hat{X}_n) \leq \left[\prod_{j=1}^{\min\{m,n\}} (1 + \min\{\tau_j, \hat{\tau}_j\})^{-1}\right] \times \left[(t_m - \hat{t}_n)^2 + \sigma_m + \hat{\sigma}_n\right]^{1/2} C$$

where $\sigma_m = \sum_{i=1}^{m}(t_i - t_{i-1})^2, \hat{\sigma}_n = \sum_{i=1}^{n}(\hat{t}_i - \hat{t}_{i-1})^2$ and $C = \int_{\mathbb{P}} d_\infty(X, A) d\mu(A)$.

An ODE flow for the L^1-Karcher mean

Theorem (Lim-Pálfia 2017)

For any $X, Y \in \mathbb{P}(E)$ *and* $t > 0$ *the curve*

$$S(t)X := \lim_{n\to\infty} \left(J^{\mu}_{t/n}\right)^{n}(X)$$

exists where the limit is in the d_∞*-topology and it is Lipschitz-continuous on compact time intervals* $[0, T]$ *for any* $T > 0$*. Moreover it satisfies the exponential contraction property*

$$d_\infty\left(S(t)X, S(t)Y\right) \leq e^{-t} d_\infty(X, Y),$$

and for $s > 0$ *verifies the semigroup property*

$$S(t+s)X = S(t)(S(s)X).$$

An ODE flow for the L^1-Karcher mean

Theorem (Lim-Pálfia 2017)

Let $\mu \in \mathcal{P}^1(\mathbb{P}(E))$ and $X \in \mathbb{P}(E)$. Then for $t > 0$, the curve $X(t) := S(t)X$ provides a strong solution of the Cauchy problem

$$X(0) := X,$$
$$\dot{X}(t) = \int_{\mathbb{P}} \log_{X(t)} A d\mu(A),$$

where the derivative $\dot{X}(t)$ is the Fréchet-derivative.

The finite dimensional case follows from the theory of gradient flows of semi-convex functions $\phi : M \mapsto \mathbb{R}$ in CAT(κ) spaces (M, d) (Ohta-Pálfia 2017, Amer. J. Math.) with the resolvent $J_\lambda^\mu(X) := \arg\min_{Y \in M} \phi(Y) + \frac{1}{2\lambda} d^2(Y, X)$.

Continuous-time slln for the L^1-Karcher mean

Theorem (Lim-Pálfia 2017)

Let $\mu \in \mathcal{P}^1(\mathbb{P}(E))$ *and let* $\{Y_i\}_{i\in\mathbb{N}}$ *be a sequence of i.i.d.* $\mathbb{P}(E)$*-valued random variables with law* μ*. Let* $\mu_n := \sum_{i=1}^{n} \frac{1}{n}\delta_{Y_i} \in \mathcal{P}^1(\mathbb{P}(E))$ *denote the empirical measures. Let* $S^{\mu}(t)$ *and* $S^{\mu_n}(t)$ *denote the semigroups corresponding to the resolvents* J_λ^{μ} *and* $J_\lambda^{\mu_n}$ *for* $t > 0$*. Then*

$$\lim_{n\to\infty} S^{\mu_n}(t) = S^{\mu}(t)$$

almost surely uniformly in d_∞ *on compact time intervals. Moreover let* $F_\rho^{\mu_n} := J_{\rho/n}^{\delta_{Y_n}} \circ \cdots \circ J_{\rho/n}^{\delta_{Y_1}}$ *where* $J_\rho^{\delta_A}(X) = X \#_{\frac{\rho}{\rho+1}} A$*. Then*

$$\lim_{m\to\infty} (F_{t/m}^{\mu_n})^m = S^{\mu_n}(t)$$

uniformly in d_∞ *on compact time intervals for any* $n \in \mathbb{N}$*.*

The idea behind: Approximating resolvents and semigroups

Lemma (Lim-Pálfia 2017)

Let $F : \mathbb{P} \mapsto \mathbb{P}$ be a nonexpansive map with respect to d_∞. Let $\lambda, \rho > 0$ and $Y \in \mathbb{P}$. Then the map

$$G_{\lambda,\rho,Y}(X) := \Lambda\left(\frac{1}{1+\lambda/\rho}\delta_Y + \frac{\lambda/\rho}{1+\lambda/\rho}\delta_{F(X)}\right)$$

is a strict contraction with Lipschitz constant $\frac{\lambda/\rho}{1+\lambda/\rho} < 1$. Thus the map $G_{\lambda,\rho,Y}$ has a unique fixed point denoted by $J_{\lambda,\rho}(Y)$.

