Advances in Mathematics 457 (2024) 109933

Contents lists available at ScienceDirect

Advances in Mathematics

journal homepage: www.elsevier.com/locate/aim

Strong law of large numbers for generalized operator means

Zoltán Léka [a], Miklós Pálfia [b,c,*]

[a] *Institute of Applied Mathematics, Óbuda University, John von Neumann Faculty of Informatics, 1034 Budapest Bécsi út 96, Hungary*
[b] *Department of Mathematics, Corvinus University of Budapest, 1093 Budapest Fővám tér 8, Hungary*
[c] *Bolyai Institute, Interdisciplinary Excellence Centre, University of Szeged, H-6720 Szeged, Hungary*

A R T I C L E I N F O

Article history:
Received 17 May 2024
Received in revised form 26 August 2024
Accepted 30 August 2024
Available online xxxx
Communicated by Ilijas Farah

MSC:
primary 47A56, 47A64, 60F15
secondary 58B20

Keywords:
ODE
Thompson metric
Karcher means
Sturm's law of large numbers

A B S T R A C T

Sturm's strong law of large numbers in CAT(0) spaces and in the Thompson metric space of positive invertible operators is not only an important theoretical generalization of the classical strong law but also serves as a root-finding algorithm in the spirit of a proximal point method with splitting. It provides an easily computable stochastic approximation based on inductive means. The purpose of this paper is to extend Sturm's strong law and its deterministic counterpart, known as the "nodice" version, to unique solutions of nonlinear operator equations that generate exponentially contracting ODE flows in the Thompson metric. This includes a broad family of so-called generalized (Karcher) operator means introduced by Pálfia in 2016. The setting of the paper also covers the framework of order-preserving flows on Thompson metric spaces, as investigated by Gaubert and Qu in 2014, and provides a generally applicable resolvent theory for this setting.

* Corresponding author.
E-mail addresses: leka.zoltan@uni-obuda.hu (Z. Léka), miklos.palfia@uni-corvinus.hu (M. Pálfia).

https://doi.org/10.1016/j.aim.2024.109933

1. Introduction

Let $\mathbb{S}$ denote the vector space of self-adjoint operators equipped with the operator norm $\|\cdot\|$ on a Hilbert space $\mathcal{H}$ and let $\mathbb{P} \subset \mathbb{S}$ denote the cone of invertible positive definite operators. On $\mathbb{S}$ the closure $\overline{\mathbb{P}}$ of the cone generates the positive definite partial order $\leq$ also called the Loewner order. Means of elements of $\mathbb{P}$ and related ergodic theorems were studied in a large number of papers, see for example [2,5,6,16,21,19,20,25,26]. Arguably the natural metric on $\mathbb{P}$ is

$$d_\infty(a,b) := \|\log(a^{-1/2}ba^{-1/2})\| = \mathrm{spr}\{\log(a^{-1}b)\} \tag{1}$$

called the Thompson metric, which turns $(\mathbb{P}, d_\infty)$ into a complete metric space such that the topology generated by d_∞ agrees with the relative operator norm topology [33], where $\mathrm{spr}(X)$ denotes the spectral radius of X. Sturm's law of large numbers [32] states the almost sure convergence proved in [23] of the stochastic inductive mean sequence $\{S_n\}_{n\in\mathbb{N}}$ to $\Lambda(\mu)$, the unique solution of the *Karcher equation*

$$\int_{\mathbb{P}} \log_x a \, d\mu(A) = 0,$$

with $\log_x a := x^{-1/2}\log(x^{1/2}ax^{-1/2})x^{1/2}$. Here S_n is defined recursively as $S_1 := Y_1$,

$$S_{n+1} := S_n \#_{\frac{1}{n+1}} Y_{n+1} \tag{2}$$

for i.i.d. random variables $\{Y_n\}_{n\in\mathbb{N}}$ with L^1-law μ. The function $t \in [0,1] \mapsto a\#_t b$ for $a, b \in \mathbb{P}$ is defined by

$$a\#_t b = \Lambda((1-t)\delta_a + t\delta_b) = a^{1/2}\left(a^{-1/2}ba^{-1/2}\right)^t a^{1/2} = a\left(a^{-1}b\right)^t,$$

it is the weighted geometric mean of positive operators $a, b \in \mathbb{P}$, which is monotone [5] with respect to the partial order $\leq$ generated by the cone $\mathbb{P}$. It is also known [24] that the geometric mean Λ is 1-Lipschitz with respect to the L^1-Wasserstein distance W_1. A deterministic, also called "nodice", version of Sturm's law that periodically recycles all the points a_i was proved by Holbrook [15] for positive matrices and then in [23] for the operator case. This "nodice" theorem states that S_n converges to $\Lambda(\sum_{i=0}^{k-1}\frac{1}{k}\delta_{a_i})$ for the recycling deterministic version $Y_n := a_{\overline{n}}$ in (2), where $\overline{n}$ denotes the residual of n modulo k. The proofs in [23] develop an ODE theory for the initial value problem

$$\dot{\gamma}(t) = \int_{\mathbb{P}} \log_{\gamma(t)} a \, d\mu(a)$$

and use resolvent iterates in the spirit of (2) to approximate ODE curves converging to the unique stationary point of the flow which is $\Lambda(\mu)$. It is rather natural to study more general operator equations of similar form

$$\int_{\mathbb{P}} f_x(a)\, d\mu(a) = 0, \tag{3}$$

where $f_x(a) := x^{1/2} f(x^{-1/2} a x^{-1/2}) x^{1/2}$ for $f : (0, \infty) \mapsto \mathbb{R}, f(1) = 0, f'(1) = 1$ an operator monotone function and this was initiated in [29] showing various properties and uniqueness of solution of (3) for probability measures with bounded support.

Our goal is to study the properties of solutions $\Lambda_\varphi(\mu)$ of an even more general class of operator equations

$$\int_{\mathbb{P}} \varphi(x, y)\, d\mu(y) = 0 \tag{4}$$

such that the associated initial value problem

$$\dot{\gamma}(t) = \int_{\mathbb{P}} \varphi(\gamma(t), y)\, d\mu(y) \tag{5}$$

preserves the cone $\mathbb{P}$ and has an *exponential contraction rate* $\alpha \in \mathbb{R}$; that is

$$d_\infty(x(t), y(t)) \leq e^{-\alpha t} d_\infty(x_0, y_0)$$

for any $t \geq 0$ where $x_0, y_0 \in \mathbb{P}$ and $x(t), y(t)$ denote the solutions to (5) with initial values $x(0) := x_0$, $y(0) := y_0$ respectively. In the special case when the generated flow of the Cauchy initial value problem

$$\begin{aligned} \dot{x} &= \psi(x) \\ x(0) &= x_0 \end{aligned}$$

turns out to be (Loewner) order preserving, then Gaubert and Qu in [13] were able to determine the best exponential contraction rate by the simple formula

$$\alpha(\psi) = \sup\{\, \alpha \in \mathbb{R} : D\psi(x)(x) - \psi(x) \leq -\alpha x \text{ for all } x \in \mathbb{P} \}. \tag{6}$$

Using only the exponential contraction rate property combined with a newly defined *resolvent map* $J_\lambda^\mu(z)$ for $\lambda > 0$ that we introduce as the solution to the equation

$$\lambda \int_{\mathbb{P}} \varphi(x, y)\, d\mu(y) + \log_x z = 0$$

suited for this setting, we prove *Sturm-type of strong law of large numbers* and *"nodice"* theorems approximating unique solutions of (4) when the exponential contraction rate with $\alpha > 0$ and the L^1-integrability condition

$$\int_{\mathbb{P}} \|y^{-1/2}\varphi(a,y)y^{-1/2}\| \, d\mu(a) < \infty$$

for all $y \in \mathbb{P}$ are satisfied. This reproves and generalizes results in [29] to the unbounded probability measure setting, establishes a crucial *Wasserstein-type continuity* estimate

$$d_\infty(\Lambda_\varphi(\mu), \Lambda_\varphi(\nu)) \leq \frac{1}{\alpha} \left\| \int_{\mathbb{P}} \Lambda_\varphi(\nu)^{-1/2}\varphi(a, \Lambda_\varphi(\nu))\Lambda_\varphi(\nu)^{-1/2} \, d(\mu - \nu)(a) \right\|$$

generalizing the one in [23,24] proved for the particular case Λ when $\alpha = 1$ and $\varphi(x, y) = \log_x(y)$. This framework applies to the various generalized operator means spawned by (3) which turn out to have *order preserving flows* associated with them through (5) with an *exponential contraction rate* $\alpha > 0$; and also establishes a (nonlinear) strong law of large numbers for them through an appropriate generalization of (2). This family includes the geometric (also called Karcher) mean Λ, the arithmetic and the harmonic means and the matrix power means [25,19,20] to name a few. To achieve this, we also provide a powerful enough resolvent theory for the setting of Gaubert and Qu in [13], so that further studies can be made about ODEs and differential inclusion problems discussed there. Among others, we establish the resolvent estimate

$$d_\infty(J^\mu_\lambda(x), J^\nu_\lambda(y)) \leq \frac{1}{1+\alpha\lambda} d_\infty(x, y) + \frac{\lambda K}{1+\alpha\lambda} W_1(\mu, \nu)$$

where $K > 0$ is a suitable Lipschitz constant for $y \mapsto y^{-1/2}\varphi(a, y)y^{-1/2}$. This, for example readily implies the contractivity of the resolvent when $\alpha > 0$ and other estimates so that one can readily develop a nonlinear Crandall-Liggett theory such as in [17].

2. Preliminaries

Let $\mathbb{S}$ denote the bounded self-adjoint elements and $\mathbb{P}$ denote the cone of positive invertible elements of $\mathcal{B}(\mathcal{H})$ on a Hilbert space $\mathcal{H}$. The Thompson part metric [33] on $\mathbb{P}$ is defined by

$$d_\infty(x, y) = \|\log(x^{-1/2}yx^{-1/2})\|.$$

According to the spectral mapping theorem,

$$\|\log(x^{-1/2}yx^{-1/2})\| = \operatorname{spr}(\log(x^{-1/2}yx^{-1/2})) = \log(\operatorname{spr}(x^{-1/2}yx^{-1/2})) = \log(\operatorname{spr}(x^{-1}y)),$$

where spr stands for the spectral radius. Therefore, for any invertible $z \in \mathcal{B}(\mathcal{H})$, we have the conjugance invariance property

$$d_\infty(z^{-1}xz^{-1}, z^{-1}yz^{-1}) = d_\infty(x, y)$$

of the Thompson metric. The *Exponential Metric Increasing* (EMI) property [5,22] says that for any $x, y \in \mathbb{P}$ the inequality

$$\|\log x - \log y\| \leq d_\infty(x, y)$$

holds.

Let us recall that the Lipschitz constant of a $\mathcal{B}(\mathcal{H})$-valued function $f : \mathbb{P} \to \mathcal{B}(\mathcal{H})$ is defined by

$$\mathrm{Lip}_\infty(f) = \sup_{x \neq y} \|f(x) - f(y)\|/d_\infty(x, y).$$

The set of fully supported Borel probability measures over $(\mathbb{P}, d_\infty)$ is denoted by $\mathcal{P}(\mathbb{P})$, where fully supported means that for $\mu \in \mathcal{P}(\mathbb{P})$ we have $\mu(\mathrm{supp}(\mu)) = 1$. Note that $\mathrm{supp}(\mu)$ is separable, thus all integrations with respect to such $d\mu$ can be restricted to the separable metric space $(\overline{\mathrm{supp}(\mu)}, d_\infty)$. This essentially rules out any pathological probability measures, even though the whole space $(\mathbb{P}, d_\infty)$ is not separable. This also means that it is enough to consider Bochner integrals of vector-valued functions with respect to elements of $\mathcal{P}(\mathbb{P})$. This approach is now the standard when dealing with probability measures over $\mathbb{P}$, see for instance [18,23] and others.

The Kantorovich–Rubinstein duality theorem on a separable metric space (S, d) ([1] and [10, Theorem 11.8.2]) states that for any two Borel probability measures μ and ν,

$$W_1(\mu, \nu) = \sup \left\{ \left| \int_S f\, d\mu - \int_S f\, d\nu \right| : f \text{ real-valued and } \mathrm{Lip}_\infty(f) \leq 1 \right\}.$$

Let $\varphi : \mathbb{P} \mapsto \mathbb{S}$ be a local Lipschitz function, so that the flow generated by the ODE $\dot{x} = \varphi(x)$ is Lipschitz. We can characterize the exponential contraction rate of this flow as follows:

Theorem 2.1. *Let $x_0, y_0 \in \mathbb{P}$ and let $x(t), y(t)$ denote the solutions to the two Cauchy initial value problems*

$$\dot{x} = \varphi(x)$$
$$x(0) = x_0$$

and

$$\dot{y} = \varphi(y)$$
$$y(0) = y_0.$$

Then we have the exponential contraction rate α *for* φ*; that is*

$$d_\infty(x(t), y(t)) \leq e^{-\alpha t} d_\infty(x_0, y_0)$$

for any $t \geq 0$*, if and only if*

$$j(D\log(x_0^{-1/2}y_0x_0^{-1/2})(D(x_0^{-1/2}y_0x_0^{-1/2})(\dot{x}(0), \dot{y}(0)))) \leq -\alpha d_\infty(x_0, y_0) \tag{7}$$

where j *is a norming linear functional for* $\log(x_0^{-1/2}y_0x_0^{-1/2})$.

