Weierstraß-Institut
für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Positivity and polynomial decay of energies for square-field operators

Artur Stephan, Holger Stephan

submitted: December 9, 2021

Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: artur.stephan@wias-berlin.de
holger.stephan@wias-berlin.de

No. 2901
Berlin 2021

2020 *Mathematics Subject Classification.* 60J25, 39B62, 35B40, 46B42, 35Q84, 47D07.

Key words and phrases. Square-field operator, carré du champs operator, Markov generator, Markov operator, log-convexity, Banach lattice, positivity, quadratic energies, asymptotic polynomial decay, normal operator.

The research of A.S. has been funded by Deutsche Forschungsgemeinschaft (DFG) through the Collaborative Research Center SFB 1114 'Scaling Cascades in Complex Systems' (Project No. 235221301), Subproject C05 'Effective models for materials and interfaces with multiple scales'.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Positivity and polynomial decay of energies for square-field operators

Artur Stephan, Holger Stephan

Abstract

We show that for a general Markov generator the associated square-field (or carré du champs) operator and all their iterations are positive. The proof is based on an interpolation between the operators involving the generator and their semigroups, and an interplay between positivity and convexity on Banach lattices. Positivity of the square-field operators allows to define a hierarchy of quadratic and positive energy functionals which decay to zero along solutions of the corresponding evolution equation. Assuming that the Markov generator satisfies an operator-theoretic normality condition, the sequence of energies is log-convex. In particular, this implies polynomial decay in time for the energy functionals along solutions.

1 Introduction

The study of the long time behavior of solutions of differential equations

$$\dot{g}(t) = \mathbf{A}g(t),\ g(0) = g_0 \in \mathcal{X}, \tag{1.1}$$

where $\mathbf{A}$ is a (in general unbounded) linear operator in a Banach space $\mathcal{X}$, plays a major role in mathematical physics. Usually, the time asymptotic behavior is studied by investigating energy functionals along the solution. Let $\mathrm{E} : \mathcal{X} \to \mathbb{R}$ be such an energy functional. If E is positive, i.e. $\mathrm{E}(g) \geq 0$ for $g \in \mathcal{X}$, and it decays along the solution, i.e., if the function $t \to \mathrm{E}(g(t))$ has a negative derivative $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}(g(t)) \leq 0$, then one can at least deduce the existence of a limit $\lim_{t\to\infty} \mathrm{E}(g(t)) = \mathrm{E}(g(\infty)) \leq \mathrm{E}(g(0))$.

The prototypical example is given by the scalar diffusion equation $\dot{g} = \frac{1}{2}\Delta g$ on $\Omega \subset \mathbb{R}^d$ with no-flux boundary condition. We define the quadratic energy $\mathrm{E}_0(g) = \int_\Omega g^2 \mathrm{d}x$, and, furthermore, the functionals $\mathrm{E}_1(g) := \int_\Omega |\nabla g|^2 \mathrm{d}x$ and $\mathrm{E}_2(g) := \int_\Omega |\Delta g|^2 \mathrm{d}x$. Obviously, we have $\mathrm{E}_i \geq 0$. Then, along (sufficiently smooth) solutions $t \mapsto g(t)$ of $\dot{g} = \frac{1}{2}\Delta g$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_0(g(t)) = -\mathrm{E}_1(g(t)), \quad \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_1(g(t)) = -\mathrm{E}_2(g(t)),$$

showing the decay of $t \mapsto \mathrm{E}_0(g(t))$ and $t \mapsto \mathrm{E}_1(g(t))$ along solutions. Often, the functional E_1 is called *dissipation rate*, and E_2 would describe the "dissipative loss" of the dissipation. In the following, we will call an operator $\mathbf{A}$ or the corresponding equation of type (1.1) *dissipative* if the dissipation rate is positive. The asymptotic decay can be quantified by, for example, proving a Poincaré-type inequality $\mathrm{E}_{k+1} \geq c\,\mathrm{E}_k$ with some $c > 0$. Then exponential decay is obtained for $t \mapsto \mathrm{E}_k(g(t))$ by the classical Gronwall lemma.

Extending dissipativity to other systems usually operators in divergence form $\mathbf{A} = -\frac{1}{2}\mathbf{D}^*\mathbf{L}\mathbf{D}$ in a suitable Hilbert space are considered. Here $\mathbf{D}$ is an operator, containing the constant function in the

DOI 10.20347/WIAS.PREPRINT.2901 Berlin 2021

kernel (e.g., the gradient), $\mathbf{D}^*$ is its adjoint with respect to a scalar product $(\cdot,\cdot)$ in a suitable Hilbert space, $\mathbf{L}$ is a positive multiplication operator such that $-\mathbf{A}$ is a symmetric operator that is positive in the form sense. Defining the energy $\mathrm{E}_0(g) = (g,g)$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_0(g(t)) = (g, \mathbf{A}g) = -(g, \mathbf{D}^*\mathbf{L}\mathbf{D}g) = -(\mathbf{D}g, \mathbf{L}\mathbf{D}g) =: -\mathrm{E}_1(g)\,.$$

Since $\mathbf{L}$ is a positive multiplication operator, we have $\mathrm{E}_1(g) \geq 0$, and, hence, operators in divergence form are dissipative by construction. Moreover, it is easy to check that for functionals

$$\mathrm{E}_k(g) = (g, (-\mathbf{A})^k g) \tag{1.2}$$

we have $\mathrm{E}_k(g) \geq 0$ and $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_k(g(t)) = -\mathrm{E}_{k+1}(g(t))$.

As *dissipativity*, another important physical property is *positivity*. A solution g, describing for example the concentration or density, should be certainly positive for being physical reasonable[1], and this property of the solution has often to be either proved by hand or assumed. The aim of the paper is to show that dissipativity is the mathematical consequence of the general notion of *positivity* from the theory of Banach lattices.

Linear operators $\mathbf{A}$ in (1.1) preserve positivity and mass if and only if they are Markov generators generating a semigroup $(\mathbf{T}(t))_{t\geq 0}$ of Markov operators. Here, we consider Markov operators on the Banach lattice of continuous functions $\mathcal{C}(\mathcal{Z})$ on a compact topological space $\mathcal{Z}$. Their adjoints map the space of probability measures to itself, ensuring the system to remain physical reasonable.

To investigate dissipativity of the system, the fundamental object will be the so-called *square-field operators* (or *carré du champs operators*). They have been introduced by Meyer and Bakry [Mey82, Mey84, Bak85] and played an important role in the theory of evolution equations and stochastic analysis since then. They are recursively defined by

$$\begin{aligned}\Gamma_0(f,g) &:= f\cdot g\\ \Gamma_{n+1}(f,g) &:= \mathbf{A}\Gamma_n(f,g) - \Gamma_n(\mathbf{A}f,g) - \Gamma_n(f,\mathbf{A}g),\end{aligned}$$

and measure the difference from $\mathbf{A}$ being a derivation (i.e. $\mathbf{A}$ satisfying $\mathbf{A}(f\cdot g) = \mathbf{A}f\cdot g + f\cdot \mathbf{A}g$). In the following, we use the notation $\Gamma_n(g) = \Gamma_n(g,g)$ for evaluating at the diagonal $f = g$. Many interesting features have been investigated for square-field operators in the last decades, connecting analysis, stochastics and geometry. The pioneering Bakry-Émery [BaÉ85, Led00] condition $\Gamma_2 \geq c\,\Gamma_1$ connects geometric properties with analytic functional inequalities. They are of great interest particularly for diffusion Markov generators, where $\mathbf{A}$ is given by, generally speaking, $\Delta - \nabla V\cdot\nabla$ (we refer to [BGL14] and reference therein concerning diffusion Markov generators). In this paper, we do not assume any particular structure of the Markov generator, but aim at showing general inequalities. Inequalities for square-field operators and geometric curvature bounds provide insights also in spectral properties and exponential decay to equilibrium, which has been further investigated and exploited, e.g. [AM*01, MaV00]. In particular, the Bakry-Émery condition ensures that $\Gamma_2 \geq 0$, which has geometric implications for the underlying manifold and was one of the starting point for the research on square-field operators (see [Bak85]). We also remark that in [Wu00] different functional inequalities have been derived under the condition that higher iterations, namely Γ_3, are positive.

In this paper, our first main result is that for a general Markov generator $\mathbf{A}$, all iterated square-field operators evaluated at the diagonal are positive, i.e. $\Gamma_n(g,g) \geq 0$ whenever they are defined (see Theorem 2.8). To show this, we do not use any restrictive symmetry assumption on $\mathbf{A}$ (like detailed

[1] In theory, there is no reason for a system not to be dissipative.

balance) or assumptions on the spectrum. The proof exploits an iterative interpolation between Γ_n involving the Markov generator and a version involving only the associated semigroup of (bounded) Markov operators. The idea is to start with the semigroup of Markov operators, derive inequalities for them and transfer them afterwards to the corresponding Markov generators. The crucial observation is the interplay of positivity and convexity by a parallelogram identity as an inherent feature of Banach lattices (see Lemma 2.6). For Markov operators we may use classical inequalities and circumvent technical difficulties that occur for unbounded operators. However, technical difficulties arise because differentiation does not preserve positivity.