Theorem (Lim-Pálfia 2017)

Let $F : \mathbb{P} \mapsto \mathbb{P}$ be a nonexpansive map. Then for any $X, Y \in \mathbb{P}$ and $t, \rho > 0$ the curve

$$S_\rho(t)X := \lim_{n\to\infty} \left(J_{t/n,\rho}\right)^n (X)$$

Theorem (cont.)

exists where the limit is in the d_∞-topology. It also satisfies the contraction property

$$d_\infty\left(S_\rho(t)X, S_\rho(t)Y\right) \leq d_\infty(X,Y),$$

for $s>0$ verifies the semigroup property

$$S_\rho(t+s)X = S_\rho(t)(S_\rho(s)X),$$

and the flow operator $S_\rho : \mathbb{P}\times(0,\infty)\mapsto\mathbb{P}$ extends by d_∞-continuity to $S_\rho : \mathbb{P}\times[0,\infty)\mapsto\mathbb{P}$.

Convergence of approximating resolvents

Theorem (Lim-Pálfia 2017)

Let $\mu = \sum_{i=1}^{n} \frac{1}{n}\delta_{A_i} \in \mathcal{P}^1(\mathbb{P})$. *For* $\rho > 0$ *let* $F_\rho := J_{\rho/n}^{\delta_{A_n}} \circ \cdots \circ J_{\rho/n}^{\delta_{A_1}}$. *In particular* $F_\rho : \mathbb{P} \mapsto \mathbb{P}$ *is a contraction with respect to* d_∞. *For* $\lambda > 0$ *let* $J_{\lambda,\rho}$ *denote the approximating resolvent of* F_ρ. *Then*

$$J_\lambda^\mu(X) = \lim_{\rho \to 0+} J_{\lambda,\rho}(X) \quad \text{in norm, thus } d_\infty.$$

Theorem (Trotter formula, Lim-Pálfia 2017)

For each $\rho > 0$ *let* $F_\rho : \mathbb{P} \mapsto \mathbb{P}$ *be a nonexpansive map and let* $J_{\lambda,\rho}$ *be its approximating resolvent. If* $\lim_{\rho \to 0+} J_{\lambda,\rho}(X) = J_\lambda^\mu(X)$ *in* d_∞ *for all* $X \in \mathbb{P}$, *then*

$$(F_{\frac{t}{n}})^n(X) \to S(t)X$$

in d_∞ *for all* $X \in \mathbb{P}$ *as* $n \to \infty$, *where* $S(t)$ *is generated by* J_λ^μ.

Technical background and the Lawson-Lim conjecture

For $\mu \in \mathcal{P}^1(\mathbb{P})$ and $X \in \mathbb{P}$ let

$$\phi_\mu(X) := \int_{\mathbb{P}} \log_X A d\mu(A).$$

Proposition (Lim-Pálfia 2017)

Let $\mu \in \mathcal{P}^1(\mathbb{P}(\mathcal{H}))$ and $X \in \mathbb{P}(\mathcal{H})$. Then the Fréchet derivative $D\phi_\mu(\Lambda(\mu)) : \mathbb{S}(\mathcal{H}) \mapsto \mathbb{S}(\mathcal{H})$ if exists, is a Banach space isomorphism.

Theorem (Continuity of P_t, Lawson-Lim conjecture, Lim-Pálfia 2017)

Let $\mu \in \mathcal{P}^1(\mathbb{P})$ with $\int_{\mathbb{P}} d_\infty^p(X, A)d\mu(A) < +\infty$ for all $p \geq 1$ and $X \in \mathbb{P}$. Then the family $P_t(\mu)$ is norm continuous in $t \in [-1, 1]$, in particular

$$\Lambda(\mu) = \lim_{t \to 0} P_t(\mu) \quad \textit{in norm.}$$

Discrete-time flows for Λ

For $\mu := \sum_{i=1}^{n} \frac{1}{n}\delta_{Y_i} \in \mathcal{P}^1(\mathbb{P}(E))$ and $F_\rho^\mu := J_{\rho/n}^{\delta_{Y_n}} \circ \cdots \circ J_{\rho/n}^{\delta_{Y_1}}$ we obtained

$$\lim_{m\to\infty} (F_{t/m}^\mu)^m = S^\mu(t).$$

It's similar to Holbrook's nodice approximation...

Proposition (Resolvent iteration, Lim-Pálfia 2019)

Let $\mu \in \mathcal{P}^1(\mathbb{P})$, $d \geq 0$ an integer and $X \in \mathbb{P}$. Let $X_0 := X$ and define $X_{k+1} := J_{1/(k+d)}^\mu(X_k)$ for $k \in \mathbb{N}$. Then $d_\infty(X_k, \Lambda(\mu)) \to 0$.

Proof.