Proof. "$\Rightarrow$": In order to differentiate the inequality $d_\infty(x(t), y(t)) \leq e^{-\alpha t} d_\infty(x_0, y_0)$ from the right at $t = 0$ where it holds with equality, we apply the multivariate chain rule to the composition $j(\log(x(t)^{-1/2}y(t)x(t)^{-1/2}))$. Let j be a norm-attaining linear functional such that

$$\|\log(x_0^{-1/2}y_0x_0^{-1/2})\| = j(\log(x_0^{-1/2}y_0x_0^{-1/2}))$$

and $\|j\| = 1$. Now we shall differentiate both sides of the inequality

$$j(\log(x(t)^{-1/2}y(t)x(t)^{-1/2})) \leq d_\infty(x(t), y(t)) \leq d_\infty(x_0, y_0)e^{-\alpha t}$$

at $t = 0$, since equality occurs when $t = 0$. Therefore, from the chain rule we obtain (7).

"$\Leftarrow$": To reverse the above argument, we use (7) pointwise for each $t > 0$ along the curves $x(t), y(t)$, that is

$$j(D\log(x^{-1/2}yx^{-1/2})(D(x^{-1/2}yx^{-1/2})(\dot{x}(t), \dot{y}(t))))\Big|_{(x,y)=(x(t),y(t))} \leq -\alpha d_\infty(x(t), y(t))$$

where j is a norming linear functional for $\log(x(t)^{-1/2}y(t)x(t)^{-1/2})$. Notice that the function $t \mapsto d_\infty(x(t), y(t))$ is Lipschitz on bounded time intervals since the solution curves of the Cauchy problem are Lipschitz themselves. Thus $t \mapsto d_\infty(x(t), y(t))$ is differentiable for almost every $t > 0$ with integrable absolutely continuous differential. This implies that

$$\lim_{\varepsilon \to 0+} \frac{d_\infty(x(t+\varepsilon), y(t+\varepsilon)) - d_\infty(x(t), y(t))}{\varepsilon} = j(D\log(x^{-1/2}yx^{-1/2})(D(x^{-1/2}yx^{-1/2})(\dot{x}(t), \dot{y}(t))))\Big|_{(x,y)=(x(t),y(t))}$$

for almost every $t > 0$ by Clarke's chain rule [7, Theorem 2.9.9], because $d_\infty(x(t), y(t))$ is a composition of a convex and Lipschitz-continuous thus regular (in the sense of 2.3.4

in [7]) function $\|\cdot\|$ and a Frechét-differentiable function $t \to \log(x(t)^{-1/2}y(t)x(t)^{-1/2})$. This further implies

$$\lim_{\varepsilon\to 0+} \frac{d_\infty(x(t+\varepsilon), y(t+\varepsilon)) - d_\infty(x(t), y(t))}{\varepsilon} \leq -\alpha d_\infty(x(t), y(t))$$

for almost every $t > 0$. Using Gronwall's lemma we integrate this with respect to $t \geq 0$ to get $d_\infty(x(t), y(t)) \leq e^{-\alpha t} d_\infty(x_0, y_0)$ as wanted. □

Corollary 2.2. *If $\alpha(\varphi) \in \mathbb{R}$ is such that $d_\infty(x(t), y(t)) \leq e^{-\alpha(\varphi)t} d_\infty(x_0, y_0)$ for any solution curves $x(t), y(t)$ as in Theorem 2.1, then $\alpha(\cdot)$ is superadditive, that is*

$$\alpha(\varphi_1 + \varphi_2) \geq \alpha(\varphi_1) + \alpha(\varphi_2). \tag{8}$$

Proof. We add up (7) for φ_1 and φ_2 to notice that on the left-hand side it is additive in $(\dot{x}(0), \dot{y}(0))$ and on the right-hand side $\alpha(\cdot)$ is also additive and use the "$\Leftarrow$" implication of Theorem 2.1. □

Under exponential contraction rate a flow is always extendable if the boundary of the set containing the dynamics is infinitely far. This, although might be known, is included in the following assertion with proof for the sake of completeness.

Proposition 2.3. *Let $\alpha \in \mathbb{R}$ be given. Let $x_0, y_0 \in \mathbb{P}$ and for $0 \leq t \leq T_{x_0,y_0}$ let $x(t), y(t)$ denote the solutions to the two Cauchy initial value problems*

$$\dot{x} = \varphi(x)$$
$$x(0) = x_0$$

and

$$\dot{y} = \varphi(y)$$
$$y(0) = y_0,$$

where the existence time T_{x_0,y_0} may depend on the initial conditions x_0, y_0. Suppose further we have the exponential contraction rate

$$d_\infty(x(t), y(t)) \leq e^{-\alpha t} d_\infty(x_0, y_0)$$

for any $0 \leq t \leq T_{x_0,y_0}$. Then $x(t)$ exists for all $t \in [0, \infty)$ independently of $x_0 \in \mathbb{P}$.

Proof. For $x_0 \in \mathbb{P}$, pick a solution $x(t) \in \mathbb{P}$. Assume that the maximal interval of its existence is finite $[0, T_{\max})$. We show that $\lim_{t\to T_{\max}-} x(t)$ does exist in d_∞. For any $0 \leq t_1 < t_2 < T_{\max}$ we have

$$d_\infty(x(t_1), x(t_2)) = d_\infty(x(t_1), x(t_1 + (t_2 - t_1))) \leq e^{-\alpha t_1} d_\infty(x(0), x(t_2 - t_1)).$$

Hence, if $t_1 < t_2$ and both approach $T_{\max}$ from the left, we obtain a continuous extension of $x(t)$ at $t = T_{\max}$. Note that the boundary of $\mathbb{P}$ in d_∞ is infinitely far from any interior point, thus the continuous extension $x(t)$ is still in $\mathbb{P}$ and thus it must be C^1. By the Cauchy-Lipschitz theorem, there exists a unique solution to the initial value problem on the extended interval $[T_{\max} - \varepsilon, T_{\max} + \varepsilon]$. This implies that $T_{\max} = \infty$ must hold. □

The solutions of the general ordinary differential equation $\dot{x} = \varphi(x)$ form an *order-preserving (or monotone) flow* on the standard positive cone $\mathbb{P}$ if any two solutions x_1 and x_2 satisfy $x_1(t) \le x_2(t)$ for all $t \ge 0$ whenever $x_1(0) \le x_2(0)$ holds on $\mathbb{P}$. According to Redheffer and Walter [31, Theorem 3], and Gaubert and Qu [13, Proposition 3.3], the flow generated by φ is order preserving on $\mathbb{P}$ if and only if for any positive linear functional ω and $0 \le v \in \mathbb{S}$ such that $\omega(v) = 0$ the inequality

$$\omega(D\varphi(x)(v)) \ge 0 \tag{9}$$

follows for any $x \in \mathbb{P}$. Moreover, the least exponential contraction rate of general order-preserving flows on positive cones has been characterized in [13]. The best contraction rate is given by [13, Theorem 3.5.]:

$$\alpha(\varphi) = \sup\{\, \alpha \in \mathbb{R} : D\varphi(x)(x) - \varphi(x) \le -\alpha x \text{ for all } x \in \mathbb{P}\}.$$

This formula implies (8) as well. We also note that

$$x^{-1/2} D\varphi(x)(x) x^{-1/2} - x^{-1/2}\varphi(x)x^{-1/2} = D(x^{-1/2}\varphi(x)x^{-1/2})(x).$$

Hence the constant $\alpha(\varphi)$ can be defined as

$$\alpha(\varphi) = \sup\{\, \alpha \in \mathbb{R} : D(x^{-1/2}\varphi(x)x^{-1/2})(x) \le -\alpha \text{ for all } x \in \mathbb{P}\}. \tag{10}$$

More generally, if one needs the least exponential contraction rate of general order-preserving flows on open subsets U of a positive cone, then [13, Remark 3.6.] gives an answer defining first $\lambda_0 := \sup_{x,y \in U} e^{d_\infty(x,y)}$ and then

$$\alpha(U, \varphi) = \sup\{\, \alpha \in \mathbb{R} : D\varphi(x)(x) - \varphi(x) \le -\alpha x \text{ for all } x \in \lambda U, \lambda_0^{-1} \le \lambda \le 1\}$$

and similarly

$$\alpha(U, \varphi) = \sup\{\, \alpha \in \mathbb{R} : D(x^{-1/2}\varphi(x)x^{-1/2})(x) \le -\alpha, \forall x \in \lambda U, \lambda_0^{-1} \le \lambda \le 1\}. \tag{11}$$

Let us recall that the logarithm function $x \mapsto \log x$ has the particular directional derivative $D\log(x)(x) = I$ for any $x \in \mathbb{P}$. This is established through the well-known integral representation formula [4]

$$\log x = \int_0^\infty \frac{s}{1+s^2} - \frac{1}{x+s}\, ds$$

for any $x > 0$. Therefore, we have the Fréchet derivative

$$\frac{d}{dt}\log(x+ty)|_{t=0} = D\log(x)(y) = \int_0^\infty (x+s)^{-1}y(x+s)^{-1}\, ds \tag{12}$$

and $D\log(x)(x) = I$ follows by a simple computation. This integral formula also implies that the map $y \mapsto D\log(x)(y)$ is a positive linear map on $\mathcal{B}(\mathcal{H})$.

Now we can state the following theorem.

Theorem 2.4. *Let us assume that the first-order system $\dot{x} = \varphi(x)$ generates a contractive flow on $\mathbb{P}$ and let x_* denote a fixed point of the system, and let*

$$d_\infty(x_*, x(t)) \le e^{-\alpha t} d_\infty(x_*, x(0)),$$

for any $t \ge 0$ and $x(0) \in \mathbb{P}$. Then, for any $y \in \mathbb{P}$, there exists a linear functional $\omega_{x_*,y} \in \mathbb{S}^*$ such that $\|\omega_{x_*,y}\| = 1$ and

$$d_\infty(x_*, y) \le \frac{1}{\alpha}\omega_{x_*,y}(y^{-1/2}\varphi(y)y^{-1/2}) \le \frac{1}{\alpha}\|y^{-1/2}\varphi(y)y^{-1/2}\|. \tag{13}$$

Proof. Set the initial condition $x(0) = y$, and denote a corresponding solution curve by $\chi(t)$. Let j be a norm-attaining linear functional such that

$$\|\log(x_*^{-1/2}\chi(0)x_*^{-1/2})\| = j(\log(x_*^{-1/2}yx_*^{-1/2}))$$

and $\|j\| = 1$. Now we shall differentiate the inequality

$$j(\log(x_*^{-1/2}\chi(t)x_*^{-1/2}) \le d_\infty(x_*, \chi(t)) \le d_\infty(x_*, y)e^{-\alpha t}.$$

at $t = 0$. Notice that equality occurs when $t = 0$. Therefore, from the chain rule we obtain

$$j(D\log(x_*^{-1/2}\chi(0)x_*^{-1/2})(x_*^{-1/2}\dot{\chi}(0)x_*^{-1/2})) \le -\alpha d_\infty(x_*, y). \tag{14}$$

Moreover, the map

$$\Phi\colon u \mapsto D\log(x_*^{-1/2}yx_*^{-1/2})(x_*^{-1/2}y^{1/2}uy^{1/2}x_*^{-1/2})$$

is positive by (12), $\Phi(1) = 1$ and thus $\|\Phi\| = 1$ (see [30, Corollary 2.9.]). Clearly,

$$D\log(x_*^{-1/2}\chi(0)x_*^{-1/2})(x_*^{-1/2}\dot{\chi}(0)x_*^{-1/2}) = \Phi(y^{-1/2}\varphi(y)y^{-1/2}).$$

Hence the linear functional $\omega_{x_*,y} := -j \circ \Phi$ is of norm 1. Moreover,

$$d_\infty(x_*, y) \leq \frac{1}{\alpha}\omega_{x_*,y}(y^{-1/2}\varphi(y)y^{-1/2}) \leq \frac{1}{\alpha}\|y^{-1/2}\varphi(y)y^{-1/2}\|$$

holds. □

Here is our next proposition.

Proposition 2.5. *Let $z \in \mathbb{P}$. The first-order system*

$$\dot{x}(t) = x(t)^{1/2} \log(x(t)^{-1/2} z x(t)^{-1/2}) x(t)^{1/2}$$

generates an exponentially contractive monotone flow on the positive cone $\mathbb{P}$ such that

$$d_\infty(z, x(t)) \leq e^{-t} d_\infty(z, x(0))$$

for any $t \geq 0$ and $x(0) \in \mathbb{P}$.

Proof. Consider the function $\varphi(x) = x^{1/2}\log(x^{-1/2}zx^{-1/2})x^{1/2}$, where $x \in \mathbb{P}$. This function is locally Lipschitz; thus, on any bounded metric ball around z, it generates a unique flow. Consequently, the flow is also unique on the whole $\mathbb{P}$. First, we prove its order-preserving property. We recall the integral formula

$$\log x = \int_0^\infty \frac{s}{1+s^2} - \frac{1}{x+s}\, ds$$

which holds for any $x > 0$.