To show the main idea for Γ_1 (basically as in [Led00]), we aim at showing that $\Gamma_1(g,g) = \mathbf{A}g^2 - 2g \cdot \mathbf{A}g \geq 0$ for all $g \in \mathcal{D}(\mathbf{A})$ with $g^2 \in \mathcal{D}(\mathbf{A})$. From Jensen's inequality we know that for all Markov operator $\mathbf{M}$ and all $g \in \mathcal{D}(\mathbf{A})$ we have $\mathbf{M}g^2 - (\mathbf{M}g)^2 \geq 0$. Hence, $\mathbf{T}(t)g^2 - (\mathbf{T}(t)g)^2 \geq 0$. Differentiating in time at $t = 0$, we obtain for $g \in \mathcal{D}(\mathbf{A})$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{T}(t)g^2 - (\mathbf{T}(t)g)^2\right)|_{t=0} = \mathbf{T}'(t)g^2 - 2\mathbf{T}(t)g \cdot \mathbf{T}'(t)g|_{t=0} = \mathbf{A}g^2 - 2g \cdot \mathbf{A}g = \Gamma_1(g,g).$$

Although, differentiating generally does not preserve inequalities, the reasoning holds true because we have the decomposition

$$\mathbf{T}(t)g^2 - (\mathbf{T}(t)g)^2 = (\mathbf{T}(t) - \mathbf{I})g^2 - g \cdot (\mathbf{T}(t) - \mathbf{I})g - \mathbf{T}(t)g \cdot (\mathbf{T}(t) - \mathbf{I})g, \tag{1.3}$$

where $\mathbf{I}$ is the identity operator. Then we have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{T}(t)g^2 - (\mathbf{T}(t)g)^2\right)|_{t=0} = \lim_{t\to 0}\frac{1}{t}\left(\mathbf{T}(t)g^2 - (\mathbf{T}(t)g)^2\right) - \left(\mathbf{T}(0)g^2 - (\mathbf{T}(0)g)^2\right) \\
&\quad = \lim_{t\to 0}\frac{1}{t}\left(\mathbf{T}(t)g^2 - (\mathbf{T}(t)g)^2\right) = \lim_{t\to 0}\frac{1}{t}\left((\mathbf{T}(t){-}\mathbf{I})g^2 - g \cdot (\mathbf{T}(t){-}\mathbf{I})g - \mathbf{T}(t)g \cdot (\mathbf{T}(t){-}\mathbf{I})g\right) \\
&\quad = \mathbf{A}g^2 - g \cdot \mathbf{A}g - g \cdot \mathbf{A}g = \Gamma(g,g),
\end{aligned}$$

which proves the desired estimate. This idea can be used also for higher iterations Γ_n, $n \geq 1$, and, moreover, a whole family of new inequalities involving $\mathbf{A}$ and its semigroup can be obtained as a byproduct.

Positivity of Γ_n implies that the corresponding energies[2]

$$\mathrm{E}_n(g) := \langle \Gamma_n(g), \mu \rangle, \quad \mathfrak{e}_n(t) := \mathrm{E}_n(g(t))$$

decay along solutions $\dot{g} = \mathbf{A}g$ and, moreover, are convex (see Theorem 2.12). Here μ is a stationary measure, i.e. $\mathbf{A}^*\mu = 0$. This fact is well known for E_0 and also studied for E_1, which corresponds to the Dirichlet form of $\mathbf{A}$. In the present paper we show that all iterations $\mathfrak{e}_n$ monotonically decay to zero. This shows in fact, that dissipativity is not a special physical property but instead a mathematical consequence of a the natural physical property of positivity.

In the second part of the paper, we quantify the convergence rate of $t \mapsto \mathfrak{e}_n(t)$ towards zero. Usually exponential decay is derived assuming a Poincaré-type inequality between, say, $\mathrm{E}_n(g)$ and $\mathrm{E}_{n+1}(g)$. It is clear, that exponential decay can only be expected whenever a spectral gap for the operator $\mathbf{A}$ is present. In particular, for general Markov generators no exponential decay is expected. To obtain nevertheless quantitative convergence for general Markov generators without assumptions on their spectrum, we adapt the operator-theoretic notion of *normality* for the lifted version of $\mathbf{A}$ in the natural Hilbert space $\mathrm{L}^2(\mu)$. This enables to prove that the sequence $(\mathrm{E}_n(g))_{n\in\mathbb{N}}$ is log-convex implying

[2]Note that we call them all *energies* and neglect the physical interpretation.

many interesting features. First, polynomial decay of the energies $t \mapsto \mathfrak{e}_n(t)$ can be derived (Theorem 3.4). Secondly, log-convexity in time and also an upper bound for powers of energies can be shown (see Theorem 3.7). We finally note that normality is much general than symmetry, which would translate to detailed balance or reversibility for stochastic processes. We refer to Section 3.5 where normality is discussed.

2 Positivity of quadratic operators in $\mathcal{C}(\mathcal{Z})$

Before defining the square-field operators and proving their positivity, we first briefly present the mathematical setting.

2.1 Markov semigroups and their generators

Let us start with a few well known facts from the theory of Markov generators and their semigroups (see e.g. [AG*86, Paz83] for more details). In what follows, let $\mathcal{Z}$ be a compact (if necessary, suitably compactified) and metrizable topological space, i.e. a compact, first-countable, Hausdorff space. Let $\mathcal{C}(\mathcal{Z})$ be the space of continuous real-valued functions on $\mathcal{Z}$ and $\mathcal{C}^*(\mathcal{Z})$ (the dual of $\mathcal{C}(\mathcal{Z})$) the space of Radon measures on Borel sets generated by the open sets of $\mathcal{Z}$. The space $\mathcal{C}(\mathcal{Z})$ is a Banach algebra by the pointwise multiplication (denoted by $\cdot$); the constant function $\mathbb{1}(z) \equiv 1$ is denoted by $\mathbb{1} \in \mathcal{C}(\mathcal{Z})$. The dual pairing is denoted by $\langle g, p\rangle = \int_{\mathcal{Z}} g(z)p(dz)$ with $g \in \mathcal{C}(\mathcal{Z})$ and $p \in \mathcal{C}^*(\mathcal{Z})$.

The spaces $\mathcal{C}(\mathcal{Z})$ and $\mathcal{C}^*(\mathcal{Z})$ are Banach lattices with the order relations $\mathcal{C}(\mathcal{Z}) \ni g \geq 0$ if and only if for all $z \in \mathcal{Z}$ we have $g(z) \geq 0$, and $\mathcal{C}^*(\mathcal{Z}) \ni p \geq 0$ if and only if for all Borel sets $B \subset \mathcal{Z}$ we have $p(B) \geq 0$. The order relation in $\mathcal{C}^*(\mathcal{Z})$ as a space of measures coincides with the order relation in dual spaces, i.e. $p \geq 0$ if and only if $\langle g, p\rangle \geq 0$ for all $0 \leq g \in \mathcal{C}(\mathcal{Z})$. In the following, elements g or p with $g \geq 0$ and $p \geq 0$ are called *positive*[3]. The convex subset $\mathcal{P}(\mathcal{Z}) = \left\{p \in \mathcal{C}^*(\mathcal{Z}) \,\middle|\, p \geq 0,\ \langle \mathbb{1}, p\rangle = \|p\| = 1\right\}$ is the set of probability measures describing the statistical states of the system.

As usual, by $\mathcal{L}(\mathcal{C}(\mathcal{Z}))$ and $\mathcal{L}(\mathcal{C}(\mathcal{Z})^*)$ we denote the spaces of linear bounded operators. A linear operator $\mathbf{T} \in \mathcal{L}(\mathcal{C})$ on a Banach lattice is called *positive* (written $\mathbf{T} \geq 0$) if it conserves positivity, i.e., $g \geq 0$ implies $\mathbf{T}g \geq 0$. A linear, bounded operator $\mathbf{M} \in \mathcal{L}(\mathcal{C}(\mathcal{Z}))$ with $\mathbf{M}\mathbb{1} = \mathbb{1}$ and $\mathbf{M} \geq 0$ is called *Markov operator*. The set of Markov operators $\mathcal{M} = \left\{\mathbf{M} \in \mathcal{L}(\mathcal{C}) \,\middle|\, \mathbf{M} \geq 0,\ \mathbf{M}\mathbb{1} = \mathbb{1}\right\}$ is convex and constitute a (noncommutative) semigroup with unit $\mathbf{I}$, where $\mathbf{I}$ is the identity operator on $\mathcal{C}(\mathcal{Z})$. A Markov operator is contractive, because we have $\|\mathbf{M}\| = 1$. Important for us will be Jensen's inequality, which says that for all convex function $\Phi : \mathbb{R} \to [-\infty, \infty]$, all Markov operators $\mathbf{M} \in \mathcal{M}$ and all $g \in \mathcal{C}(\mathcal{Z})$ we have

$$\mathbf{M}\Phi(g) \geq \Phi(\mathbf{M}g), \tag{2.1}$$

where $\Phi(g)$ is defined pointwise, i.e. $\Phi(g)(z) = \Phi(g(z))$. The simple proof for that, based on the fact that Markov operators provide convex combinations, is contained in [Ste05].

Adjoints of Markov operators $\mathbf{M}^*$ are also positive and contractive with $\|\mathbf{M}^*\| = 1$. Given an operator $\mathbf{T}^* \in \mathcal{L}(\mathcal{C}(\mathcal{Z})^*)$ we have $\mathbf{T}^*\mathcal{P}(\mathcal{Z}) \subset \mathcal{P}(\mathcal{Z})$ if and only if $\mathbf{T} \in \mathcal{M}$. In this sense, we can say that Markov operators and only these provide physical reasonable state changes. For all $\mathbf{M} \in \mathcal{M}$ there

[3]Throughout the paper, *positive* is meant to be non-negative. Moreover, *positivity* of functions is defined pointwise and not "almost everywhere".

is at least one invariant measure $\mu \in \mathcal{P}$ with $\mathbf{M}^*\mu = \mu$ by the Frobenius–Perron–Krein–Rutman Theorem.