Let $t_k := \sum_{i=d+1}^{k} \frac{1}{i}$ and $\hat{t}_j := j\frac{t_n}{n}$ for a fixed $n \in \mathbb{N}$ such that $\lfloor n/t_n \rfloor \geq d$, so that $\tau_j = \frac{1}{j+d}$ and $\hat{\tau} = \hat{\tau}_j = \frac{t_n}{n}$ for all $1 \leq j \leq n$. Let $S(t)$ denote the semigroup generated by J^μ. Then

$$d_\infty(X_n, \Lambda(\mu)) \leq d_\infty(X_n, \left(J_{\hat{\tau}}^\mu\right)^n(X)) + d_\infty(\left(J_{\hat{\tau}}^\mu\right)^n(X), S(t_n)X) + d_\infty(S(t_n)X, \Lambda(\mu)) \leq O(1/\log(n)).$$

Discrete-time flows: Nodice theorem for Λ

Theorem (Nodice, Lim-Pálfia 2019)

Let $\frac{1}{n}\sum_{i=1}^{n}\delta_{Y_i} =: \mu \in \mathcal{P}^1(\mathbb{P})$ for a fixed integer n. Let $S_1 := Y_1$ and $S_{k+1} := S_k \#_{\frac{1}{k+1}} Y_{\overline{k+1}}$, where $\overline{k}$ is defined to equal the residual of $k \mod n$ and the 0 residual is identified with n. Then $S_k \to \Lambda(\mu)$ in d_∞ with convergence rate at least $O(1/\log(k))$.

sketch of the proof.

Let $1 \leq i \leq n$ be arbitrary and $N \in \mathbb{N}$ such that $\frac{2}{Nn}\mathrm{diam}(\mathrm{supp}(\mu)) < 1$. Let $k \geq N$ be an integer. By definition

$$\frac{1}{nk+i+1}\log_{S_{nk+i+1}} Y_{\overline{nk+i+1}} + \log_{S_{nk+i+1}}(S_{nk+i}) = 0.$$

By expanding log in the above, we get

proof contd.

$$\frac{1}{nk+i+1}\log_{S_{nk+i+1}} Y_{\overline{nk+i+1}} + S_{nk+i} - S_{nk+i+1} + O\left(4\frac{\mathrm{diam}(\mathrm{supp}(\mu))^2}{(nk+i+1)^2}\right) e^{d_\infty(Y_1,I)+2\mathrm{diam}(\mathrm{supp}(\mu))} = 0.$$

Summing the above identity in $0 \leq i \leq n-1$, we get

$$S_{kn} - S_{(k+1)n} + \sum_{i=0}^{n-1} \frac{1}{nk+i+1}\log_{S_{nk+i+1}} Y_{\overline{nk+i+1}} + O\left(4\frac{\mathrm{diam}(\mathrm{supp}(\mu))^2}{(nk+i+1)^2}\right) e^{d_\infty(A_1,I)+2\mathrm{diam}(\mathrm{supp}(\mu))} = 0.$$

proof contd.

There exists an $\overline{S}_k \in \mathbb{P}$ such that

$$\frac{1}{k+1}\sum_{i=1}^{n}\frac{1}{n}\log_{S_{n(k+1)}} Y_{\overline{nk+i}} + \log_{S_{n(k+1)}} \overline{S}_k = 0 \qquad (*)$$

and $d_\infty(S_{nk}, \overline{S}_k) \leq O\left(\frac{1}{k^2}\right)$ for all $k \geq N$. In other words $S_{n(k+1)} = J^{\mu}_{\frac{1}{k+1}}\overline{S}_k$. Now let $\hat{S}_k := \overline{S}_N$ for $1 \leq k \leq N$, and $\hat{S}_{k+1} := J^{\mu}_{\frac{1}{k+1}}\hat{S}_k$ for $k \geq N$ recursively. Then

$$\begin{aligned}
d_\infty(S_{n(k+1)}, \hat{S}_{k+1}) &\leq \frac{1}{1+\frac{1}{k+1}} d_\infty(\overline{S}_k, \hat{S}_k) \\
&\leq \frac{1}{1+\frac{1}{k+1}}\left(d_\infty(S_{kn}, \hat{S}_k) + d_\infty(S_{kn}, \overline{S}_k)\right) \\
&\leq \left(1 - \frac{1}{k+2}\right) d_\infty(S_{kn}, \hat{S}_k) + O\left(\frac{1}{(k+2)^2}\right).
\end{aligned}$$

proof contd.