Hence,

$$\varphi(x) = \int_0^\infty xs(1+s^2)^{-1} - x(z+sx)^{-1}x\, ds = \int_0^\infty l_s(x)\, ds,$$

where $l_s(x) = xs(1+s^2)^{-1} - x(z+sx)^{-1}x$. Now, we verify that

$$\omega(Dl_s(x)(v)) \geq 0$$

holds for any state ω and $0 \leq v \in \mathbb{S}$ such that $\omega(v) = 0$. A simple computation shows that

$$Dl_s(x)(v) = vs(1+s^2)^{-1} - v(z+sx)^{-1}x + x(z+sx)^{-1}sv(z+sx)^{-1}x - x(z+sx)^{-1}v.$$

From the Cauchy–Schwartz inequality for positive functionals (see e.g. [3, p. 28.]), we get

$$|\omega(vr)| = |\omega(v^{1/2}v^{1/2}r)| \leq \omega(v)^{1/2}\omega(r^*vr)^{1/2} = 0$$

for any $r \in \mathcal{B}(\mathcal{H})$; that is

$$\omega(vr) = \omega(rv) = 0.$$

Therefore,

$$\omega(Dl_s(x)(v)) = \omega(x(z+sx)^{-1}sv(z+sx)^{-1}x) \geq 0$$

which clearly implies $\omega(D\varphi(x)(v)) \geq 0$ as well. Then the generated flow is order-preserving from (9).

Next, we shall apply the Gaubert–Qu condition (10) to determine the least contraction rate of the flow. An easy calculation reveals

$$\begin{aligned} D(x^{-1/2}\varphi(x)x^{-1/2})(x) &= D(\log(x^{-1/2}zx^{-1/2}))(x) \\ &= D\log(x^{-1/2}zx^{-1/2})(D(x^{-1/2}zx^{-1/2})(x)) \\ &= -D\log(x^{-1/2}zx^{-1/2})(x^{-1/2}zx^{-1/2}) \\ &= -1, \end{aligned}$$

from the result on the directional derivative (12). This finishes the proof. □

Corollary 2.6 *(cf. Corollary 2.3. [11]). Any operator monotone function $f\colon (0,\infty) \to \mathbb{R}$ is operator concave and the non-commutative perspectives*

$$f_x(z) := x^{1/2}f(x^{-1/2}zx^{-1/2})x^{1/2} \tag{15}$$

are jointly concave.

Proposition 2.7. *Under the assumptions of the previous Corollary 2.6, the map*

$$x \mapsto \int_{\mathbb{P}} f_x(z)\, d\mu(z)$$

where $\mu \in \mathcal{P}(\mathbb{P})$ such that $\int_{\mathbb{P}} \|f_y(z)\|\, d\mu(z) < \infty$ for any $y \in \mathbb{P}$, is also concave in $x \in \mathbb{P}$ and locally Lipschitz with respect to $\|\cdot\|$.

Proof. Firstly when $\mathrm{supp}(\mu)$ is bounded $x \mapsto \int_{\mathbb{P}} f_x(z)\, d\mu(z)$ is easily seen to be continuous by the Banach space version of Lebesgue's dominated convergence theorem. Then given a general $\mu \in \mathcal{P}(\mathbb{P})$ satisfying $\int_{\mathbb{P}} \|f_y(z)\|\, d\mu(z) < \infty$ for each $y \in \mathbb{P}$, we can approximate it weakly by a sequence $\mu_n \in \mathcal{P}(\mathbb{P})$ such that $\mathrm{supp}(\mu_n)$ are all bounded and also

$$\int_{\mathbb{P}} f_x(z)\, d\mu_n(z) \to \int_{\mathbb{P}} f_x(z)\, d\mu(z)$$

pointwisely for each $x \in \mathbb{P}$. Then $\left\{\int_{\mathbb{P}} f_x(z)\, d\mu_n(z)\right\}_{n\in\mathbb{N}}$ is a pointwisely convergent family of concave continuous functions, thus by [8, Theorem 7.1.] $\left\{\int_{\mathbb{P}} f_x(z)\, d\mu_n(z)\right\}_{n\in\mathbb{N}}$ is a uniformly locally Lipschitz family. Then it follows that $\int_{\mathbb{P}} f_x(z)\, d\mu(z)$ is also locally Lipschitz and concave. □

Proposition 2.8. *For any operator monotone function $f\colon (0,\infty) \to \mathbb{R}$, the perspectives $x \mapsto f_x(z)$ and their integrals*

$$x \mapsto \int_{\mathbb{P}} x^{1/2} f(x^{-1/2} z x^{-1/2}) x^{1/2}\, d\mu(z),$$

where $\mu \in \mathcal{P}(\mathbb{P})$ such that $\int_{\mathbb{P}} \|f_y(z)\|\, d\mu(z) < \infty$ for any $y \in \mathbb{P}$, define monotone flows on $\mathbb{P}$ given by the first-order system

$$\dot{x}(t) = \int_{\mathbb{P}} x(t)^{1/2} f(x(t)^{-1/2} z x(t)^{-1/2}) x(t)^{1/2}\, d\mu(z).$$

Proof. The local Lipschitz property of the previous Proposition 2.7 implies the existence and uniqueness of a solution to the ODE. Then the remaining part of the proof is essentially the same as that of the previous Proposition 2.5 and relies on the celebrated integral representation formula for operator monotone functions

$$f(x) = a + bx + \int_0^\infty \frac{s}{1+s^2} - \frac{1}{x+s}\, d\nu(s), \tag{16}$$

where a is real, b is non-negative and ν is a unique positive measure on $(0,\infty)$ such that

$$\int_0^\infty \frac{1}{s^2+1} d\nu(s) < \infty,$$

see [4, p. 144]. □

3. Generalized resolvent estimates

Let us define the family of functions

$$\mathcal{C} := \{\varphi\colon \mathbb{P} \to \mathbb{S} \text{ local Lipschitz and the flow of } \dot{x} = \varphi(x) \text{ has exponential contraction rate } \alpha \in \mathbb{R}_+\}.$$

Then by Proposition 2.3 the flow of any $\varphi \in \mathcal{C}$ exists for all time $t \in [0, \infty)$ and thus leaves $\mathbb{P}$ invariant. It is then simple to check that $\mathcal{C}$ is a cone. Indeed, if $\varphi_1, \varphi_2 \in \mathcal{C}$ generate exponentially contracting flows, then so do their sum $\varphi_1 + \varphi_2$ and $\lambda\varphi_1$ for a positive λ by Corollary 2.2. We have the following existence and uniqueness result.

Proposition 3.1. *Let $\varphi \in \mathcal{C}$ with $\alpha(\varphi) > 0$. Then there exists a unique solution x_0 to the equation*

$$\varphi(x) = 0$$

over the domain $\mathbb{P}$ and it is the only fixed point of the ODE $\dot{x} = \varphi(x)$.

Proof. For any $\varphi \in \mathcal{C}$, the solutions of the ODE $\dot{x} = \varphi(x)$ are exponentially strictly contractive in d_∞, thus we have at most one stationary point which solves the fixed point equation. Indeed, let $\chi_{x_0}(t)$ denote the solution of the ODE with the initial value condition $\chi_{x_0}(0) = x_0$. Then the flow map $x_0 \mapsto \chi_{x_0}(1)$, $x_0 \in \mathbb{P}$, is a strict contraction in the metric space $(\mathbb{P}, d_\infty)$. Hence an application of the Banach fixed point theorem gives the existence of x_* such that $\chi_{x_*}(1) = x_*$. It is simple to see that x_* belongs to a periodic orbit of the flow with period $\Delta t = 1$. Let x_1 be any point of this periodic orbit. Then the function $t \mapsto d_\infty(\chi_{x_1}(t), \chi_{x_*}(t))$ is periodic as well. On the other hand, since $d_\infty(\chi_{x_1}(t), \chi_{x_*}(t)) \leq e^{-\alpha t} d_\infty(x_1, x_*)$, we get that $d_\infty(\chi_{x_1}(t), \chi_{x_*}(t)) \to 0$. Thus we obtain that $x_1 = x_*$ and $\chi_{x_*}(t) = x_*$ for any $t \geq 0$; that is, x_* is a (unique) stationary point. □

For simplicity and later use, let us introduce the notation

$$\log_x z = x^{1/2} \log(x^{-1/2} z x^{-1/2}) x^{1/2}.$$

For any $\varphi \in \mathcal{C}$, it is easy to see that the function $\psi_z \colon x \mapsto \lambda\varphi(x) + \log_x z$ belongs to the cone $\mathcal{C}$ for any $z \in \mathbb{P}$ and $\lambda > 0$. In fact, from Proposition 2.5, we have that $\alpha(\log_{(\cdot)} z) = 1$ and then $\alpha(\psi_z) \geq \lambda\alpha(\varphi) + 1$. Relying on this and the previous Proposition 3.1, for any fixed z and $\lambda > 0$, the *resolvent equation*

$$\lambda\varphi(x) + \log_x z = 0$$

has a unique solution $x =: J_\lambda^\varphi(z)$ in $\mathbb{P}$, because the flow generated by the ODE $\dot{x} = \lambda\varphi(x) + \log_x z$ is exponentially contractive in $\mathbb{P}$, and thus possesses a unique fixed point denoted by $J_\lambda^\varphi(z)$.

Definition 3.2 *(Generalized resolvent).* For any $\varphi \in \mathcal{C}$ and $\lambda > 0$, the unique solution $J_\lambda^\varphi(z)$ of

$$\lambda\varphi(x) + \log_x z = 0$$

in $x \in \mathbb{P}$ is called the generalized resolvent at $z \in \mathbb{P}$ associated with φ.

In the sequel, we shall prove several inequalities related to the generalized resolvent. Our first statement establishes its contractive property.

Proposition 3.3. *Let $\varphi \in \mathcal{C}$. For any $\lambda > 0$,*

$$d_\infty(J_\lambda^\varphi(x), J_\lambda^\varphi(y)) \leq \frac{1}{1+\alpha\lambda} d_\infty(x,y)$$

holds.

Proof. Fix $x \in \mathbb{P}$. Let us consider the function $\psi_x \colon z \mapsto \lambda\varphi(z) + \log_z x$ defined on $\mathbb{P}$. We write $z = J_\lambda^\varphi(x)$ to the (unique) solution to the resolvent equation

$$\psi_x(z) = \lambda\varphi(z) + \log_z x = 0.$$

For any $y \in \mathbb{P}$, we recall that $J_\lambda^\varphi(y)$ satisfies its own resolvent equation; that is,

$$\lambda\varphi(J_\lambda^\varphi(y)) + \log_{J_\lambda^\varphi(y)} y = 0.$$

Thus,

$$\psi_x(J_\lambda^\varphi(y)) = \lambda\varphi(J_\lambda^\varphi(y)) + \log_{J_\lambda^\varphi(y)} y + \log_{J_\lambda^\varphi(y)} x - \log_{J_\lambda^\varphi(y)} y = \log_{J_\lambda^\varphi(y)} x - \log_{J_\lambda^\varphi(y)} y.$$

From Theorem 2.1 and Proposition 2.5, $\alpha(\psi_x) \geq 1 + \lambda\alpha(\varphi)$. Applying (13) to ψ_x, we have

$$\begin{aligned} d_\infty(J_\lambda^\varphi(x), J_\lambda^\varphi(y)) &\leq \frac{1}{1+\alpha\lambda} \|J_\lambda^\varphi(y)^{-1/2}\psi_x(J_\lambda^\varphi(y))J_\lambda^\varphi(y)^{-1/2})\| \\ &= \frac{1}{1+\alpha\lambda} \|\log(J_\lambda^\varphi(y)^{-1/2} x J_\lambda^\varphi(y)^{-1/2}) - \log(J_\lambda^\varphi(y)^{-1/2} y J_\lambda^\varphi(y)^{-1/2})\|. \end{aligned}$$

Now, the EMI inequality and the conjugate invariance of d_∞ complete the proof. □

Interestingly enough, one can apply the generalized resolvents to approximate the solutions of the ordinary differential equation $\dot{x} = \varphi(x)$. In the spirit of the non-linear Crandall–Liggett theory [9], this approach has essentially been developed in the metric space $(\mathbb{P}, d_\infty)$ to study the (generalized) Karcher equation [24], and has been extended by Lawson in [17]. Now, by Theorem 4.11 of [17], for any $x_0 \in \mathbb{P}$,

$$(J_{t/n}^\varphi)^n x_0 \to S(t)x_0 \qquad n \to \infty$$

in the Thompson metric for $t \geq 0$, where $(S(t))_{t\geq 0}$ is a non-linear operator semigroup.