A Markov semigroup $(\mathbf{T}(t))_{t\geq 0}$ is strongly-continuous semigroup of Markov operators, i.e. $\mathbf{T}(0) = \mathbf{I}$, $\mathbf{T}(t_1 + t_2) = \mathbf{T}(t_2)\mathbf{T}(t_1)$ for all $t_1, t_2 \geq 0$, and $\mathbf{T}(t)g$ converges strongly to g as $t \to 0$. A Markov semigroup is contractive, since for all $t \geq 0$ we have $\|\mathbf{T}(t)\| = 1$. Moreover, a Markov semigroup is a commuting family of Markov operators, and therefore, there exists an invariant measure $\mu \in \mathcal{P}(\mathcal{Z})$ not depending on t with $\mathbf{T}^*(t)\mu = \mu$ by the Markov–Kakutani Theorem (see e.g. [DuS59]). For a given strongly-continuous semigroup $(\mathbf{T}(t))_{t\geq 0}$ the Markov generator is denoted by $(\mathbf{A}, \mathcal{D}(\mathbf{A}))$, where

$$\mathcal{D}(\mathbf{A}) := \left\{ g \in \mathcal{C}(\mathcal{Z}) : \lim_{t\to 0} \frac{1}{t} \left(\mathbf{T}(t)g - g\right) \text{ exists} \right\}, \quad \mathbf{A}g := \lim_{t\to 0} \frac{1}{t} \left(\mathbf{T}(t)g - g\right), \ g \in \mathcal{D}(\mathbf{A}) .$$

Denoting $\mathcal{D}(\mathbf{A}^n)$ the domain of $\mathbf{A}^n$ and setting $\bigcap_{n\in\mathbb{N}} \mathcal{D}(\mathbf{A}^n) =: \mathcal{D}(\mathbf{A}^\infty)$, we have that $\mathcal{D}(\mathbf{A}^\infty)$ is a core of $(\mathbf{A}, \mathcal{D}(\mathbf{A}))$ and a dense subset of $\mathcal{C}(\mathcal{Z})$ (see, e.g., [Paz83]). In particular, in the following we will sometimes omit denoting the domain of definition for operators $\mathbf{A}^n$, because obviously all identities only hold, if both sides of the relation are well-defined. Since $\mathcal{D}(\mathbf{A}^\infty)$ is dense, we may always consider $f, g \in \mathcal{D}(\mathbf{A}^\infty)$.

Given a differential equation

$$\dot{g} = \mathbf{A}g, \quad g(0) = g_0 \in \mathcal{D}(\mathbf{A}) \subset \mathcal{C}(\mathcal{Z}), \tag{2.2}$$

where $\mathbf{A}$ is a Markov generator with semigroup $(\mathbf{T}(t))_{t\geq 0}$, then the solution $t \mapsto g(t)$ is given by $g(t) = \mathbf{T}(t)g_0$. If μ is an invariant measure of a $(\mathbf{T}(t))_{t\geq 0}$, then $\mu \in \mathcal{D}(\mathbf{A}^*)$ and $\mathbf{A}^*\mu = 0$ and vice versa: If a measure $\mu \in \mathcal{D}(\mathbf{A}^*)$ satisfy $\mathbf{A}^*\mu = 0$ then we have $\mathbf{T}^*(t)\mu = \mu$.

In mathematical physics and stochastics, equations of the type (2.2) describe the time evolution of observations and is called *(Chapman-Kolmogorov) backward equation*. From physical perspective, the so-called *(Chapman-Kolmogorov) forward equation* for the evolution of a probability measures $t \mapsto p(t) \in \mathcal{P}(\mathcal{Z})$ is interesting. It can be formally expressed as $\dot{p}(t) = \mathbf{A}^*p(t)$, $p(0) = p_0$, which is defined in the weak-* sense by

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle g, p(t)\rangle = \langle \mathbf{A}g, p(t)\rangle, \quad g \in \mathcal{D}(\mathbf{A}), \quad p(0) = p_0. \tag{2.3}$$

Although in general $\mathbf{A}^*$ is not a generator (it does not need to be densely defined), the solution of (2.3) can be found by solving the equation (2.2), finding $\mathbf{T}(t)$ and setting $p(t) = \mathbf{T}^*(t)p_0$.

Assuming that p_0 has a density h_0 with respect to μ, $p(t)$ can also be calculated as $p(t) = h(t)\mu$, where $h(t)$ is the solution to an equation of type (1.1) but with a different Markov operator as $\mathbf{A}$. In this form, the equation is commonly used as Fokker–Planck equation [AM*01], Levy–Fokker–Planck equation [Gel08], master equation and others. Due to their universality, we restrict ourselves in this paper to equation (2.2).

2.2 Square-field operators and their interpolation

For a given Markov generator $(\mathbf{A}, \mathcal{D}(\mathbf{A}))$, we define the so-called *square-field operators* by

$$\begin{aligned} \Gamma_0(f, g) &= f \cdot g \\ \Gamma_{n+1}(f, g) &= \mathbf{A}\Gamma_n(f, g) - \Gamma_n(\mathbf{A}f, g) - \Gamma_n(f, \mathbf{A}g). \end{aligned}$$

Of course, to ensure that Γ_n is well-defined, we have to restrict to a subalgebra and to elements of $\mathcal{D}(\mathbf{A}^n)$. To do so, we define $\mathcal{D} := \mathcal{D}(\mathbf{A}^\infty) \cap \big(\mathcal{D}(\mathbf{A}^\infty) \cdot \mathcal{D}(\mathbf{A}^\infty)\big)$ the largest subalgebra[4] contained in $\mathcal{D}(\mathbf{A}^\infty)$. Here, in Section 2.2 and Section 2.3 we always assume that $f, g \in \mathcal{D}$. Note that for the associated energies in Section 2.4, no subalgebra is needed and it suffices to consider functions from $\mathcal{D}(\mathbf{A}^\infty)$.

The first elements are given by

$$\begin{aligned}
&\Gamma_1(f,g) = \mathbf{A}(f\cdot g) - f\cdot \mathbf{A}g - g\cdot \mathbf{A}f\\
&\Gamma_2(f,g) = \mathbf{A}^2(f\cdot g) - 2\mathbf{A}(f\cdot \mathbf{A}g) - 2\mathbf{A}(g\cdot \mathbf{A}f) + g\cdot \mathbf{A}^2 f + 2\mathbf{A}g\cdot \mathbf{A}f + f\cdot \mathbf{A}^2 g.
\end{aligned}$$

Moreover, we define the following square-field type operators which contain bounded Markov operators. For this, consider a sequence of Markov operators $(\mathbf{M}_n)_{n\in\mathbb{N}}$. We define a sequence $G_n = G_n(\mathbf{M}_n, \cdots, \mathbf{M}_1, \cdot, \cdot)$ of quadratic operators on $\mathcal{C}(\mathcal{Z})$ by the recursion formula

$$\begin{aligned}
G_0(f,g) &= f\cdot g\\
G_{n+1}(\mathbf{M}_{n+1},...,\mathbf{M}_1,f,g) &= \mathbf{M}_{n+1}G_n(\mathbf{M}_n,...,\mathbf{M}_1,f,g) - G_n(\mathbf{M}_n,...,\mathbf{M}_1,\mathbf{M}_{n+1}f,\mathbf{M}_{n+1}g)\,.
\end{aligned} \tag{2.4}$$

The first terms of G_n are given by

$$\begin{aligned}
G_1(\mathbf{M}_1,f,g) &= \mathbf{M}_1(f\cdot g) - \mathbf{M}_1 f\cdot \mathbf{M}_1 g\\
G_2(\mathbf{M}_2,\mathbf{M}_1,f,g) &= \mathbf{M}_2\mathbf{M}_1(f\cdot g) - \mathbf{M}_2(\mathbf{M}_1 f\cdot \mathbf{M}_1 g) - \mathbf{M}_1(\mathbf{M}_2 f\cdot \mathbf{M}_2 g)\\
&\quad + (\mathbf{M}_1\mathbf{M}_2 f)\cdot(\mathbf{M}_1\mathbf{M}_2 g).
\end{aligned}$$

Note that G_n has n Markov operators in their argument. Since G_0 is symmetric and linear in both arguments separately the same holds for all G_n by definition. Moreover, if $f, g \in \mathcal{C}(\mathcal{Z})$, then also $G_n(\mathbf{M}_n, ..., \mathbf{M}_1, \ldots, f, g) \in \mathcal{C}(\mathcal{Z})$.