It follows that $d_\infty(S_{n(k+1)}, \hat{S}_{k+1}) \to 0$. In particular, since $\hat{S}_k \to \Lambda(\mu)$ the assertion is proved for the subsequence $\{S_{kn}\}_{k\in\mathbb{N}}$. The convergence of the rest follows from

$$\begin{aligned} d_\infty(S_{nk}, S_{kn+i}) &\leq \sum_{j=0}^{i} \frac{2\mathrm{diam}(\mathrm{supp}(\mu))}{nk+j} \\ &\leq n\frac{2\mathrm{diam}(\mathrm{supp}(\mu))}{nk} \\ &= \frac{2\mathrm{diam}(\mathrm{supp}(\mu))}{k} \end{aligned}$$

valid for any $1 \leq i \leq n$. □

Stochastic discrete-time flows: Sturm's L^1-slln

Theorem (Sturm's L^1-slln for Λ, Lim-Pálfia 2019)

Let $\mu \in \mathcal{P}^1(\mathbb{P})$ and let Y_i be an i.i.d. sequence of random variables with law μ. Let $S_1 := Y_1$ and $S_{k+1} := S_k \#_{\frac{1}{k+1}} Y_{k+1}$. Then $S_k \to \Lambda(\mu)$ a.s. in d_∞.

idea of the proof.

The idea is that given an L^1-random variable, some truncation of it with finite/bounded support can be made arbitrarily close to it in W^1 in a controlled manner. So then the metric convexity yields that their inductive mean sequences are arbitrarily close to each other a.s.:

Stochastic discrete-time flows: some technicalities 1.

Lemma (Lim-Pálfia 2019)

Let $\epsilon > 0$ and $\mu \in \mathcal{P}^1(\mathbb{P})$. Then there exists an $R > 0$ such that

$$\limsup_{n\to\infty} d_\infty(X_n, X_n^R) < \epsilon$$

almost surely, where $X_1 := Y_1$, $X_1^R := Y_1^R$ and recursively $X_{n+1} := X_n \#_{\frac{1}{n+1}} Y_{n+1}$, $X_{n+1}^R := X_n^R \#_{\frac{1}{n+1}} Y_{n+1}^R$, where Y_n is an i.i.d. sequence of $\mathbb{P}$-valued random variables with law μ and

$$Y_n^R(\omega) := \begin{cases} Y_n(\omega), & \text{if} \quad d_\infty(Y_n(\omega), \Lambda(\mu)) < R, \\ \Lambda(\mu), & \text{if} \quad d_\infty(Y_n(\omega), \Lambda(\mu)) \geq R. \end{cases}$$

Thus, it suffices to show the assertion for r.v. with finite/bounded support.

Stochastic discrete-time flows: some technicalities 2.

Then, we combine a stochastic version of $(*)$ in the proof of the Nodice theorem with the following lemmata.
I. Empirical measures are uniformly close in W_1 a.s.:

Lemma (Lim-Pálfia 2019)

Let $\mu \in \mathcal{P}^1(\mathbb{P})$ and let Y_n be a sequence of i.i.d. $\mathbb{P}$-valued random variables with law μ. Then for any $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that for any $n \geq N$

$$\mathbb{E}W_1(\mu, \mu_n) < \epsilon,$$

where $\mu_n \in \mathcal{P}^1(\mathbb{P})$ is a random measure defined as $\mu_n := \sum_{i=1}^n \frac{1}{n}\delta_{Y_i}$.

Stochastic discrete-time flows: some technicalities 3.

II. Diminishing step-size resolvent iterates with varying measures close to a measure, remain close to each other:

Lemma (Lim-Pálfia 2019)

Let $\mu, \nu_i \in \mathcal{P}^1(\mathbb{P})$ for $i \in \mathbb{N}$, $l \geq 0$ an integer and $X_0, Y_0 \in \mathbb{P}$. Let

$$X_{k+1} := J^{\mu}_{1/(l+k+1)}(X_k) \quad \textit{and} \quad Y_{k+1} := J^{\nu_{k+1}}_{1/(l+k+1)}(Y_k).$$

Then

$$d_\infty(X_{k+1}, Y_{k+1}) \leq \frac{l+1}{k+l+1} d_\infty(X_0, Y_0) + \frac{k-l}{k+l+1} \sum_{i=l+1}^{k+1} \frac{W_1(\mu, \nu_i)}{k-l}.$$

These are the main ingredients of the proof and a lot of technical nonsense :) □

- Y. Lim and M. Pálfia, *Strong law of large numbers for the L^1-Karcher mean*, submitted (2019), available upon request, 30 pages.
- Y. Lim and M. Pálfia, *Existence and uniqueness of the L^1-Karcher mean*, revised in Adv. Math. (2017), https://arxiv.org/abs/1703.04292, 36 pages.
- S. Ohta and M. Pálfia, *Gradient flows and a Trotter-Kato formula of semi-convex functions on CAT(1)-spaces*, Amer. J. Math. 139:4 (2017), pp. 937–965.
- M. Pálfia, *Operator means of probability measures and generalized Karcher equations*, Adv. Math. 289 (2016), pp. 951–1007.
- K.-T. Sturm, *Nonlinear martingale theory for processes with values in metric spaces of nonpositive curvature*, Ann. Prob. **30** (2002), pp. 1195–1222.