Corollary 3.4. *Let $\varphi \in \mathcal{C}$. For any $x, y \in \mathbb{P}$,*

$$d_\infty(J_\lambda^\varphi(x), y) \le \frac{1}{1+\alpha\lambda} d_\infty(x, y) + \frac{\lambda}{1+\alpha\lambda} \|y^{-1/2}\varphi(y)y^{-1/2}\|.$$

Proof. We note that the inverse of J_λ^φ on $\mathbb{P}$ is given by $z = (J_\lambda^\varphi)^{-1}(y) = \exp_y(-\lambda\varphi(y))$. Thus there exists a unique $z = (J_\lambda^\varphi)^{-1}(y) \in \mathbb{P}$ such that $J_\lambda^\varphi(z) = y$; i.e.

$$\lambda\varphi(y) + \log_y z = 0$$

holds. From Proposition 3.3 and the previous equation

$$\begin{aligned} d_\infty(J_\lambda^\varphi(y), J_\lambda^\varphi(z)) = d_\infty(J_\lambda^\varphi(y), y) \le & \frac{1}{1+\alpha\lambda} d_\infty(y, z) \\ = & \frac{\lambda}{1+\alpha\lambda} \|y^{-1/2}\varphi(y)y^{-1/2}\|. \end{aligned}$$

Hence

$$\begin{aligned} d_\infty(J_\lambda^\varphi(x), y) &\le d_\infty(J_\lambda^\varphi(x), J_\lambda^\varphi(y)) + d_\infty(J_\lambda^\varphi(y), J_\lambda^\varphi(z)) \\ &\le \frac{1}{1+\alpha\lambda} d_\infty(x, y) + \frac{\lambda}{1+\alpha\lambda} \|y^{-1/2}\varphi(y)y^{-1/2}\|, \end{aligned}$$

which proves the assertion. □

We shall introduce another notion for the resolvent if ϕ has an integral representation

$$\phi(x) = \int_{\mathbb{P}} \varphi(x, y)\, d\mu(y)$$

with respect to a $\mu \in \mathcal{P}(\mathbb{P})$ and a measurable locally Lipschitz family $\varphi(\cdot, y) \in \mathcal{C}$ for any $y \in \mathbb{P}$ provided that

$$\int_{\mathbb{P}} \|x^{-1/2}\varphi(x, y)x^{-1/2}\| d\mu(y) < \infty$$

which assures the existence of the Bochner integral on the right-hand side above. For simplicity, let $J_\lambda^\mu(z) := J_\lambda^\phi(z)$ denote a solution to the equation

$$\lambda \int_{\mathbb{P}} \varphi(x, y)\, d\mu(y) + \log_x z = 0.$$

Definition 3.5. If the equation $\phi(x) = 0$ has a unique solution in $\mathbb{P}$, we call it a (generalized) barycenter and denote it by $\Lambda_\varphi(\mu)$.

In the remainder of this section, from this point onward, assume that ϕ generates a flow with an exponential contraction rate $\alpha > 0$. By Proposition 3.1, this ensures the existence and uniqueness of the zero of ϕ.

Proposition 3.6. *With the notations of Theorem 2.1*

$$d_\infty(\Lambda_\varphi(\mu), \Lambda_\varphi(\nu)) \leq \frac{1}{\alpha} \left\| \int_{\mathbb{P}} \Lambda_\varphi(\nu)^{-1/2} \varphi(\Lambda_\varphi(\nu), y) \Lambda_\varphi(\nu)^{-1/2} \, d(\mu - \nu)(y) \right\|.$$

Proof. Let us introduce the function

$$\psi^m(x) = \int_{\mathbb{P}} \varphi(x, y) \, dm(y)$$

for any $m \in \mathcal{P}(\mathbb{P})$. Clearly, $\psi^\nu(\Lambda_\varphi(\nu)) = 0$. We also note that

$$\psi^\mu(\Lambda_\varphi(\nu)) = \psi^\mu(\Lambda_\varphi(\nu)) - \psi^\nu(\Lambda_\varphi(\nu)) = \int_{\mathbb{P}} \varphi(\Lambda_\varphi(\nu), y) \, d(\mu - \nu)(y).$$

From inequality (13) applied to the function $x \mapsto \psi^\mu(x)$, we obtain that there is a functional ω of norm 1 such that

$$d_\infty(\Lambda_\varphi(\mu), \Lambda_\varphi(\nu)) \leq \frac{1}{\alpha} \int_{\mathbb{P}} \omega(\Lambda_\varphi(\nu)^{-1/2} \varphi(\Lambda_\varphi(\nu), y) \Lambda_\varphi(\nu)^{-1/2}) \, d(\mu - \nu)(y),$$

which completes the proof. □

Let us note that the previous proposition implies the Wasserstein contraction result (17) below for barycenters as shown in [24, Proposition 2.6]. To see this, consider the function

$$\varphi(x, y) = x^{1/2} \log(x^{-1/2} y x^{-1/2}) x^{1/2}.$$

The exponential contraction rate of the generated flow on the cone $\mathbb{P}$ is $\alpha = 1$, as indicated in Proposition 2.5. Moreover, there is a functional ω of norm 1 such that

$$\begin{aligned} &\left\| \int_{\mathbb{P}} \Lambda(\nu)^{-1/2} \varphi(\Lambda(\nu), y) \Lambda(\nu)^{-1/2}) \, d(\mu - \nu)(y) \right\| \\ &= \int_{\mathbb{P}} \omega(\Lambda(\nu)^{-1/2} \varphi(\Lambda(\nu), y) \Lambda(\nu)^{-1/2}) \, d(\mu - \nu)(y) \end{aligned}$$

$$= \int_{\mathbb{P}} \omega(\log(\Lambda(\nu)^{-1/2} y \Lambda(\nu)^{-1/2})\, d\mu(y) - \int_{\mathbb{P}} \omega(\log(\Lambda(\nu)^{-1/2} x \Lambda(\nu)^{-1/2}) d\nu(x)$$
$$=: \int_{\mathbb{P}} \psi_1(y)\, d\mu(y) - \int_{\mathbb{P}} \psi_2(x)\, d\nu(x).$$

Notice that

$$\psi_1(y) - \psi_2(x) \le d_\infty(x, y).$$

Indeed,

$$\begin{aligned}\psi_1(y) - \psi_2(x) &= \omega(\log(\Lambda(\nu)^{-1/2} y \Lambda(\nu)^{-1/2}) - \log(\Lambda(\nu)^{-1/2} x \Lambda(\nu)^{-1/2})) \\ &\le \| \log(\Lambda(\nu)^{-1/2} y \Lambda(\nu)^{-1/2}) - \log(\Lambda(\nu)^{-1/2} x \Lambda(\nu)^{-1/2})\| \\ &\le d_\infty(x, y),\end{aligned}$$

from the EMI inequality and the invariance of the Thompson metric d_∞; that is, the map $x \mapsto \omega(\log(\Lambda(\nu)^{-1/2} x \Lambda(\nu)^{-1/2}))$ is 1-Lipschitz on the metric cone $(\mathbb{P}, d_\infty)$. Hence, by the Kantorovich duality (see [34, Theorem 5.10]), we immediately obtain that

$$\left\| \int_{\mathbb{P}} \Lambda(\nu)^{-1/2} \varphi(\Lambda(\nu), y) \Lambda(\nu)^{-1/2}\, d(\mu - \nu)(y) \right\| \le W_1(\mu, \nu).$$

Thus, we obtain

$$d_\infty(\Lambda(\mu), \Lambda(\nu)) \le W_1(\mu, \nu) \tag{17}$$

as claimed.

Here is our general convergence result for barycenters.

Corollary 3.7. *For a sequence $\{\mu_n\}_n$ of probability measures, if $\mu_n \to \mu$ weakly and*

$$\begin{aligned}\lim_{n\to\infty} &\int_{\mathbb{P}} \|\Lambda_\varphi(\mu)^{-1/2} \varphi(\Lambda_\varphi(\mu), y) \Lambda_\varphi(\mu)^{-1/2}\|\, d\mu_n(y) \\ &= \int_{\mathbb{P}} \|\Lambda_\varphi(\mu)^{-1/2} \varphi(\Lambda_\varphi(\mu), y) \Lambda_\varphi(\mu)^{-1/2}\|\, d\mu(y)\end{aligned} \tag{18}$$

holds for any $x \in \mathbb{P}$, then $d_\infty(\Lambda_\varphi(\mu), \Lambda_\varphi(\mu_n)) \to 0$.

Proof. For simplicity, let us use the notation

$$\Lambda(y) = \Lambda_\varphi(\mu)^{-1/2} \varphi(\Lambda_\varphi(\mu), y) \Lambda_\varphi(\mu)^{-1/2}.$$

Set $B_M = \{y \in \mathbb{P} \colon \|\Lambda(y)\| < M\}$. Note that the convergence conditions give the 'tightness' of the measures; that is, the uniform integrability condition

$$\lim_{M\to\infty} \limsup_{n\to\infty} \int\limits_{\mathbb{P}\setminus B_M} \|\Lambda(y)\|\, d\mu_n(y) = 0$$

holds.

From Proposition 3.6 we have the inequality

$$d_\infty(\Lambda_\varphi(\mu), \Lambda_\varphi(\mu_n)) \leq \frac{1}{\alpha}\left\| \int\limits_{\mathbb{P}} \Lambda(y)\, d(\mu - \mu_n)(y) \right\|.$$

For any Borel set $B_M \subseteq \mathbb{P}$, we have $\sup_{y\in B_M} \|\Lambda(y)\| < \infty$. Thus the weak convergence of probability measures also implies

$$\int\limits_{B_M} \Lambda(y)\, d\mu_n(y) \to \int\limits_{B_M} \Lambda(y)\, d\mu(y),$$

$n \to \infty$, in norm (see [28, Theorem 2]). Hence, for any $\varepsilon > 0$,

$$\begin{aligned}\left\| \int\limits_{\mathbb{P}} \Lambda(y)\, d(\mu - \mu_n)(y) \right\| &\leq \left\| \int\limits_{B_M} \Lambda(y)\, d(\mu - \mu_n)(y) \right\| + \int\limits_{\mathbb{P}\setminus B_M} \|\Lambda(y)\|\, d(\mu + \mu_n)(y) \\ &\leq 3\varepsilon\end{aligned}$$

for any sufficiently large M and n, which completes the proof. □

It is known that a continuity result such as Corollary 3.7 implies uniqueness of the solution to the corresponding operator equation for probability measures, even with possibly unbounded support. This holds as long as an L^1-integrability condition, as described below, is satisfied. See, for example, [23, Theorem 4.11].

Proposition 3.8. *Let $\mu, \nu \in \mathcal{P}(\mathbb{P})$ and $\{\varphi(\cdot, y)\}_{y\in\mathbb{P}}$ be a family of functions such that*

$$\alpha \equiv \alpha\left(\int\limits_{\mathbb{P}} \varphi(x, y) d\mu(y) \right) > 0.$$

Then, for any $z \in \mathbb{P}$ and $\lambda > 0$ we have

$$d_\infty(J_\lambda^\mu(z), J_\lambda^\nu(z)) \leq \frac{\lambda}{1 + \alpha\lambda} \left\| \int\limits_{\mathbb{P}} J_\lambda^\nu(z)^{-1/2} \varphi(J_\lambda^\nu(z), y) J_\lambda^\nu(z)^{-1/2}\, d(\mu - \nu)(y) \right\|.$$

Proof. Fix $z \in \mathbb{P}$ and $\lambda > 0$. Let us introduce the function

$$\psi^m(x) = \lambda \int_{\mathbb{P}} \varphi(x, y)\, dm(y) + \log_x z$$

for any $m \in \mathcal{P}(\mathbb{P})$. Clearly, $\psi^\nu(J_\lambda^\nu(z)) = 0$. Moreover, we note that

$$\begin{aligned}\psi^\mu(J_\lambda^\nu(z)) &= \psi^\mu(J_\lambda^\nu(z)) - \psi^\nu(J_\lambda^\nu(z)) \\ &= \lambda \int_{\mathbb{P}} \varphi(J_\lambda^\nu(z), y)\, d(\mu - \nu)(y).\end{aligned}$$

Again, $\alpha(\psi^\mu) \geq 1 + \alpha\lambda$. From inequality (13) applied to the function $x \mapsto \psi^\mu(x)$, we obtain

$$\begin{aligned}d_\infty(J_\lambda^\mu(z), J_\lambda^\nu(z)) &\leq \frac{1}{1+\alpha\lambda} \| J_\lambda^\nu(z)^{-1/2} \psi^\mu(J_\lambda^\nu(z)) J_\lambda^\nu(z)^{-1/2} \| \\ &= \frac{\lambda}{1+\alpha\lambda} \left\| \int_{\mathbb{P}} J_\lambda^\nu(z)^{-1/2} \varphi(J_\lambda^\nu(z), y) J_\lambda^\nu(z)^{-1/2}\, d(\mu-\nu)(y) \right\|. \quad \square\end{aligned}$$

Corollary 3.9. *Let $\mu, \nu \in \mathcal{P}(\mathbb{P})$ and $\{\varphi(\cdot, y)\}_{y \in \mathbb{P}}$ be a family of functions such that*

$$\alpha \equiv \alpha \left(\int_{\mathbb{P}} \varphi(x, y) d\mu(y) \right) > 0.$$

Then, for any x and $z \in \mathbb{P}$ and $\lambda > 0$ we have

$$\begin{aligned}d_\infty(J_\lambda^\mu(x), J_\lambda^\nu(y)) &\leq d_\infty(J_\lambda^\mu(x), J_\lambda^\mu(y)) + d_\infty(J_\lambda^\mu(y), J_\lambda^\nu(y)) \\ &\leq \frac{1}{1+\alpha\lambda} d_\infty(x, y) \\ &\quad + \frac{\lambda}{1+\alpha\lambda} \left\| \int_{\mathbb{P}} J_\lambda^\nu(y)^{-1/2} \varphi(J_\lambda^\nu(y), z) J_\lambda^\nu(y)^{-1/2}\, d(\mu-\nu)(z) \right\| \\ &\leq \frac{1}{1+\alpha\lambda} d_\infty(x, y) + \frac{\lambda K}{1+\alpha\lambda} W_1(\mu, \nu)\end{aligned}$$

where $\mathrm{Lip}_\infty(J_\lambda^\nu(y)^{-1/2} \varphi(J_\lambda^\nu(y), \cdot) J_\lambda^\nu(y)^{-1/2}) \leq K$.