As we will see the operators Γ_n can be obtained from G_n by subsequently substituting $\mathbf{M}_k$ by the semigroup $\mathbf{T}(t)$ of the Markov generator $\mathbf{A}$ and taking the limit $t \to 0$. For this, we introduce operators G_n^k, which interpolate between G_n and Γ_n, and are defined by

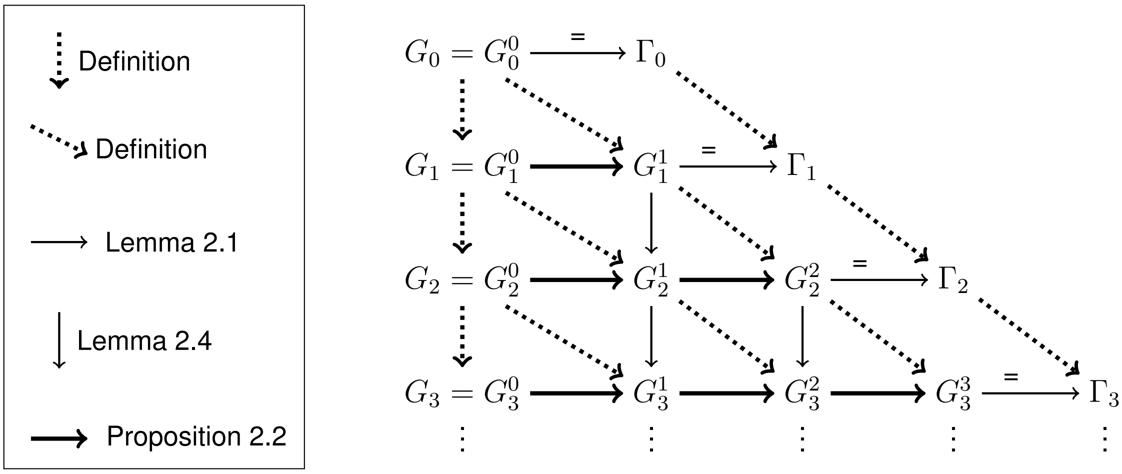
$$\begin{aligned}
G_n^0(\mathbf{M}_n,\ldots,\mathbf{M}_1,f,g) &= G_n(\mathbf{M}_n,\ldots,\mathbf{M}_1,f,g),\\
G_n^k(\mathbf{M}_{n-k},\ldots,\mathbf{M}_1,f,g) &= \mathbf{A}G_{n-1}^{k-1}(\mathbf{M}_{n-k},\ldots,\mathbf{M}_1,f,g)\\
&\quad - G_{n-1}^{k-1}(\mathbf{M}_{n-k},\ldots,\mathbf{M}_1,\mathbf{A}f,g) - G_{n-1}^{k-1}(\mathbf{M}_{n-k},\ldots,\mathbf{M}_1,f,\mathbf{A}g)\,.
\end{aligned}$$

By definition, $G_n^k(f,g)$ is only defined if $G_{n-1}^{k-1}(f,g) \in \mathcal{D}(\mathbf{A})$ and $G_{n-1}^{k-1}(\mathbf{M}_{n-k},\ldots,\mathbf{M}_1,\mathbf{A}f,g)$ as well as $G_{n-1}^{k-1}(\mathbf{M}_{n-k},\ldots,\mathbf{M}_1,f,\mathbf{A}g)$ are well-defined. We note that G_n^k has $n-k$ Markov operators in their arguments and all G_n^k are symmetric.

Next, we are going to show how to obtain Γ_n from G_n to carry over $G_n \geq 0$ to $\Gamma_n \geq 0$. First, in Lemma 2.1, we show that $\Gamma_n = G_n^n$. Secondly, in Proposition 2.2 we show how to obtain G_{n+1}^k from G_n^k. The connection between the operators can be summarized in the following diagram:

Lemma 2.1. *Let $(\mathbf{M}_n)_{n\in\mathbb{N}}$ be a sequence of Markov operators. The operators G_n^k have the following properties:*

[4]To decide whether a subalgebra $\mathcal{D} \subset \mathcal{C}(\mathcal{Z})$ is dense, it suffices to show that it separates points by the Stone-Weierstraß-Theorem, see e.g. [Sem71].



1. *G_n^k is continuous, i.e., if $f_t \to f$, $g_t \to g$ in $\mathcal{C}(\mathcal{Z})$ as $t \to 0$, then $G_n^k(\mathbf{M}_{n-k}, \dots, \mathbf{M}_1, f_t, g_t) \to G_n^k(\mathbf{M}_{n-k}, \dots, \mathbf{M}_1, f, g)$ whenever $G_n^k(\mathbf{M}_{n-k}, \dots, \mathbf{M}_1, f, g)$ is defined.*

2. *For all $n \in \mathbb{N}$, we have $G_n^n(f,g) = \Gamma_n(f,g)$.*

Proof. To simplify notation, we just write $G_n^k = G_n^k(\mathbf{M}_{n-k}, \dots, \mathbf{M}_1, \cdot, \cdot)$ because all $\mathbf{M}_k$ are fixed. For the first claim, we observe that $f_t \to f$ and $g_t \to g$ implies $G_n^0(f_t, g_t) \to G_n^0(f,g)$ since all $\mathbf{M}_j$ are bounded and $\mathcal{C}(\mathcal{Z})$ is a Banach algebra. By induction, we obtain that also $G_n^k(f_t, g_t) \to G_n^k(f,g)$ because $\mathbf{A}$ is a generator and hence a closed operator.

The proof of the second claim is done by induction. By definition, the claim holds for $n = 0$. The recursion formula for G_n^n is given by

$$G_n^n(f,g) = \mathbf{A}G_{n-1}^{n-1}(f,g) - G_{n-1}^{n-1}(\mathbf{A}f, g) - G_{n-1}^{n-1}(f, \mathbf{A}g)\,.$$

This is the same for Γ_n. □

The following proposition relates G_n^k with G_n^{k+1} by replacing $\mathbf{M}_{n-k}$ with $\mathbf{T}(t)$ and taking the rescaled limit $t \to 0$.

Proposition 2.2. *Let a sequence of Markov operators $\left(\mathbf{M}_n\right)_{n\in\mathbb{N}}$ be given such that $\mathbf{M}_n$ commutes with the semigroup $\mathbf{T}(t)$. Then, for all $n \in \mathbb{N}$ and $k \in \{0, \dots, n-1\}$ we have*

$$\lim_{t\to 0} \frac{1}{t}\, G_n^k(\mathbf{T}(t), \mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, f, g) = G_n^{k+1}(\mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, f, g)\,.$$

Before proving the proposition, we immediately observe the following. Replacing subsequently all Markov operators $\mathbf{M}_k$ by the semigroup $\mathbf{T}(t_k)$ we derive the following connection between G_n and Γ_n.

Corollary 2.3. *We have*

$$\lim_{t_1\to 0} \frac{1}{t_1} \left(\lim_{t_2\to 0} \frac{1}{t_2} \left(\dots \lim_{t_n\to 0} \frac{1}{t_n}\, G_n(\mathbf{T}(t_n), \dots, \mathbf{T}(t_1), f, g) \right) \right) = \Gamma_n(f,g)\,.$$

Proof. This follows directly from the definition $G_n^0 = G_n$, Proposition 2.2 and the relation $G_n^n = \Gamma_n$ from Lemma 2.1. □

To prove Proposition 2.2, we need the following lemma, which can be understood as the generalization of (1.3) from the Introduction.

Lemma 2.4. *Let a sequence of Markov operators $(\mathbf{M}_n)_{n\in\mathbb{N}}$ be given such that $\mathbf{M}_{n-k}$ commutes with the semigroup $\mathbf{T}(t)$. For all $n \in \mathbb{N}$ and $k \in \{0, \dots, n-1\}$ we have*

$$G_n^k(\mathbf{M}_{n-k}, \mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, f, g) = \mathbf{M}_{n-k} G_{n-1}^k(\mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, f, g) - G_{n-1}^k(\mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, \mathbf{M}_{n-k} f, \mathbf{M}_{n-k} g).$$

In particular, we have that

$$G_n^k(\mathbf{M}_{n-k}, \dots, f, g) = (\mathbf{M}_{n-k} - \mathbf{I})\, G_{n-1}^k(\dots, f, g) - - G_{n-1}^k(\dots, \mathbf{M}_{n-k} f, (\mathbf{M}_{n-k} - \mathbf{I})\, g) - G_{n-1}^k(\dots, (\mathbf{M}_{n-k} - \mathbf{I})\, f, g)\,. \tag{2.5}$$

Proof. To simplify notation, we again just write $G_n^k(\mathbf{M}, f, g) = G_n^k(\mathbf{M}, \mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, f, g)$ and $G_{n-1}^k(f, g) = G_{n-1}^k(\mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, f, g)$. We are going to show that $G_n^k(\mathbf{M}, f, g) = \mathbf{M} G_{n-1}^k(f, g) - G_{n-1}^k(\mathbf{M} f, \mathbf{M} g)$; the second formula follows directly by linearity of G_n^k.

For $k = 0$ and all $n \in \mathbb{N}$ this follows from the recursion formula of $G_n^0 = G_n$. Assume that it holds for a fixed $k \geq 0$ and all $n \in \mathbb{N}$. We want to show that also

$$G_{n+1}^{k+1}(\mathbf{M}, f, g) = \mathbf{M} G_n^{k+1}(f, g) - G_n^{k+1}(\mathbf{M} f, \mathbf{M} g)\,.$$

By definition, the left-hand side is – assuming that $\mathbf{M}$ and $\mathbf{A}$ commute – given by

$$\begin{aligned} G_{n+1}^{k+1}(\mathbf{M}, f, g) &= \mathbf{A} G_n^k(\mathbf{M}, f, g) - G_n^k(\mathbf{M}, \mathbf{A} f, g) - G_n^k(\mathbf{M}, f, \mathbf{A} g) \\ &= \mathbf{A}\left(\mathbf{M} G_{n-1}^k(f, g) - G_{n-1}^k(\mathbf{M} f, \mathbf{M} g)\right) - \mathbf{M} G_{n-1}^k(\mathbf{A} f, g) - G_{n-1}^k(\mathbf{M}\mathbf{A} f, \mathbf{M} g) \\ &\quad - \mathbf{M} G_{n-1}^k(f, \mathbf{A} g) - G_{n-1}^k(\mathbf{M} f, \mathbf{M}\mathbf{A} g) \\ &= \mathbf{M}\left\{\mathbf{A} G_{n-1}^k(f, g) - G_{n-1}^k(\mathbf{A} f, g) - G_{n-1}^k(f, \mathbf{A} g)\right\} \\ &\quad - \mathbf{A} G_{n-1}^k(\mathbf{M} f, \mathbf{M} g) + G_{n-1}^k(\mathbf{A}\mathbf{M} f, \mathbf{M} g) + G_{n-1}^k(\mathbf{M} f, \mathbf{A}\mathbf{M} g) \\ &= \mathbf{M} G_n^{k+1}(f, g) - G_n^{k+1}(\mathbf{M} f, \mathbf{M} g)\,, \end{aligned}$$

which is the desired formula. □

Using that lemma, we are able to prove Proposition 2.2.