Proof. We only show how to obtain the last inequality, the others are straightforward implications of the previous results. There exists a norming linear functional ω such that

$$\begin{aligned}
&\left\|\int_{\mathbb{P}} J_\lambda^\nu(y)^{-1/2}\varphi(J_\lambda^\nu(y),z)J_\lambda^\nu(y)^{-1/2}\,d(\mu-\nu)(z)\right\| \\
&=\omega\left(\int_{\mathbb{P}} J_\lambda^\nu(y)^{-1/2}\varphi(J_\lambda^\nu(y),z)J_\lambda^\nu(y)^{-1/2}\,d(\mu-\nu)(z)\right) \\
&=\int_{\mathbb{P}} \omega\left(J_\lambda^\nu(y)^{-1/2}\varphi(J_\lambda^\nu(y),z)J_\lambda^\nu(y)^{-1/2}\right)\,d(\mu-\nu)(z) \\
&\leq KW_1(\mu,\nu)
\end{aligned}$$

where to obtain the last inequality we used Kantorovich duality. □

4. Holbrook's nodice theorem

In this section we shall prove a deterministic version of resolvent iteration theorems. Pick a $y_0 \in \mathbb{P}$. Let us define the sequence of resolvent iterations:

$$y_k = J_{\frac{1}{k}}^{\varphi}(y_{k-1}) \equiv J_{\frac{1}{k}}(\varphi, y_{k-1}), \quad k = 1, 2, \ldots.$$

Theorem 4.1. *Let $\varphi \in \mathcal{C}$. Let x_* denote the unique solution to the equation $\varphi(x) = 0$. Then the resolvent sequence $y_{n+1} = J_{\frac{1}{n+1}}(\varphi, y_n)$, $n = 0, 1 \ldots$, converges to x_* in the Thompson metric for any initial $y_0 \in \mathbb{P}$.*

Proof. From the definition of the resolvent sequence, we have

$$\frac{1}{k}\varphi(y_k) + \log_{y_k} y_{k-1} = 0,$$

that is,

$$\frac{1}{k}\varphi(y_k) + \log_{y_k} x_* = \log_{y_k} x_* - \log_{y_k} y_{k-1}.$$

Let us define the function $\psi(x) = \frac{1}{k}\varphi(x) + \log_x x_*$. Here the best contraction rate $\alpha(\psi) \geq 1 + \alpha/k$ follows from (8). Since $\psi(x_*) = 0$, relying on inequality (13), we obtain

$$\begin{aligned}
d_\infty(x_*, y_k) &\leq \left(1 + \frac{\alpha}{k}\right)^{-1} \left\|\log(y_k^{-1/2}x_*y_k^{-1/2}) - \log(y_k^{-1/2}y_{k-1}y_k^{-1/2})\right\| \\
&\leq \left(1 + \frac{\alpha}{k}\right)^{-1} d_\infty(x_*, y_{k-1}),
\end{aligned}$$

where in the last step we used the EMI inequality and the conjugate invariance property of the metric d_∞.

Hence, a straightforward computation shows that

$$d_\infty(x_*, y_n) \le d_\infty(x_*, y_0) \prod_{k=1}^{n} \left(1 + \frac{\alpha}{k}\right)^{-1} = O\left(\frac{1}{n^\alpha}\right);$$

the proof is done. □

Let us choose the functions $\varphi_1, ..., \varphi_k \in \mathcal{C}$. Let $\Lambda \in \mathbb{P}$ stand for the unique solution to the equation

$$\frac{1}{k}\sum_{i=1}^{k} \varphi_i(\Lambda) = 0.$$

Theorem 4.2 *(Nodice Theorem). Let $\varphi_1, \ldots, \varphi_k \in \mathcal{C}$ be Lipschitz functions on any bounded d_∞-metric ball. Let us define the cyclic resolvent iteration sequence*

$$s_{n+1} = J_{\frac{1}{n+1}}(\varphi_{\overline{n+1}}, s_n),$$

($\overline{n}$ stands for the remainder class modulo k of n and $\varphi_0 := \varphi_k$). Then the sequence s_n converges to Λ in the Thompson metric for any initial data $s_0 \in \mathbb{P}$.

First, we need to prove a preliminary lemma.

Lemma 4.3. *Let us assume that $x, y \in \mathbb{P}$ and $\lambda := d_\infty(x, y) \le \log 3/2$. Then*

$$\|\log_x y - (y - x)\| \le \frac{1}{4\log^2 3/2}\lambda^2\|x\|.$$

Proof. First, we have

$$e^{-\lambda}x \le y \le e^{\lambda}x;$$

that is,

$$e^{-\lambda} - 1 \le x^{-1/2}yx^{-1/2} - 1 \le e^{\lambda} - 1.$$

Thus $\|x^{-1/2}yx^{-1/2} - 1\| \le \max(1 - e^{-\lambda}, e^{\lambda} - 1) = e^{\lambda} - 1 < 1$. Next, from the power series expansion of the function $x \mapsto \log(1 + x)$,

$$\begin{aligned} x^{1/2}\log(x^{-1/2}yx^{-1/2})x^{1/2} &= \sum_{j=1}^{\infty} \frac{(-1)^j}{j} x^{1/2}(x^{-1/2}yx^{-1/2} - 1)^j x^{1/2} \\ &= y - x + x^{1/2}(1 - x^{-1/2}yx^{-1/2})^2 \sum_{j=0}^{\infty} \frac{(-1)^j}{j+2}(1 - x^{-1/2}yx^{-1/2})^j x^{1/2}. \end{aligned}$$

Therefore,

$$\begin{aligned}\|\log_x y-(y-x)\| &\le \frac{1}{2}\|x\|(1-e^{\lambda})^2\sum_{j=0}^{\infty}\|1-x^{-1/2}yx^{-1/2}\|^j\\ &=\frac{1}{2}(1-e^{\lambda})^2\frac{\|x\|}{1-\|1-x^{-1/2}yx^{-1/2}\|}\\ &\le \frac{1}{2}\frac{(1-e^{\lambda})^2}{2-e^{\lambda}}\|x\|\end{aligned}$$

We note that $(1-e^{\lambda})^2(2-e^{\lambda})^{-1}\le(1/2\log^2 3/2)\lambda^2$ for $0\le\lambda\le\log 3/2$ by simple calculation, hence we are done. □

We shall also use the following lemma from [27, Lemma 2.1].

Lemma 4.4. *Let $\{a_k\}_k$ be a sequence of nonnegative real numbers satisfying*

$$a_{k+1}\le\left(1-\frac{\alpha}{k+1}\right)a_k+\frac{\beta}{(k+1)^2},$$

where $\alpha,\beta>0$. Then

$$a_{k+1}\le\begin{cases}\frac{1}{(k+2)^{\alpha}}\left(a_0+\frac{2^{\alpha}\beta(2-\alpha)}{(1-\alpha)}\right) & \text{if } 0<\alpha<1\\ \beta\frac{1+\log(k+1)}{k+1} & \text{if } \alpha=1\\ \frac{1}{(\alpha-1)(k+2)}\left(\beta+\frac{(\alpha-1)a_0-\beta}{(k+2)^{\alpha-1}}\right) & \text{if } \alpha>1.\end{cases}$$

Now we are ready to present the proof of the main theorem of this section.

Proof of Theorem 4.2. Let $\alpha=\frac{1}{k}\sum_{i=1}^{k}\alpha(\varphi_i)$. We claim

$$d_\infty(\Lambda,s_{(n+1)k})\le\frac{1}{1+\frac{\alpha}{n}}d_\infty(\Lambda,s_{nk})+O\left(\frac{1}{n^2}\right)$$

holds for any sufficiently large n.

First, we prove the sequence $\{s_n\}_n$ is bounded with respect to d_∞. Indeed, Corollary 3.4 implies that

$$d_\infty(s_{n+1},\Lambda)\le\frac{n+1}{n+1+\alpha(\varphi_{\overline{n+1}})}d_\infty(s_n,\Lambda)+\frac{\alpha(\varphi_{\overline{n+1}})}{n+1+\alpha(\varphi_{\overline{n+1}})}\frac{\|\Lambda^{-1/2}\varphi_{\overline{n+1}}(\Lambda)\Lambda^{-1/2}\|}{\alpha(\varphi_{\overline{n+1}})},$$

which by induction gives

$$d_\infty(s_{n+1},\Lambda)\le\max_{\overline{n+1}}\left\{\frac{\|\Lambda^{-1/2}\varphi_{\overline{n+1}}(\Lambda)\Lambda^{-1/2}\|}{\alpha(\varphi_{\overline{n+1}})},d_\infty(s_0,\Lambda)\right\}.$$

That is, s_n is bounded in d_∞.

Next, from the definition of the sequence $\{s_n\}_n$

$$\frac{1}{nk+i}\varphi_i(s_{nk+i}) + \log_{s_{nk+i}} s_{nk+i-1} = 0, \quad i = 1, \ldots, k.$$

Summing up these inequalities for all $1 \leq i \leq k$, we obtain that

$$\sum_{i=1}^{k} \frac{1}{nk+i}\varphi_i(s_{nk+i}) + \sum_{i=1}^{k} \log_{s_{nk+i}} s_{nk+i-1} = 0. \tag{19}$$

For the first term, we have

$$\begin{aligned}\sum_{i=1}^{k} \frac{1}{nk+i}\varphi_i(s_{nk+i}) &= \sum_{i=1}^{k} \frac{1}{nk}\varphi_i(s_{nk+i}) + \sum_{i=1}^{k} \left(\frac{1}{nk+i} - \frac{1}{nk}\right)\varphi_i(s_{nk+i}) \\ &= \frac{1}{nk}\sum_{i=1}^{k} \varphi_i(s_{nk+i}) + O\left(\frac{1}{n^2}\right)\end{aligned}$$

due to local boundedness of φ_i-s. Then, for any index n, by the definition of the resolvent $J_{1/n+1}(\varphi_{\overline{n+1}}, s_n)$,

$$d_\infty(s_n, s_{n+1}) = \| \log s_{n+1}^{-1/2} s_n s_{n+1}^{-1/2} \| \leq \frac{1}{n+1} \|s_{n+1}^{-1}\| \|\varphi_{\overline{n+1}}(s_{n+1})\| = O\left(\frac{1}{n}\right),$$

since s_n is bounded in d_∞. Hence a simple argument also implies $d_\infty(s_n, s_{n+i}) \leq O(1/n)$ for any $1 \leq i \leq k$. Since φ_i-s are locally Lipschitzian in d_∞, we have

$$d_\infty(\varphi_i(s_{nk+i}), \varphi_i(s_{nk})) \leq L d_\infty(s_{nk+i}, s_{nk}) \leq O\left(\frac{1}{n}\right),$$

thus

$$\sum_{i=1}^{k} \frac{1}{nk+i}\varphi_i(s_{nk+i}) = \frac{1}{nk}\sum_{i=1}^{k} \varphi_i(s_{(n+1)k}) + O\left(\frac{1}{n^2}\right)$$

Furthermore, to handle the second term in (19) we apply Lemma 4.3 several times for any sufficiently large n, resulting in the following chain of equalities:

$$\begin{aligned}\sum_{i=1}^{k} \log_{s_{nk+i}} s_{nk+i-1} &= \sum_{i=1}^{k} s_{nk+i-1} - s_{nk+i} + O\left(\frac{1}{n^2}\right) \\ &= s_{nk} - s_{(n+1)k} + O\left(\frac{1}{n^2}\right) \\ &= \log_{s_{(n+1)k}} s_{nk} + O\left(\frac{1}{n^2}\right).\end{aligned}$$

Therefore, equation (19) implies that

$$\frac{1}{nk}\sum_{i=1}^{k}\varphi_i(s_{(n+1)k}) + \log_{s_{(n+1)k}} s_{nk} = O\left(\frac{1}{n^2}\right). \tag{20}$$

Let us introduce the auxiliary function ψ_n on $\mathbb{P}$:

$$\psi_n \colon x \mapsto \frac{1}{nk}\sum_{i=1}^{k}\varphi_i(x) + \log_x \Lambda.$$

We shall apply the inequality (13) to ψ_n. We note hat $\psi_n(x) = 0$ if (and only if) $x = \Lambda$ and $\alpha(\psi_n) \geq 1 + \frac{\alpha}{n}$ by (8). Hence

$$\psi_n(s_{(n+1)k}) = \log_{s_{(n+1)k}} \Lambda - \log_{s_{(n+1)k}} s_{nk} + O\left(\frac{1}{n^2}\right)$$

from (20).

We are now in a position to use (13):

$$\begin{aligned} d_\infty(\Lambda, s_{(n+1)k}) &\leq \frac{1}{1+\frac{\alpha}{n}}\left\| s_{(n+1)k}^{-1/2}\psi_n(s_{(n+1)k})s_{(n+1)k}^{-1/2}\right\| \\ &\leq \frac{1}{1+\frac{\alpha}{n}}\left\| s_{(n+1)k}^{-1/2}(\log_{s_{(n+1)k}} \Lambda - \log_{s_{(n+1)k}} s_{nk})s_{(n+1)k}^{-1/2}\right\| + O\left(\frac{1}{n^2}\right) \\ &\leq \frac{1}{1+\frac{\alpha}{n}} d_\infty(\Lambda, s_{nk}) + O\left(\frac{1}{n^2}\right), \end{aligned}$$

where we applied the EMI inequality and the conjugate invariancy of d_∞ in the last step. Hence our initial claim is proven.

We can obtain from these iterative inequalities and Lemma 4.4 that

$$d_\infty(\Lambda, s_{nk}) \to 0$$

if $n \to \infty$. A simple reasoning shows that $d_\infty(s_{nk}, s_{nk+i}) = O(1/n)$ if $1 \leq i \leq k$, hence the proof is done. □

5. Stochastic resolvent iterations

In this section, we shall prove a stochastic version of the Nodice Theorem. We consider the result as an L^1 Strong Law of Large Numbers in the Thompson metric. Let us recall the following extension of the classical real-valued SLLN due to Etemadi [12, Theorem 1]. Consider a sequence of positive weights $\{w_i\}_i$ such that $W_n = \sum_{i=1}^{n} w_i \to \infty$. Assume that

$$\sup_n \frac{nw_n}{W_n} < \infty \quad \text{and} \quad \sup_n \sum_{i=1}^{n} \frac{i|w_{i+1} - w_i|}{W_n} < \infty.$$

If $X_1, X_2, \dots$ is a sequence of real random variables, then

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to X_0 \text{ a.s.} \qquad \text{implies} \qquad \frac{1}{W_n}\sum_{i=1}^{n} w_i X_i \to X_0 \text{ a.s.}$$

Additionally, if the sequence $\{w_i\}_i$ is monotone, the second condition on the averages can be omitted.

Lemma 5.1. *Given $0 < \alpha < 1$, $0 < s_0$ and a sequence $\{b_n\}_n$ of real numbers, let us define the sequence*

$$s_n = \left(1 - \frac{\alpha}{n}\right) s_{n-1} + \frac{\alpha}{n} b_{n-1}, \quad n \geq 1.$$

Then $s_n = \frac{1}{W_n}\left(s_0 + \sum_{i=0}^{n-1} w_i b_i\right)$, where the sequence of weights $\{w_i\}_i$ is monotone and $W_n = \sum_{i=0}^{n} w_i \to \infty$, $n \to \infty$.

Proof. A straightforward computation gives that

$$w_i = \frac{\alpha}{i+1}\prod_{j=1}^{i+1}\left(1 - \frac{\alpha}{j}\right)^{-1} \qquad \text{and} \qquad W_n = \prod_{j=1}^{n}\left(1 - \frac{\alpha}{j}\right)^{-1} = O(n^{\alpha}).$$

Hence, $\sup_n \frac{nw_n}{W_n} < \infty$ and

$$\frac{w_{i+1}}{w_i} = \frac{\frac{\alpha}{i+2}}{\frac{\alpha}{i+1}(1 - \frac{\alpha}{i+2})} = \frac{i+1}{i+2-\alpha} < 1,$$

which immediately gives the proof of the rest of the lemma. □

Let us consider

$$\phi(x) = \int_{\mathbb{P}} \varphi(x, y)\, d\mu(y) \in \mathcal{C},$$

where $\mu \in \mathcal{P}(\mathbb{P})$ for some measurable locally Lipschitzian function family $\varphi(\cdot, y) \in \mathcal{C}$, for any $y \in \mathbb{P}$. For any $\mathbb{P}$-valued random variable Y with law μ, we recall that $J_\lambda^{\delta_Y}(z) \equiv J_\lambda(Y, z)$ denote the solution to the equation

$$\lambda\varphi(x, Y) + \log_x z = 0.$$

Theorem 5.2. *Assume that $\varphi : \mathbb{P} \times \mathbb{P} \to \mathbb{S}$ is Lipschitz on any bounded d_∞-metric ball and*

$$\int_{\mathbb{P}} \|x^{-1/2}\varphi(x,y)x^{-1/2}\| \, d\mu(y) < \infty \tag{21}$$

holds for all $x \in \mathbb{P}$. Additionally, $\alpha(\varphi(\cdot, y)) \geq \alpha > 0$ for any $y \in \mathbb{P}$, and also, $\phi : \mathbb{P} \to \mathbb{S}$ is Lipschitz on any bounded d_∞-metric ball. Let $Y_1, Y_2, \ldots$ be a sequence of i.i.d. random variables with law $\mu \in \mathcal{P}(\mathbb{P})$. Then the stochastic resolvent iteration $S_0 \in \mathbb{P}$ and $S_{n+1} = J_{\frac{1}{n+1}}(Y_{n+1}, S_n)$ almost surely converges to $\Lambda_\varphi(\mu)$ with respect to the Thompson metric.

Proof. Let us define the empirical measure $\mu_k := \frac{1}{k}\sum_{i=1}^k \delta_{Y_i}$, where δ_{Y_i} is the Dirac measure supported on Y_i. Then by Varadarajan's theorem [10, Theorem 11.4.1] the sequence μ_k converges weakly to μ on the support of μ. Since the norm-function is $L^1(\mu)$-integrable, we also have that

$$\int_{\mathbb{P}} \|\Lambda_\varphi(\mu)^{-1/2}\varphi(\Lambda_\varphi(\mu), y)\Lambda_\varphi(\mu)^{-1/2}\| \, d\mu_k(y)$$
$$\to \int_{\mathbb{P}} \|\Lambda_\varphi(\mu)^{-1/2}\varphi(\Lambda_\varphi(\mu), y)\Lambda_\varphi(\mu)^{-1/2}\| \, d\mu(y)$$

if $k \to \infty$, almost surely. Hence,

$$C(\Lambda; \mu, \mu_k) := \left\| \int_{\mathbb{P}} \Lambda_\varphi(\mu)^{-1/2}\varphi(\Lambda_\varphi(\mu), y)\Lambda_\varphi(\mu)^{-1/2} \, d(\mu - \mu_k)(y) \right\| \to 0$$

a.s. follows if $k \to \infty$ (see the proof of Corollary 3.7). Since $\sup_k \mathbb{E}C(\Lambda; \mu, \mu_k) < \infty$, from the Dominated Convergence Theorem, for any positive constant ε, we have

$$\mathbb{E}C(\Lambda; \mu, \mu_k) < \varepsilon\alpha$$

for any sufficiently large k. Pick such a k. Let us introduce a random sequence of empirical measures $\nu_n := \frac{1}{k}\sum_{i=1}^k \delta_{Y_{nk+i}}$. Then, for any $\hat{S}_1 \in \mathbb{P}$, we can define the resolvent sequence

$$\hat{S}_{n+1} = J_{1/n}^{\nu_n}(\hat{S}_n) \qquad n \geq 1.$$

Notice that $\Lambda_\varphi(\mu) = J_{1/n}^{\mu}(\Lambda_\varphi(\mu))$ holds for any n. Hence, from Corollary 2.2 and Corollary 3.9, we obtain that

$$d_\infty(\Lambda_\varphi(\mu), \hat{S}_{n+1}) \leq \frac{1}{1+\frac{\alpha}{n}} d_\infty(\Lambda_\varphi(\mu), \hat{S}_n) + \frac{\frac{\alpha}{n}}{1+\frac{\alpha}{n}} \frac{C(\Lambda; \mu, \nu_n)}{\alpha} \qquad \text{a.s.}$$

Let us define an auxiliary sequence of random variables: $D_1 := d_\infty(\Lambda_\varphi(\mu), \hat{S}_1)$ and

$$D_{n+1} := \frac{1}{1+\frac{\alpha}{n}} D_n + \frac{\frac{\alpha}{n}}{1+\frac{\alpha}{n}} \frac{C(\Lambda; \mu, \nu_n)}{\alpha} \qquad n \geq 1.$$

A simple induction shows that $d_\infty(\Lambda_\varphi(\mu), \hat{S}_n) \leq D_n$ a.s. for any n. Notice that $C(\Lambda; \mu, \nu_n)$ are real-valued i.i.d. random variables. By applying Etemadi's extension to the Strong Law of Large Numbers and Lemma 5.1 to these variables, we have that

$$D_n \to \mathbb{E} \frac{C(\Lambda; \mu, \nu_n)}{\alpha} = \varepsilon$$

almost surely. This also implies that

$$\limsup_{n\to\infty} d_\infty(\Lambda_\varphi(\mu), \hat{S}_n) \leq \limsup_{n\to\infty} D_n = \varepsilon \qquad \text{a.s.} \tag{22}$$

To complete the proof, we claim that

$$d_\infty(S_{nk}, \hat{S}_n) \to 0 \quad \text{a.s.}$$

In fact, we can closely follow the steps of the proof of the deterministic Nodice Theorem 4.2. Let an auxiliary random function Ψ_n be defined on $\mathbb{P}$ as follows:

$$\Psi_n \colon x \mapsto \frac{1}{nk} \sum_{i=1}^{k} \varphi(x, Y_{nk+i}) + \log_x \hat{S}_n.$$

Let us recall that the sequence $\{S_n\}_n$ is almost surely bounded with respect to d_∞. Indeed, Corollary 3.4 shows

$$\begin{aligned} d_\infty(S_{n+1}, \Lambda_\varphi(\mu)) \leq & \frac{n+1}{n+1+\alpha} d_\infty(S_n, \Lambda_\varphi(\mu)) \\ & + \frac{\alpha}{n+1+\alpha} \frac{\|\Lambda_\varphi(\mu)^{-1/2} \varphi(\Lambda_\varphi(\mu), Y_{n+1}) \Lambda_\varphi(\mu)^{-1/2}\|}{\alpha}. \end{aligned}$$

From the Strong Law of Large Numbers, it follows

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^{n} \|\Lambda_\varphi(\mu)^{-1/2} \varphi(\Lambda_\varphi(\mu), Y_i) \Lambda_\varphi(\mu)^{-1/2}\| \\ & \to \int_{\mathbb{P}} \|\Lambda_\varphi(\mu)^{-1/2} \varphi(\Lambda_\varphi(\mu), y) \Lambda_\varphi(\mu)^{-1/2}\| \, d\mu(y) \end{aligned}$$

a.s. Again, Etemadi's extension to the Strong Law [12, Theorem 1] implies that

$$\limsup_{n\to\infty} d_\infty(S_n, \Lambda_\varphi(\mu)) \le \int_{\mathbb{P}} \|\Lambda_\varphi(\mu)^{-1/2}\varphi(\Lambda_\varphi(\mu), y)\Lambda_\varphi(\mu)^{-1/2}\|\, d\mu(y)$$

a.s.; that is, S_n is bounded a.s.

We now use a stochastic version of (20)

$$\frac{1}{nk}\sum_{i=1}^{k}\varphi(S_{(n+1)k}, Y_{nk+i}) + \log_{S_{(n+1)k}} S_{nk} = O\left(\frac{1}{n^2}\right) \quad \text{a.s. .}$$

Hence

$$\Psi_n(S_{(n+1)k}) = \log_{S_{(n+1)k}} \hat{S}_n - \log_{S_{(n+1)k}} S_{nk} + O\left(\frac{1}{n^2}\right).$$

Clearly, $\Psi_n(\hat{S}_{n+1}) = 0$ holds and $\alpha(\Psi_n) \ge 1 + (nk)^{-1}\sum_{i=1}^{k}\alpha = 1 + n^{-1}\alpha$. Again, inequality (13) implies

$$\begin{aligned} d_\infty(\hat{S}_{n+1}, S_{(n+1)k}) &\le \frac{1}{1+\frac{\alpha}{n}}\left\|S_{(n+1)k}^{-1/2}\Psi_n(S_{(n+1)k})S_{(n+1)k}^{-1/2}\right\| \\ &\le \frac{1}{1+\frac{\alpha}{n}}\left\|S_{(n+1)k}^{-1/2}(\log_{S_{(n+1)k}} \hat{S}_n - \log_{S_{(n+1)k}} S_{nk})S_{(n+1)k}^{-1/2}\right\| + O\left(\frac{1}{n^2}\right) \\ &\le \frac{1}{1+\frac{\alpha}{n}} d_\infty(\hat{S}_n, S_{nk}) + O\left(\frac{1}{n^2}\right), \end{aligned}$$

where we used the EMI property and the conjugate invariancy of d_∞ in the last step. Thus, we get $d_\infty(\hat{S}_{n+1}, S_{(n+1)k}) \to 0$ if $n \to \infty$. Additionaly, it is easy to see that $d_\infty(S_{nk}, S_{nk+i}) \to 0$ holds for any $0 \le i \le k$ as $n \to \infty$. Combined with (22), it follows that

$$\limsup_{n\to\infty} d_\infty(S_n, \Lambda_\varphi(\mu)) \le \varepsilon$$

almost surely. Since $\varepsilon > 0$ was arbitrary, the proof is complete. □

6. Applications to operator means

In this section we shall present some applications of the previous results related to operator means and operator monotone functions. Let us recall the celebrated Loewner integral representation formula of operator monotone functions is (16). The (non-commutative) perspectives of f are given by the maps

$$x \mapsto f_x(z) := x^{1/2} f(x^{-1/2} z x^{-1/2}) x^{1/2}.$$

By Proposition 2.8, for any operator monotone function $f\colon (0,\infty) \to \mathbb{R}$ with $f(1) = 0, f'(1) = 1$, the perspectives $x \in \mathbb{P} \mapsto f_x(z) := x^{1/2}f(x^{-1/2}zx^{-1/2})x^{1/2}$ and their integrals

$$x \mapsto \varphi(x) := \int_{\mathbb{P}} f_x(z)\, d\mu(z),$$

where $\mu \in \mathcal{P}(\mathbb{P})$ such that

$$\int_{\mathbb{P}} \|f_y(z)\|\, d\mu(z) < \infty \tag{23}$$

for any $y \in \mathbb{P}$, define monotone flows on $\mathbb{P}$ through the first-order system

$$\dot{x}(t) = \varphi(x(t)) = \int_{\mathbb{P}} f_{x(t)}(z)\, d\mu(z)$$

for $t \geq 0$. This is indeed the case, since it is a straightforward exercise to see that the $f.(\cdot)$ are locally Lipschitz, thus φ as well by Proposition 2.7. If

$$0 = \varphi(x) = \int_{\mathbb{P}} f_x(z)\, d\mu(z)$$

has a (unique) solution, we denote it by $\Lambda_f(\mu)$.