Proof of Proposition 2.2. Again, we just write $G_n^k(\mathbf{T}(t), f, g) = G_n^k(\mathbf{T}(t), \mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, f, g)$ and $G_{n-1}^k(f, g) = G_{n-1}^k(\mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, f, g)$. With (2.5) from Lemma 2.4 and using the linearity of G_n^k, we have

$$\frac{1}{t} G_n^k(\mathbf{T}(t), f, g) = \frac{1}{t}\,(\mathbf{T}(t) - \mathbf{I})\, G_{n-1}^k(f, g) - G_{n-1}^k\big(\mathbf{T}(t) f, \frac{1}{t}\,(\mathbf{T}(t) - \mathbf{I})\, g\big) - G_{n-1}^k\big(\frac{1}{t}\,(\mathbf{T}(t) - \mathbf{I})\, f, g\big).$$

In the limit $t \to 0$, the first term converges to $\mathbf{A} G_n^k(f, g)$, if $G_n^k(f, g) \in \mathcal{D}(\mathbf{A})$. Moreover, we have

$$G_{n-1}^k(\mathbf{T}(t) f, \frac{1}{t}\,(\mathbf{T}(t) - \mathbf{I})\, g) \to G_{n-1}^k(f, \mathbf{A} g), \quad G_{n-1}^k(\frac{1}{t}\,(\mathbf{T}(t) - \mathbf{I})\, f, g) \to G_{n-1}^k(\mathbf{A} f, g),$$

as $t \to 0$ by continuity (see Lemma 2.1). Hence, we obtain

$$\lim_{t\to 0} \frac{1}{t}\, G_n^k(\mathbf{T}(t), f, g) = \mathbf{A} G_{n-1}^k(f, g) - G_{n-1}^k(\mathbf{A} f, g) - G_{n-1}^k(f, \mathbf{A} g),$$

which is equal to $G_n^{k+1}(f, g)$ by definition. □

Remark 2.5. We remark that similarly, Γ_n can be obtained from G_n by subsequently inserting $\mathbf{M}_k = \mathbf{T}(t)$ and then differentiating with respect to t at $t = 0$. Indeed, one can show that

$$\Gamma_n(f,g) = \frac{\mathrm{d}}{\mathrm{d}t} \cdots \left\{ \frac{\mathrm{d}}{\mathrm{d}t} \left\{ \frac{\mathrm{d}}{\mathrm{d}t}\ G_n(\mathbf{M}_n, \dots, \mathbf{M}_1, f, g)|_{\mathbf{M}_1=\mathbf{T}(t)} \Big|_{t=0} \right\} \Big|_{\mathbf{M}_2=\mathbf{T}(t)} \Big|_{t=0} \right\} \cdots \Big|_{\mathbf{M}_n=\mathbf{T}(t)} \Big|_{t=0} .$$

Similar formulas also hold true for G_n^k, but can not directly by used to derive positivity. Moreover, it is possible to derive the following product rule for Γ_n

$$\frac{\mathrm{d}}{\mathrm{d}t}\Gamma_n(\mathbf{T}(t)f, \mathbf{T}(t)g)\Big|_{t=0} = \mathbf{A}\Gamma_n(f,g) - \Gamma_{n+1}(f,g).$$

2.3 Positivity for Γ_n and G_n^k

So far, we introduced G_n^k as an iterative approximation of Γ_n. Next, we want to show that indeed G_n^k and Γ_n are positive, when evaluated at the diagonal $f = g$. For this, with a slight abuse of notation, we introduce

$$\begin{aligned}
\Gamma_n(g) &:= \Gamma_n(g,g) \\
G_n(\mathbf{M}_n, \mathbf{M}_{n-1}, \dots, \mathbf{M}_1, g) &:= G_n(\mathbf{M}_n, \mathbf{M}_{n-1}, \dots, \mathbf{M}_1, g, g) \\
G_n^k(\mathbf{M}_{n-k}, \mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, g) &:= G_n^k(\mathbf{M}_{n-k}, \mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, g, g).
\end{aligned}$$

There will be now confusion with the notation because in this section, we only consider the operators evaluated at the diagonal. The first operators are given by

$$\begin{aligned}
G_0(g) &= \Gamma_0(g) = g^2 \\
G_1(\mathbf{M}, g) &= \mathbf{M}g^2 - (\mathbf{M}g)^2 \\
G_2(\mathbf{M}_2, \mathbf{M}_1, g) &= \mathbf{M}_2\mathbf{M}_1 g^2 - \mathbf{M}_2(\mathbf{M}_1 g)^2 - \mathbf{M}_1(\mathbf{M}_2 g)^2 + (\mathbf{M}_1\mathbf{M}_2 g)^2. \\
\Gamma_1(g) &= \mathbf{A}(g^2) - 2g \cdot \mathbf{A}g \\
\Gamma_2(g) &= \mathbf{A}^2(g^2) - 4\mathbf{A}(g \cdot \mathbf{A}g) + 2g \cdot \mathbf{A}^2 g + 2\mathbf{A}g \cdot \mathbf{A}g.
\end{aligned}$$

The next lemma shows the estimate for $G_n = G_n^0$ which is based on the interplay between convexity and positivity. With the help of Proposition 2.2 we are then able to transfer the estimates to G_n^k and to Γ_n.

Lemma 2.6. *Let a sequence of Markov operators $(\mathbf{M}_n)_{n\in\mathbb{N}}$ be given. With the above notation, we have:*

1 *For all $n \in \mathbb{N}$, the function $g \mapsto G_n(\mathbf{M}_n, \mathbf{M}_{n-1}, \dots, \mathbf{M}_1, g)$ satisfies the parallelogram identity, i.e. for all $f, g \in \mathcal{C}(\mathcal{Z})$ we have*

$$\begin{aligned}
&G_n(\mathbf{M}_n, \mathbf{M}_{n-1}, \dots, \mathbf{M}_1, g) + G_n(\mathbf{M}_n, \mathbf{M}_{n-1}, \dots, \mathbf{M}_1, f) = \\
&\qquad 2G_n\left(\mathbf{M}_n, \mathbf{M}_{n-1}, \dots, \mathbf{M}_1, \frac{f+g}{2}\right) + 2G_n\left(\mathbf{M}_n, \mathbf{M}_{n-1}, \dots, \mathbf{M}_1, \frac{f-g}{2}\right).
\end{aligned}$$

2 *The function $g \mapsto G_n(\mathbf{M}_n, \mathbf{M}_{n-1}, \dots, \mathbf{M}_1, g)$ is convex and non-negative.*

Proof. We prove the claims in several steps, mainly using induction over $n \in \mathbb{N}$. Again, we simply write $G_n(g) = G_n(\mathbf{M}_n, \mathbf{M}_{n-1}, \dots, \mathbf{M}_1, g)$.

For the first part, we observe that the parallelogram identity holds for $n = 0$, since $G_0(g) = g^2$ and we have

$$f^2 + g^2 = 2\left(\frac{f+g}{2}\right)^2 + 2\left(\frac{f-g}{2}\right)^2 .$$

For the step from n to $n+1$, we add the equations

$$\begin{aligned} G_{n+1}(f) &= \mathbf{M}_{n+1}G_n(f) - G_n(\mathbf{M}_{n+1}f) \\ G_{n+1}(g) &= \mathbf{M}_{n+1}G_n(g) - G_n(\mathbf{M}_{n+1}g), \end{aligned}$$

and obtain,

$$\begin{aligned} &G_{n+1}(f) + G_{n+1}(g) = \\ &= \mathbf{M}_{n+1}G_n(f) + \mathbf{M}_{n+1}G_n(g) - \{G_n(\mathbf{M}_{n+1}f) + G_n(\mathbf{M}_{n+1}g)\} \\ &= 2\mathbf{M}_{n+1}\left\{G_n\left(\frac{f+g}{2}\right) + G_n\left(\frac{f-g}{2}\right)\right\} - 2\left\{G_n\left(\mathbf{M}_{n+1}\frac{f+g}{2}\right) + G_n\left(\mathbf{M}_{n+1}\frac{f-g}{2}\right)\right\} \\ &= 2\mathbf{M}_{n+1}G_n\left(\frac{f+g}{2}\right) - 2G_n\left(\mathbf{M}_{n+1}\frac{f+g}{2}\right) + 2\mathbf{M}_{n+1}G_n\left(\frac{f-g}{2}\right) - 2G_n\left(\mathbf{M}_{n+1}\frac{f-g}{2}\right) \\ &= 2G_{n+1}\left(\frac{f+g}{2}\right) + 2G_{n+1}\left(\frac{f-g}{2}\right) , \end{aligned}$$

where we have used the linearity $\frac{1}{2}\left(\mathbf{M}_{n+1}f \pm \mathbf{M}_{n+1}g\right) = \mathbf{M}_{n+1}\left(\frac{f \pm g}{2}\right)$. This proves the claim.

The proof of the second part is done by induction in two steps. First we show that convexity of G_n implies that $G_{n+1} \geq 0$. Secondly, we show that if $G_n \geq 0$ then G_n is convex, by using the parallelogram identity. By induction this would conclude the proof.

Clearly, we have that $G_0(g) = g^2$ is positive and convex. Assuming that $g \mapsto G_n(g)$ is convex, we have for all $\mathbf{M}_{n+1}$ by Jensen's inequality (2.1) that

$$\forall g \in \mathcal{C}(\mathcal{Z}) : \ \mathbf{M}_{n+1}G_n(g) \geq G_n(\mathbf{M}_{n+1}g),$$

which proves the positivity of G_{n+1} because

$$G_{n+1}(g) = \mathbf{M}_{n+1}G_n(g) - G_n(\mathbf{M}_{n+1}g) \geq 0.$$

Next, we show that positivity implies convexity. By the parallelogram identity, we have for all $f, g \in \mathcal{C}(\mathcal{Z})$ that

$$G_n(f) + G_n(g) = 2G_n\left(\frac{f+g}{2}\right) + 2G_n\left(\frac{f-g}{2}\right),$$

which implies

$$\frac{G_n(f) + G_n(g)}{2} - G_n\left(\frac{f+g}{2}\right) = G_n\left(\frac{f-g}{2}\right).$$

Since $G_n \geq 0$, we obtain that G_n is convex. □

Remark 2.7. The key to the proof of positivity is the equivalence of positivity and convexity expressed by the parallelogram identity, which suggests that $G_n(g)$ behaves like a norm in a Hilbert space.

Using Lemma 2.6 and Proposition 2.2, we immediately obtain positivity for $G_n^k(g)$ and $\Gamma_n(g)$.

Theorem 2.8. *Let a sequence of Markov operators $\mathbf{M}_n$ be given such that all $\mathbf{M}_n$ commute with the semigroup $\mathbf{T}(t)$. Then, for all $n \in \mathbb{N}$, $k \in \{0, \dots, n\}$ we have that $G_n^k(\mathbf{M}_{n-k}, \mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, g) \geq 0$. In particular, for all $n \in \mathbb{N}$ we have $\Gamma_n(g) \geq 0$.*

Proof. By Proposition 2.2, we have

$$\lim_{t\to 0} \frac{1}{t} G_n^k(\mathbf{T}(t), \mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, f, g) = G_n^{k+1}(\mathbf{M}_{n-k-1}, \dots, \mathbf{M}_1, f, g) .$$

Since $G_n^0 = G_n$ is positive on the diagonal $f = g$ by Lemma 2.6, we also obtain that $G_n^k(g) \geq 0$ iteratively for all $k \in \{1, \dots, n\}$. In particular, we have $G_n^n(g) = \Gamma_n(g) \geq 0$. □

2.4 Associated energies

By Theorem 2.8 we know that $\Gamma_n(g)$ is positive for all $g \in \mathcal{D}(\mathbf{A}^n)$. Hence, we may define the associated energies. For $\mu \in \mathcal{P}(\mathcal{Z})$ being the stationary probability measure of $\mathbf{A}$, i.e. $\mathbf{A}^*\mu = 0$, we inductively define the bilinear forms

$$\begin{aligned} \mathrm{E}_0(f, g) &= \langle f \cdot g, \mu\rangle - \langle f, \mu\rangle\langle g, \mu\rangle \\ \mathrm{E}_{n+1}(f, g) &= -\mathrm{E}_n(\mathbf{A}f, g) - \mathrm{E}_n(f, \mathbf{A}g) , \end{aligned}$$

whenever the right-hand side is well-defined. In particular, we have

$$\mathrm{E}_1(f, g) = -\langle \mathbf{A}f \cdot g, \mu\rangle - \langle \mathbf{A}g \cdot f, \mu\rangle ,$$

which is usually called the *Dirichlet form* associated with the Markov generator $\mathbf{A}$. Moreover, we introduce the notation $\mathrm{E}_n(g) := \mathrm{E}_n(g, g)$.

Lemma 2.9. *For all $n \geq 1$ we have that*

$$\mathrm{E}_n(f, g) = \langle \Gamma_n(f, g), \mu\rangle.$$

In particular, we have that $\mathrm{E}_n(g) \geq 0$ for all $n \in \mathbb{N}$.

Proof. For $n = 1$, we have

$$\begin{aligned} \mathrm{E}_1(f, g) &= -\mathrm{E}_0(\mathbf{A}f, g) - \mathrm{E}_0(f, \mathbf{A}g) = \\ &= -\langle \mathbf{A}f \cdot g, \mu\rangle + \langle \mathbf{A}f, \mu\rangle\langle g, \mu\rangle - \langle f \cdot \mathbf{A}g, \mu\rangle + \langle f, \mu\rangle\langle \mathbf{A}g, \mu\rangle \\ &= -\langle \mathbf{A}f \cdot g + f \cdot \mathbf{A}g, \mu\rangle = \langle \Gamma_1(f, g), \mu\rangle, \end{aligned}$$

where we have used that $\mathbf{A}^*\mu = 0$. This proves the claim for $n = 1$.

Assuming that the claim holds for $n \geq 1$ and using the recursion formula, we obtain for $n + 1$ that

$$\begin{aligned} \mathrm{E}_{n+1}(f, g) &= -\mathrm{E}_n(\mathbf{A}f, g) - \mathrm{E}_n(f, \mathbf{A}g) = \\ &= -\langle \Gamma_n(\mathbf{A}f, g), \mu\rangle - \langle \Gamma_n(\mathbf{A}f, g), \mu\rangle = \\ &= \langle \mathbf{A}\Gamma_n(f, g) - \Gamma_n(\mathbf{A}f, g) - \langle \Gamma_n(\mathbf{A}f, g), \mu\rangle = \langle \Gamma_{n+1}(f, g), \mu\rangle, \end{aligned}$$

where we again have used $\mathbf{A}^*\mu = 0$. This proves the desired relation.

Since we have $\mathrm{E}_n(g) = \langle \Gamma_n(g), \mu\rangle$, positivity follows for $n \geq 1$ by Theorem 2.8 by using that $\mu \geq 0$. Moreover, $\mathrm{E}_0(g) \geq 0$ by Jensen's inequality. □

One can easily derive an explicit formula for E_n.

Proposition 2.10. *We have the following explicit form for $n \geq 1$*

$$\mathrm{E}_n(f,g) = (-1)^n \sum_{j=0}^{n} \binom{n}{j} \langle \mathbf{A}^{n-j} f \cdot \mathbf{A}^j g, \mu \rangle \,. \tag{2.6}$$

Proof. We observe that for $n = 1$, the right-hand side is given by
$-\left(\langle \mathbf{A} f \cdot g, \mu\rangle + \langle f \cdot \mathbf{A} g, \mu\rangle\right)$, which is E_1. Assuming that the formula holds true for $n \in \mathbb{N}$, we obtain, by using $\binom{n+1}{j} = \binom{n}{j-1} + \binom{n}{j}$ that

$$\begin{aligned}
&(-1)^{n+1} \sum_{j=0}^{n+1} \binom{n+1}{j} \langle \mathbf{A}^{n+1-j} f \cdot \mathbf{A}^j g, \mu \rangle = \\
&= (-1)^{n+1} \left\{ \sum_{j=1}^{n+1} \binom{n+1}{j} \langle \mathbf{A}^{n+1-j} f \cdot \mathbf{A}^j g, \mu \rangle + \langle \mathbf{A}^{n+1} f \cdot g, \mu \rangle \right\} \\
&= (-1)^{n+1} \left\{ \sum_{j=1}^{n+1} \binom{n}{j-1} \langle \mathbf{A}^{n+1-j} f \cdot \mathbf{A}^j g, \mu \rangle + \binom{n}{j} \langle \mathbf{A}^{n+1-j} f \cdot \mathbf{A}^j g, \mu \rangle + \langle \mathbf{A}^{n+1} f \cdot g, \mu \rangle \right\} \\
&= (-1)^{n+1} \left\{ \sum_{k=0}^{n} \binom{n}{k} \langle \mathbf{A}^{n-k} f \cdot \mathbf{A}^{k+1} g, \mu \rangle + \sum_{j=0}^{n} \binom{n}{j} \langle \mathbf{A}^{n+1-j} f \cdot \mathbf{A}^j g, \mu \rangle \right\} \\
&= -\left\{ \mathrm{E}_n(f, \mathbf{A} g) + \mathrm{E}_n(\mathbf{A} f, g) \right\} = \mathrm{E}_{n+1}(f,g),
\end{aligned}$$

which proves the claimed formula. □

Remark 2.11. Similar explicit expression like (2.6) can also be derived for Γ_n and G_n^k.

Considering the solution g of $\dot{g} = \mathbf{A} g$, we may define the energy along solutions given by

$$\mathfrak{e}_n(t) := \mathrm{E}_n(g(t), g(t)) \,.$$

The next theorem shows that all energies $\mathfrak{e}_n$ decay along solutions and are convex in $t > 0$.

Theorem 2.12. *Let $n \in \mathbb{N}$. We have the following properties for $\mathfrak{e}_n$:*

1 *Along solutions $g = g(t)$ of $\dot{g} = \mathbf{A} g$ with $g(0) = g_0 \in \mathcal{D}(\mathbf{A}^\infty)$ the trajectory $t \mapsto \mathfrak{e}_n(t)$ is differentiable and we have*

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathfrak{e}_n(t) = -\mathfrak{e}_{n+1}(t) \,. \tag{2.7}$$

In particular, $t \mapsto \mathfrak{e}_n(t)$ is monotonically decreasing and convex.