It is also well known that

$$I - x^{-1} \leq f(x) \leq x - I \tag{24}$$

for $x \in \mathbb{P}$ where the upper bound follows from the concavity of f and the normalizations $f(1) = 0, f'(1) = 1$; the lower bound is implied by this and by the fact that $-f(x^{-1})$ is also operator monotone with the same normalizations.

Theorem 6.1. *Let f be an operator monotone function as above with $f(1) = 0, f'(1) = 1$ and assume that $\mu \in \mathcal{P}(\mathbb{P})$ such that (23) holds. Then the first-order system*

$$\dot{x}(t) = \int_{\mathbb{P}} f_{x(t)}(z)\, d\mu(z)$$

generates an order-preserving flow. Furthermore, it has an exponential contraction rate $\alpha > 0$ on any closed ball $\overline{B}(I, c) = \{y \in \mathbb{P} : d_\infty(I, y) \leq c\}$.

Proof. Firstly, the flow is order-preserving, as established in Proposition 2.8. We rely on the Gauber-Qu condition (11) and the integral representation (16) to obtain

$$
\begin{aligned}
D(x^{-1/2}f_x(z)x^{-1/2})(x) &= D(f(x^{-1/2}zx^{-1/2}))(x)\\
&= Df(x^{-1/2}zx^{-1/2})(D(x^{-1/2}zx^{-1/2})(x))\\
&= -Df(x^{-1/2}zx^{-1/2})(x^{-1/2}zx^{-1/2})\\
&= -x^{-1/2}zx^{-1/2}f'(x^{-1/2}zx^{-1/2}) \leq -\alpha_z.
\end{aligned}
$$

Note that $x^{-1/2}zx^{-1/2} \in \mathbb{P}$ for $z \in \mathrm{supp}(\mu)$. Now let us choose any order interval $[\frac{1}{c}I, cI]$ with $c > 0$. Then f being non-constant and the representation (16) ensure $f' > 0$, $f'' \leq 0$ on $(0, \infty)$, and it is easy to find some measurable positive lower bound on $[\frac{1}{c}I, cI]$ for α_z for $z \in \mathbb{P}$. Then choosing $\alpha := \int_{\mathbb{P}} \alpha_z \, d\mu(z)$ provides a suitable positive lower bound on the exponential contraction rate of φ on $[\frac{1}{c}I, cI]$. □

Lemma 6.2. *Let f be an operator monotone function as above with $f(1) = 0, f'(1) = 1$ and let $z \in \mathbb{P}$. Then the first-order system*

$$\dot{x}(t) = f_{x(t)}(z)$$

leaves the half-open order intervals $[z, \infty)$ and $(0, z]$ invariant.

Proof. First note that $f_z(z) = 0$, thus the orbit starting at $x(0) := z$ is stationary. The generated flow is order preserving by Theorem 6.1, thus any orbit with starting point $x(0) \geq z$ has $x(t) \geq z$ and similarly $x(0) \leq z$ implies $x(t) \leq z$ for all $t \geq 0$. This ensures the flow invariance of the order intervals $[z, \infty)$ and $(0, z]$. □

By [13, Remark 2.4.] the flow invariance of a closed convex set S is equivalent to

$$\int_{\mathbb{P}} f_x(y) \, d\mu(y) \in T_S(x), \forall x \in S,$$

where $T_S(x)$ is the tangent cone of S which by [13, Proposition 2.5.] can be calculated as

$$T_S(x) := \overline{\{v : \exists \lambda > 0, x + \lambda v \in S\}}, \forall x \in S. \tag{25}$$

Proposition 6.3. *Under the assumptions of the previous Theorem 6.1 if*

$$\int_{\mathbb{P}} \|z^{\pm 1}\| \, d\mu(z) \leq \int_{\mathbb{P}} e^{d_\infty(z,I)} \, d\mu(z) < \infty, \tag{26}$$

then the flow leaves the order-interval $\left[(\int_{\mathbb{P}} z^{-1} \, d\mu(z))^{-1}, \int_{\mathbb{P}} z \, d\mu(z)\right]$ invariant.

Proof. For any $x \in \mathbb{P}$ by (24) we have the following

$$I - \int_{\mathbb{P}} (x^{-1/2} y x^{-1/2})^{-1} \, d\mu(y) \leq \int_{\mathbb{P}} f(x^{-1/2} y x^{-1/2}) \, d\mu(y) \leq \int_{\mathbb{P}} x^{-1/2} y x^{-1/2} \, d\mu(y) - I$$

which is equivalent to

$$\begin{aligned} I - \left(x^{-1/2} \left(\int_{\mathbb{P}} y^{-1} \, d\mu(y) \right)^{-1} x^{-1/2} \right)^{-1} &\leq \int_{\mathbb{P}} f(x^{-1/2} y x^{-1/2}) \, d\mu(y) \\ &\leq x^{-1/2} \int_{\mathbb{P}} y \, d\mu(y) x^{-1/2} - I \end{aligned}$$

and by conjugation with $x^{1/2}$ this further implies

$$x - x^{1/2} \left(x^{-1/2} \left(\int_{\mathbb{P}} y^{-1} \, d\mu(y) \right)^{-1} x^{-1/2} \right)^{-1} x^{1/2} \leq \int_{\mathbb{P}} f_x(y) \, d\mu(y) \leq \int_{\mathbb{P}} y \, d\mu(y) - x. \tag{27}$$

Note that the left hand side above vanishes when $x = \left(\int_{\mathbb{P}} y^{-1} \, d\mu(y) \right)^{-1}$ and similarly the right hand side vanishes if $x = \int_{\mathbb{P}} y \, d\mu(y)$. Then by Lemma 6.2 the flow generated by the function on the left hand side of (27) leaves the half-open order interval $\left[\left(\int_{\mathbb{P}} y^{-1} \, d\mu(y) \right)^{-1}, \infty \right)$ invariant, while the flow generated by the function on the right hand side of (27) leaves the half-open order interval $\left(0, \int_{\mathbb{P}} y \, d\mu(y) \right]$ invariant. This yields that the left hand side of (27) is an element of $T_{\left[\left(\int_{\mathbb{P}} y^{-1} \, d\mu(y) \right)^{-1}, \infty \right)}(x)$, thus $\int_{\mathbb{P}} f_x(y) \, d\mu(y)$ is also an element by (27). Similarly the right hand side of (27) is an element of $T_{(0, \int_{\mathbb{P}} y \, d\mu(y)]}(x)$, thus $\int_{\mathbb{P}} f_x(y) \, d\mu(y)$ is by (27). When combined, these yield that

$$\int_{\mathbb{P}} f_x(y) d\mu(y) \in T_{\left[\left(\int_{\mathbb{P}} y^{-1} \, d\mu(y) \right)^{-1}, \int_{\mathbb{P}} y \, d\mu(y) \right]}(x), \quad \forall x \in \left[\left(\int_{\mathbb{P}} y^{-1} \, d\mu(y) \right)^{-1}, \int_{\mathbb{P}} y \, d\mu(y) \right].$$

This proves that the flow generated by $\int_{\mathbb{P}} f_x(y) \, d\mu(y)$ leaves the closed convex order interval $\left[\left(\int_{\mathbb{P}} y^{-1} \, d\mu(y) \right)^{-1}, \int_{\mathbb{P}} y \, d\mu(y) \right]$ invariant. □

What follows is a more general version of the above with a different proof.

Proposition 6.4. *Under the assumptions of Theorem 6.1, there exists an* $1 < R \in \mathbb{R}$ *such that the flow leaves* $\left[\frac{1}{R} I, RI \right]$ *invariant.*

Proof. First of all note that $f : (0, \infty) \mapsto \mathbb{R}$ is operator monotone, so it holds that

$$\int_A f\left(\frac{1}{K}y\right)\, d\mu(y) \leq \int_A f(y)\, d\mu(y) \leq \int_A f(Ky)\, d\mu(y) \tag{28}$$

for any $1 < K \in \mathbb{R}$ and Borel set $A \subseteq \mathbb{P}$. By the integrability assumption (23) for any $\varepsilon > 0$ there exists $0 < R_\varepsilon \in \mathbb{R}$ large enough such that

$$\int_{\mathbb{P}\setminus B(I,R_\varepsilon)} \|f(y)\|\, d\mu(y) < \varepsilon. \tag{29}$$

Moreover, the left hand side of (29) is a decreasing function in $R_\varepsilon > 0$. By (28) this also implies that

$$\int_{\mathbb{P}\setminus B(I,R_\varepsilon)} f\left(e^{-K}y\right)\, d\mu(y) < \varepsilon I$$

for any $K > 0$. For $y \in B(I, R_\varepsilon)$ and $K > R_\varepsilon$ we have that $e^{-K}y < I$ which by (24) implies $f(e^{-K}y) \leq e^{-K}y - I < 0$ and then

$$\int_{B(I,R_\varepsilon)} f(e^{-K}y)\, d\mu(y) \leq e^{-K} \int_{B(I,R_\varepsilon)} y\, d\mu(y) - I < -2\varepsilon I$$

for all large enough $K > R_\varepsilon$. Thus we have

$$\int_{\mathbb{P}} f\left(e^{-K}y\right)\, d\mu(y) < -\varepsilon I.$$

for all large enough $K > R_\varepsilon > 0$. This shows

$$\int_{\mathbb{P}} f_{KI}(y)\, d\mu(y) < -\delta_1 I$$

for a $\delta_1 > 0$.

In a similar way, but now using the right hand side of (28) and the left hand side of (24) for any $\varepsilon > 0$ we have

$$\int_{\mathbb{P}} f\left(e^{k}y\right)\, d\mu(y) > \varepsilon I.$$

for all large enough $k > 0$, yielding that for all small enough $1 > k \in \mathbb{R}$ we have

$$\int_{\mathbb{P}} f_{kI}(y)\,d\mu(y) > \delta_2 I.$$

At this point we define $R := \max\{K, 1/k\}$ and note that still

$$\int_{\mathbb{P}} f_{RI}(y)\,d\mu(y) < -\delta_1 I, \int_{\mathbb{P}} f_{\frac{1}{R}I}(y)\,d\mu(y) > \delta_2 I.$$

By Proposition 2.7 $x \mapsto \int_{\mathbb{P}} f_x(y)$ is locally Lipschitz, thus in a neighborhood of $x = RI$ we have

$$\int_{\mathbb{P}} f_x(y)\,d\mu(y) < -\frac{\delta_1}{2} I, \tag{30}$$

and in a neighborhood of $x = \frac{1}{R}I$ we have

$$\int_{\mathbb{P}} f_x(y)\,d\mu(y) > \frac{\delta_2}{2} I. \tag{31}$$

Then (30) and (31) imply that

$$\int_{\mathbb{P}} f_x(y)\,d\mu(y) \in T_{[\frac{1}{R}I,RI]}(x)$$

for x in a neighborhood of $x = \frac{1}{R}I$ and $x = RI$. Since the generated flow is monotone, this further implies that

$$\int_{\mathbb{P}} f_x(y)\,d\mu(y) \in T_{[\frac{1}{R}I,RI]}(x), \forall x \in \left[\frac{1}{R}I, RI\right]$$

proving the flow invariance. □

Theorem 6.5. *Under the assumptions of the previous Proposition 6.4, $\Lambda_f(\mu)$ exists and is unique.*

Proof. Theorem 6.1 and Proposition 6.4 yields the uniqueness of Λ_f in $\left[\frac{1}{R}I, RI\right]$ by Proposition 3.1, because the orbits γ of the flow do not leave the order interval $\left[\frac{1}{R}I, RI\right]$ as long as $\gamma(0) \in \left[\frac{1}{R}I, RI\right]$ and on $\left[\frac{1}{R}I, RI\right]$ the flow has an exponential contraction rate $\alpha > 0$, so we can apply the same reasoning as in Proposition 3.1. Furthermore, the existence of another stationary point of the flow in $\left[\frac{1}{R}I, RI\right]$ would lead to the existence of two distinct stationary orbits, contradicting the exponential contractivity established in Theorem 6.1. □

The following order preserving property was proved in [29] in the case of bounded $\mathrm{supp}(\mu)$. We can reprove this using our tools.