2 *For $n \geq 1$ we have $\lim_{t \to \infty} \mathfrak{e}_n(t) = 0$ and $\int_0^\infty \mathfrak{e}_{n+1}(t) \, \mathrm{d}t = \mathfrak{e}_n(0)$.*

Proof. We prove both claims separately. For the first claim, we exploit the explicit expression of E_n. Then, we have for all $n \in \mathbb{N}$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \mathfrak{e}_n(t) &= (-1)^n \sum_{j=0}^{n} \binom{n}{j} \langle \mathbf{A}^{n-j} \dot{g} \cdot \mathbf{A}^j g, \mu \rangle + (-1)^n \sum_{j=0}^{n} \binom{n}{j} \langle \mathbf{A}^{n-j} g \cdot \mathbf{A}^j \dot{g}, \mu \rangle \\
&= (-1)^n \sum_{j=0}^{n} \binom{n}{j} \langle \mathbf{A}^{n-j} \mathbf{A} g \cdot \mathbf{A}^j g, \mu \rangle + (-1)^n \sum_{j=0}^{n} \binom{n}{j} \langle \mathbf{A}^{n-j} g \cdot \mathbf{A}^j \mathbf{A} g, \mu \rangle = \\
&= \mathrm{E}_n(\mathbf{A} g, g) + \mathrm{E}_n(g, \mathbf{A} g) = -\mathrm{E}_{n+1}(g(t)) = -\mathfrak{e}_{n+1}(t).
\end{aligned}$$

Since E_{n+1} is positive $\mathfrak{e}_n$ decays. Moreover, the above formula provides that $\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathfrak{e}_n(t) = \mathfrak{e}_{n+2}(t) \geq 0$, which shows that $\mathfrak{e}_n$ is convex. This proves the first claim.

From (2.7) we have for all $n \in \mathbb{N}$ that

$$\mathfrak{e}_n(0) - \mathfrak{e}_n(t) = \int_0^t \mathfrak{e}_{n+1}(t')\,\mathrm{d}t'$$

Since the left-hand side is monotone and bounded, the limit $\lim_{t\to\infty} \mathfrak{e}_n(t) =: \mathfrak{e}_n(\infty)$ exists and we have

$$\mathfrak{e}_n(0) - \mathfrak{e}_n(\infty) = \int_0^\infty \mathfrak{e}_{n+1}(t')\,\mathrm{d}t' \tag{2.8}$$

This shows the integrability of $t \mapsto \mathfrak{e}_{n+1}(t)$ on $[0,\infty)$. Since $\mathfrak{e}_{n+1}(t)$ is monotone and positive it has to tend to 0. This proves the second part. □

In the next section, we show that an explicit quantitative convergence rate can be derived for $\mathbf{A}$ being a normal operator. For completeness, we recall that $\mathfrak{e}_k$ decays exponentially for $k \in \{1,\ldots,n\}$ under the assumption of a Poincare-type inequality $\mathrm{E}_{n+1}(g) \geq c\,\mathrm{E}_n(g)$, for fixed $n \geq 1$ with some constant $c > 0$.

Corollary 2.13. *Assume, we have for some fixed $n \in \mathbb{N}$ a Poincare-type inequality $\mathrm{E}_{n+1}(g) \geq c\,\mathrm{E}_n(g)$ with some constant $c > 0$. Then, $t \mapsto \mathfrak{e}_k(t)$ decays exponentially in time for all $k \in \{1,\ldots,n\}$ with exponential rate $\exp(-ct)$.*

Proof. By assumption, we get that $\dot{\mathfrak{e}}_n(t) \leq -c\,\mathfrak{e}_n(t)$, which implies $\mathfrak{e}_n(t) \leq \mathfrak{e}_n(0)\mathrm{e}^{-ct}$, the exponential decay of $\mathfrak{e}_n$. Integrating $\dot{\mathfrak{e}}_{n-1}(t) = -\mathfrak{e}_n(t)$ in time, we conclude

$$\mathfrak{e}_{n-1}(T) - \mathfrak{e}_{n-1}(t) = \int_t^T -\mathfrak{e}_n(s)\mathrm{d}s \geq -\int_t^T \mathfrak{e}_n(0)\mathrm{e}^{-cs}\mathrm{d}s = \frac{\mathfrak{e}_n(0)}{c}\left(\mathrm{e}^{-cT} - \mathrm{e}^{-ct}\right) .$$

Taking the limit $T \to \infty$ and using that $\mathfrak{e}_{n-1}(T=\infty) = 0$ by Theorem 3.4, we conclude that

$$\mathfrak{e}_{n-1}(t) \leq \frac{\mathfrak{e}_n(0)}{c}\mathrm{e}^{-ct} .$$

Clearly this can be iterated to show that $t \mapsto \mathfrak{e}_k(t)$ decays exponentially for all $k \in \{1,\ldots,n\}$. □

Remark 2.14. We remark that the presented theory is valid for arbitrary Markov generators. However, there are cases for which the discussed inequalities are trivially satisfied but do not contain any information.

The first obvious case is when parts of the spectrum lie on the imaginary axis (without considering 0) which is true for derivations. A Markov generator $\mathbf{A}$ is called *derivation*, if $\mathcal{D}(\mathbf{A})$ is a subalgebra of $\mathcal{C}$, $\mathbb{1} \in \mathcal{D}(\mathbf{A})$, $\mathbf{A}\mathbb{1} = 0$ and $\mathbf{A}(f\cdot g) = f\cdot\mathbf{A}g + g\cdot\mathbf{A}f,\ f,g \in \mathcal{D}(\mathbf{A})$ holds. Then we have $\Gamma_1(f,g) \equiv 0$ what implies $\Gamma_n(f,g) \equiv 0$ for all $n \geq 1$.

A second case is a degenerated generator $\mathbf{A}$. A Markov generator $\mathbf{A}$ is called *degenerate* if there exists a closed set $B \subset \mathcal{Z}$ and $(\mathbf{A}g)(z) = 0$ holds for all $z \in B$ and $g \in \mathcal{D}(\mathbf{A})$. Obviously, any measure $\mu \in \mathcal{P}$ with $\mu(B) = 1$ is a stationary measure for $\mathbf{A}^*$. Although in this case $\Gamma_n \not\equiv 0$ holds, we obtain $\mathrm{E}_n(f,g) = \langle \Gamma_n(f,g), \mu\rangle = 0$ for all $n \geq 1$.

3 Decay rate for energies in Hilbert space

So far the decay of the energies $t \mapsto \mathfrak{e}_n(t) = \mathrm{E}_n(g(t))$ for a general Markov generator $\mathbf{A}$ along solution is not quantified. It is well-known that whenever the generator $\mathbf{A}$ has a spectral gap, i.e. the largest non-trivial eigenvalue has strictly negative real part, then exponential decay of the solution and, hence, also for energies should hold in theory. For a general Markov generator $\mathbf{A}$ the real-parts of its eigenvalues may accumulate in zero, i.e. there is no spectral gap and no exponential decay. Moreover, in practice it is rather difficult to compute the spectral gap for a given Markov generator.

As we will see it is possible to derive polynomial (n-dependent) decay for all energies $t \mapsto \mathfrak{e}_n(t)$ for general $\mathbf{A}$, which satisfy an operator-theoretic normality condition. Because *normality* (like *self-adjointness*) are terms for operators in Hilbert spaces, we first lift $\mathbf{A}$ to the natural Hilbert space $\mathrm{L}^2(\mu)$. Then in Section 3.3, we derive log-convexity for the sequence $(\mathrm{E}_n(g))_{n\in\mathbb{N}}$ which enables to derive conclusion for $t \mapsto \mathfrak{e}_n(t)$.

3.1 Hilbert space embedding of $\mathbf{A}$

The generator $(\mathbf{A}, \mathcal{D}(\mathbf{A}))$ with semigroup $\mathbf{T}(t)$ with stationary measure $\mu \in \mathcal{P}(\mathcal{Z})$ is defined on $\mathcal{C}(\mathcal{Z})$. Certain important properties, however, are naturally to be studied in a Hilbert space.

As usual we define the real separable Hilbert space $\mathrm{L}^2(\mu)$ as the completion of $\mathcal{C}(\mathcal{Z})$ with the norm (written $\|g\|_\mu$) induced by the scalar product

$$(f, g)_\mu := \langle f \cdot g, \mu \rangle = \int_{\mathcal{Z}} f(z)g(z)p(dz)\,.$$

By definition $\mathcal{C}(\mathcal{Z}) \subset \mathrm{L}^2(\mu)$ is dense[5]. We implicitly use this fact in the following, writing for example $(f, g)_\mu = \langle f \cdot g, \mu\rangle$, although this identity is valid only for $f, g \in \mathcal{C}(\mathcal{Z}) \subset \mathrm{L}^2(\mu)$.