Proposition 6.6. *For any operator monotone function $f_i\colon (0,\infty) \to \mathbb{R}$ with $f_i(1) = 0, f_i'(1) = 1$, $i \in \{1,2\}$, if $f_1(x) \le f_2(x)$ for all $x \in \mathbb{P}$, then $\Lambda_{f_1}(\mu) \le \Lambda_{f_2}(\mu)$.*

Proof. The flow generated by f_1 leaves $\Lambda_{f_1}(\mu)$ invariant and by the order preserving property $[\Lambda_{f_1}(\mu), \infty)$ is also invariant. Then by the assumptions we have

$$0 = \int_{\mathbb{P}} f_1(\Lambda_{f_1}(\mu)^{-1/2} y \Lambda_{f_1}(\mu)^{-1/2})\, d\mu(y) \le \int_{\mathbb{P}} f_2(\Lambda_{f_1}(\mu)^{-1/2} y \Lambda_{f_1}(\mu)^{-1/2})\, d\mu(y),$$

thus the order preserving flow generated by f_2 leaves the order interval $[\Lambda_{f_1}(\mu), \infty)$ invariant as well by a similar containment argument for the tangent cone of $[\Lambda_{f_1}(\mu), \infty)$ as in the proof of Proposition 6.3, implying $\Lambda_{f_2}(\mu) \in [\Lambda_{f_1}(\mu), \infty)$. □

Corollary 6.7. *Under the assumptions of Proposition 6.3 we have*

$$\left(\int_{\mathbb{P}} y^{-1}\, d\mu(y)\right)^{-1} \le \Lambda_f(\mu) \le \int_{\mathbb{P}} y\, d\mu(y). \tag{32}$$

Proof. The result follows from Proposition 6.6 and (27) along with the direct implications of the latter. □

The *stochastic order* of probability measures over partially ordered spaces were studied in detail for example in [14,18]. It is denoted as $\mu \le \nu$ which means that $\mu(U) \le \nu(U)$ for any open upper set U, that is U is open and $x \in U$ with $x \le y$ implies $y \in U$. Then for instance $\mu \le \nu$ implies $\int_{\mathbb{P}} f_x(z)\, d\mu(z) \le \int_{\mathbb{P}} f_x(z)\, d\nu(z)$ for all $x \in \mathbb{P}$ and then in turn this implies $\Lambda_f(\mu) \le \Lambda_f(\nu)$, i.e. Λ_f is *monotone* which has been first shown in [29] for measures with bounded support. We need a variant of Corollary 3.7 that applies to the above in the case of probability measures with unbounded support.

Corollary 6.8. *For a sequence $\{\mu_n\}_n$ of probability measures if $\mu_n \to \mu$ weakly and (23), (18) hold with respect to $d\mu_n(y)$ and $d\mu(y)$, then $d_\infty(\Lambda_f(\mu), \Lambda_f(\mu_n)) \to 0$.*

Proof. The proof of Theorem 6.5 yields an exponential contraction rate $\alpha > 0$ on the flow invariant $\left[\frac{1}{R}I, RI\right]$, so Corollary 3.7 applies. □

It is desirable to look at $M_t^f(a,b) := \Lambda_f((1-t)\delta_a + t\delta_b)$ for $t \in [0,1]$ and $a, b \in \mathbb{P}$ providing a more "symmetric" resolvent. The motivation for this is to recover the (Sturm-type of) strong law for the arithmetic, harmonic and Karcher means in their known forms. The form of the last one is discussed above in the Introduction.

Theorem 6.9 *(Strong Law of Large Numbers). Let $Y_1, Y_2, \dots$ be a sequence of i.i.d. random variables with the probability law μ on $\mathbb{P}$ satisfying (23) and (26). Then the stochastic iteration $S_0 \in \mathbb{P}$ and $S_{n+1} = M^f_{\frac{1}{n+1}}(Y_{n+1}, S_n)$ almost surely converges to $\Lambda_f(\mu)$ with respect to the Thompson metric.*

Proof. First, we prove that the stochastic iteration $S_0 \in \mathbb{P}$ and

$$\frac{1}{n+1} f_{S_{n+1}}(Y_{n+1}) + f_{S_{n+1}}(S_n) = 0 \qquad n \geq 0,$$

defines a bounded sequence in d_∞ a.s. Indeed, since f is concave, we have

$$f\left({S_{n+1}}^{-1/2}\left(\frac{1}{n+2}Y_{n+1} + \frac{n+1}{n+2}S_n\right)S_{n+1}^{-1/2}\right) \geq 0 = f(I).$$

Thus

$$S_{n+1} \leq \frac{1}{n+2}Y_{n+1} + \frac{n+1}{n+2}S_n$$

holds and

$$S_{n+1} \leq \frac{1}{n+2}\left(S_0 + \sum_{i=1}^{n+1} Y_i\right)$$

follows for any $n \geq 1$. Since $\mathbb{E}(Y_i)$ is finite, the average is weakly bounded above and then strongly in norm as well. A similar argument with $g(x) = -f(1/x)$, which is operator monotone, shows that S_n^{-1} bounded above in norm. This implies that S_n is bounded in d_∞ almost surely.

From the Taylor series expansions, we have

$$f(x) = x - 1 + O((x-1)^2) \quad \text{and} \quad \log x = x - 1 + O((x-1)^2)$$

on bounded sets. Therefore, the stochastic resolvent iteration $\tilde{S}_0 = S_0$ and $\tilde{S}_{n+1} = J_{\frac{1}{n+1}}(Y_{n+1}, \tilde{S}_n)$, $n \geq 0$, and the iteration S_n above are close in metric; that is,

$$d_\infty(S_n, \tilde{S}_n) = O\left(\frac{1}{n^2}\right). \tag{33}$$

Obviously, $\tilde{S}_n$ is also bounded in d_∞ a.s.

Next, let us recall that the first-order system

$$\dot{x} = \int_{\mathbb{P}} f_x(s)\, d\mu(s)$$

leaves any sufficiently large order-interval $Q = [c^{-1}I, cI]$ invariant, where the exponential contraction rate of the flow is positive, see Theorem 6.1 and Proposition 6.4. We also note that $\alpha(Q, f_x(s)) > 0$ holds for any $s \in \mathbb{P}$ (see the proof of Theorem 6.1). Moreover, $\mathbb{E}(\alpha(Q, f_x(Y)) < \infty$, because without loss of generality we may assume that $\alpha(Q, f_x(s)) \leq 1$ for any $s \in \mathbb{P}$. Hence, the SLLN implies

$$\frac{1}{k}\sum_{i=1}^{k}\alpha(Q, f_x(Y_i)) \to \mathbb{E}(\alpha(Q, f_x(Y))) > 0.$$

Choosing a large invariant set Q that contains the bounded resolvent iteration $\tilde{S}_n$, we can utilize the proof of Theorem 5.2 on the set Q to obtain that $d_\infty(\tilde{S}_n, \Lambda_f(\mu)) \to 0$ almost surely. With the approximation (33) above at hand, the proof is complete. □

A similar modified version of Holbrook's Nodice Theorem 4.2 follows in the same vein as well.

Remark 6.10. These results imply the ones in [24] for the geometric also called Karcher mean Λ of positive operators by choosing $f(x) = \log(x)$. In fact in this case we can apply Theorem 5.2 directly, since $\alpha = 1$ globally by Proposition 2.5.

Furthermore, the usual strong law for the arithmetic mean $(1-t)a + tb$ with its multivariable version $\int_{\mathbb{P}} a\, d\mu(a)$ follows when $f(x) = x - 1$; and for its cousin, the harmonic mean $[(1-t)a^{-1} + tb^{-1}]^{-1}$ with its multivariable version $[\int_{\mathbb{P}} a^{-1}\, d\mu(a)]^{-1}$ as well when $f(x) = 1 - \frac{1}{x}$.

More generally, the non-commutative power means defined and studied in [25,19,20] are generated by the one-parameter family $f_t(x) = \frac{x^t - 1}{t}$ of normalized operator monotone functions for $t \in [-1, 1]$. Then Λ_{f_t} gives a non-commutative extension of the one-parameter family of t-power means of positive numbers that varies continuously, and interpolates monotonically in $t \in [-1, 1]$ between the harmonic for $t = -1$, the Karcher for $t = 0$ and the arithmetic means for $t = 1$.

Even more generally our results reprove most of the results of [29] and extend them to the case of unbounded $\mathrm{supp}(\mu)$.

Acknowledgments

The authors would like to thank Zsigmond Tarcsay for the critical remarks and discussions and also the anonymous referee for useful suggestions improving the presentation of the results. The work of M. Pálfia was supported by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund and financed under the TKP2021-NVA funding scheme, Project no. TKP2021-NVA-09; the Hungarian National Research, Development and Innovation Office NKFIH, Grant No. FK128972, Grant No. ÚNKP-23-5 and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, Grant No. BO/00998/23/3.

References

[1] A. de Acosta, Invariance principles in probability for triangular arrays of B-valued random vectors and some applications, Ann. Probab. 10 (2) (1982) 346–373.
[2] J. Antezana, E. Ghiglioni, D. Stojanoff, Ergodic theorem in CAT(0) spaces in terms of inductive means, Ergod. Theory Dyn. Syst. 43 (5) (2023) 1433–1454.
[3] W. Arveson, An Invitation to C^*-Algebras, Springer Science & Business Media, 2012.
[4] R. Bhatia, Matrix Analysis, Springer–Verlag, New York, 1997.
[5] R. Bhatia, J. Holbrook, Riemannian geometry and matrix geometric means, Linear Algebra Appl. 413 (2006) 594–618.
[6] R. Bhatia, R. Karandikar, Monotonicity of the matrix geometric mean, Math. Ann. 353 (4) (2012) 1453–1467.
[7] F.R. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.
[8] S. Cobzas, Lipschitz properties of convex mappings, Adv. Oper. Theory 2 (1) (2017) 21–49.
[9] M.G. Crandall, T.M. Liggett, Generation of semi-groups of nonlinear transformations on general Banach spaces, Am. J. Math. 93 (2) (1971) 265–298.
[10] R.M. Dudley, Real Analysis and Probability, Cambridge Studies in Advanced Mathematics, vol. 74, Cambridge University Press, Cambridge, 2002.
[11] A. Ebadian, I. Nikoufar, M.E. Gordji, Perspectives of matrix convex functions, Proc. Natl. Acad. Sci. USA 108 (18) (2011) 7313–7314.
[12] N. Etemadi, Convergence of weighted averages of random variables revisited, Proc. Am. Math. Soc. 134 (2006) 2739–2744.
[13] S. Gaubert, Z. Qu, The contraction rate in Thompson's part metric of order-preserving flows on a cone – application to generalized Riccati equations, J. Differ. Equ. 256 (8) (2014) 2902–2948.
[14] F. Hiai, J. Lawson, Y. Lim, The stochastic order of probability measures on ordered metric spaces, J. Math. Anal. Appl. 464 (1) (2018) 707–724.
[15] J. Holbrook, No dice: a deterministic approach to the Cartan centroid, J. Ramanujan Math. Soc. 27 (4) (2012) 509–521.
[16] A. Karlsson, G.A. Margulis, A multiplicative ergodic theorem and nonpositively curved spaces, Commun. Math. Phys. 208 (1999) 107–123.
[17] J. Lawson, Geodesic bicombings and a metric Crandall-Liggett theory, J. Lie Theory 33 (1) (2023) 361–376.
[18] J. Lawson, Ordered probability spaces, J. Math. Anal. Appl. 455 (1) (2017) 167–179.
[19] J. Lawson, Y. Lim, Karcher means and Karcher equations of positive definite operators, Trans. Am. Math. Soc. 1 (2014) 1–22.
[20] J. Lawson, Y. Lim, Weighted means and Karcher equations of positive operators, Proc. Natl. Acad. Sci. 110 (39) (2013) 15626–15632.
[21] J. Lawson, Y. Lim, Monotonic properties of the least squares mean, Math. Ann. 351 (2) (2011) 267–279.
[22] J. Lawson, Y. Lim, Symmetric spaces with convex metrics, Forum Math. 19 (2007) 571–602.
[23] Y. Lim, M. Palfia, Strong law of large numbers for the L^1-Karcher mean, J. Funct. Anal. 279 (7) (2020) 108672.
[24] Y. Lim, M. Pálfia, Existence, uniqueness and an ODE approach to the L^1-Karcher mean, Adv. Math. 376 (2021) 107435.
[25] Y. Lim, M. Pálfia, Matrix power means and the Karcher mean, J. Funct. Anal. 262 (4) (2012) 1498–1514.
[26] Y. Lim, M. Pálfia, Weighted deterministic walks for the least squares mean on Hadamard spaces, Bull. Lond. Math. Soc. 46 (3) (2014) 561–570.
[27] A. Nedic, D.P. Bertsekas, Convergence rate of incremental subgradient algorithms, in: Stochastic Optimization: Algorithms and Applications, Gainesville, FL, 2000, in: Appl. Optim., vol. 54, Kluwer Acad. Publ., Dordrecht, 2001, pp. 223–264.
[28] L. Nielsen, Weak convergence and Banach space-valued functions: improving the stability theory of Feynman's operational calculi, Math. Phys. Anal. Geom. 14 (4) (2011) 279–294.
[29] M. Pálfia, Operator means of probability measures and generalized Karcher equations, Adv. Math. 289 (2016) 951–1007.
[30] V. Paulsen, Completely Bounded Maps and Operator Algebras, Cambridge University Press, 2003.
[31] R.M. Redheffer, W. Walter, Flow-invariant sets and differential inequalities in normed spaces, Appl. Anal. 5 (2) (1975) 149–161.

[32] K.-T. Sturm, Nonlinear martingale theory for processes with values in metric spaces of nonpositive curvature, Ann. Probab. 30 (2002) 1195–1222.
[33] A.C. Thompson, On certain contraction mappings in a partially ordered vector space, Proc. Am. Math. Soc. 14 (1963) 438–443.
[34] C. Villani, Optimal Transport: Old and New, Springer, Berlin, 2009.