Recall, that if μ is the invariant measure of the Markov operator $\mathbf{M}$, i.e. $\mathbf{M}^*\mu = \mu$, then $\mathbf{M}$ can be boundedly extended to $\mathrm{L}^2(\mu)$. To see this, we observe for $g \in \mathrm{L}^2(\mu)$ that

$$\|\mathbf{M}g\|_\mu^2 = (\mathbf{M}g, \mathbf{M}g)_\mu = \left\langle (\mathbf{M}g)^2, \mu \right\rangle \le \langle \mathbf{M}g^2, \mu\rangle = \langle g^2, \mathbf{M}^*\mu\rangle = \langle g^2, \mu\rangle = (g, g)_\mu,$$

where we have used Jensen's inequality and that $\mathbf{M}^*\mu = \mu$. Hence, $\mathbf{M}$ is bounded on $\mathrm{L}^2(\mu)$ with constant 1. In the following we also denote the operator on the larger space $\mathrm{L}^2(\mu)$ by the same symbol.

Similarly, a strongly-continuous semigroup of operators $(\mathbf{T}(t))_{t\ge 0}$ on $\mathcal{C}(\mathcal{Z})$ with invariant measure $\mu \in \mathcal{P}(\mathcal{Z})$ can be extended to the space $\mathrm{L}^2(\mu)$. Clearly, then the family $(\mathbf{T}(t))_{t\ge 0}$ is a semigroup of bounded operators on $\mathrm{L}^2(\mu)$. Moreover, $(\mathbf{T}(t))_{t\ge 0}$ is also strongly-continuous in $\mathrm{L}^2(\mu)$, because we

[5]The choice of $\mathrm{L}^2(\mu)$ instead of any $\mathrm{L}^2(p)$ with an arbitrary positive measure p, has good reasons. In order to transfer the theory in $\mathcal{C}$ to L^2, certain properties should be preserved as boundedness and strong-continuity of $\mathbf{T}(t)$. We note that in general a bounded operator $\mathbf{M}$ on $\mathcal{C}(\mathcal{Z})$ is not bounded on $\mathrm{L}^2(p)$. Take for example $\mathcal{Z} = [0, 1]$, $\mathbf{M}g(z) := g(z_0)$ and p the Lebesgue measure on $[0, 1]$. The distinguished role of the stationary measure μ becomes clear when considering, for example, symmetry of $\mathbf{A}$, i.e. $(\mathbf{A}f, g)_p = (f, \mathbf{A}g)_p$ or, equivalently, $\langle g \cdot \mathbf{A}f, p\rangle = \langle f \cdot \mathbf{A}g, p\rangle$. Setting $f = \mathbb{1}$ we conclude $\mathbf{A}^*p = 0$, so p must then be a stationary measure of $\mathbf{A}$.

DOI 10.20347/WIAS.PREPRINT.2901

have

$$\begin{aligned}\|\mathbf{T}(t)g-g\|_\mu^2 =&\langle(\mathbf{T}(t)g-g)\cdot(\mathbf{T}(t)g-g),\mu\rangle\\ =&\langle(\mathbf{T}(t)g)^2,\mu\rangle-\langle g^2,\mu\rangle-2\langle g\cdot(\mathbf{T}(t)g-g),\mu\rangle\\ \le&\langle\mathbf{T}(t)g^2,\mu\rangle-\langle g^2,\mu\rangle-2\langle g\cdot(\mathbf{T}(t)g-g),\mu\rangle\\ =&\langle\mathbf{T}(t)g^2-g^2,\mu\rangle-2\langle g\cdot(\mathbf{T}(t)g-g),\mu\rangle.\end{aligned}$$

The right hand side tends to 0 for $t\to 0$ due to the strong continuity of $\mathbf{T}(t)$ in $\mathcal{C}$. The generator of $(\mathbf{T}(t))_{t\ge0}$ on $\mathrm{L}^2(\mu)$ is (with a slight abuse of notation) also denoted by $(\mathbf{A},\mathcal{D}(\mathbf{A}))$, $\mathcal{D}(\mathbf{A})\subset\mathrm{L}^2(\mu)$. It is closed and densely defined on $\mathrm{L}^2(\mu)$. There will be no confusion with the notation because in this section we are only interested in the Hilbert space $\mathrm{L}^2(\mu)$-version. The generator $\mathbf{A}$ coincides with the original generator on $\mathcal{C}(\mathcal{Z})$ because the semigroups coincide there. Since $\mathbf{T}(t)$ is a contraction, the spectrum of $\mathbf{A}$ is located on the left-hand side of the complex plane.

3.2 Normality of A and the connection to E_n

To quantify the decay rate of the energies $\mathfrak{e}_n$, we assume that $\mathbf{A}$ is a normal operator. For this, we first define its $\mathrm{L}^2(\mu)$-adjoint $(\mathbf{A}^\star,\mathcal{D}(\mathbf{A}^\star))$ as usual[6] (we refer e.g. to [Sch12] for unbounded operators on Hilbert spaces). The operator $\mathbf{A}^\star$ is well defined since $\mathbf{A}$ is densely defined, and moreover, it is closed. In the following we make the following assumption.

Assumption. The operator $(\mathbf{A},\mathcal{D}(\mathbf{A}))$ is a normal operator on $\mathrm{L}^2(\mu)$, meaning that $\mathcal{D}(\mathbf{A})=\mathcal{D}(\mathbf{A}^\star)$ and $\|\mathbf{A}g\|_\mu=\|\mathbf{A}^\star g\|_\mu$ for all $g\in\mathcal{D}(\mathbf{A})=\mathcal{D}(\mathbf{A}^\star)$.

Since $\mathbf{A}$ is closed, we have $\mathbf{A}^\star\mathbf{A}=\mathbf{A}\mathbf{A}^\star$. We note that normality is a substantially more general concept than self-adjointness, which would mean that $\mathcal{D}(\mathbf{A})=\mathcal{D}(\mathbf{A}^\star)$ and $\mathbf{A}=\mathbf{A}^\star$. In Section 3.5, we discuss what normality of $\mathbf{A}$ means for the original operator on $\mathcal{C}(\mathcal{Z})$. In particular we provide an operator-theoretic version for normality with respect to $\mathcal{C}(\mathcal{Z})$. Moreover, we show with an example that without the assumption log-convexity may fail, which is our main ingredient to derive polynomial decay of the energy.

For the normal operator $\mathbf{A}$, we define the operator

$$\mathbf{C}:=-\left(\mathbf{A}+\mathbf{A}^\star\right),\qquad \mathcal{D}(\mathbf{C})=\mathcal{D}(\mathbf{A})=\mathcal{D}(\mathbf{A}^\star)\,. \tag{3.1}$$

In particular, normality of $\mathbf{A}$ implies that $\mathcal{D}(\mathbf{C})=\mathcal{D}(\mathbf{A})$ and that $\mathbf{C}$ is densely defined. Moreover, we see that for all $f,g\in\mathcal{D}(\mathbf{C})=\mathcal{D}(\mathbf{A})$, we have

$$(f,\mathbf{C}g)_\mu=-(f,\left(\mathbf{A}+\mathbf{A}^\star\right)g)_\mu=-(f,\mathbf{A}g)_\mu-(\mathbf{A}f,g)_\mu,$$

which shows that $\mathbf{C}$ is symmetric and, hence, also closable. In general $\mathbf{C}$ is, as a sum of a normal operator with its adjoint, not necessarily closed[7]. However classical spectral theory for normal operators provides that the closure $\overline{\mathbf{C}}=-\overline{(\mathbf{A}+\mathbf{A}^\star)}$ is self-adjoint, see e.g. [Sch12]. In fact the complex spectral family for a (in general unbounded) normal operator can be split into two real spectral families,

[6] Note that we use a star ($\star$) to distinguish the Hilbert-space adjoint $\mathbf{A}^\star$ with the Banach space dual $\mathbf{A}^*$, where the latter is defined on $\mathcal{C}^*(\mathcal{Z})$.

[7] As an example consider $\mathbf{A}=\frac{\mathrm{d}}{\mathrm{d}x}$ on $\mathrm{L}^2(\mathbb{R})$. Then $\mathbf{A}$ is normal on $\mathcal{D}(\mathbf{A})=\mathrm{H}^1(\mathbb{R})$ because $\mathbf{A}^\star=-\frac{\mathrm{d}}{\mathrm{d}x}=-\mathbf{A}$. But $\mathbf{A}+\mathbf{A}^\star=\mathbf{0}$, which is not closed as defined on $\mathrm{H}^1(\mathbb{R})$. See e.g. [ArT20] for recent results on the variety of domain intersection for an operator with its adjoint

which provides a decomposition into a real and imaginary part of the normal operator that are essentially self-adjoint. Here, we are not interested in spectral properties of $\mathbf{A}$, and, in particular, we do not make any assumptions on the spectrum of $\mathbf{A}$, we only use the fact that $\mathbf{C}$ is essentially self-adjoint.

Moreover, $\mathbf{C}$ as well as its closure $\overline{\mathbf{C}}$ is positive in the form sense, because

$$(g, \mathbf{C}g)_\mu = -(g, \mathbf{A}g)_\mu - (g, \mathbf{A}g)_\mu = -2\langle g \cdot \mathbf{A}g, \mu\rangle = \mathrm{E}_1(g, g) \geq 0\,,$$

holds for all $g \in \mathcal{D}(\mathbf{C})$. Summarizing, $\mathbf{C}$ is a positive essentially self-adjoint operator. In particular, there is a unique positive self-adjoint square-root $\mathbf{B}$ of $\overline{\mathbf{C}}$ (see e.g. [Kat95]), which will be used later.

As it turns out powers of the operator $\mathbf{C}$ are directly related to the energies $\mathrm{E}_n(f, g)$. Since $\mathbf{A}$ is normal, we immediately obtain that the operators $\mathbf{C}$ and $\mathbf{A}$ commute on $\mathcal{D}(\mathbf{A}^2)$. Hence, also higher powers of $\mathbf{C}$ commute with $\mathbf{A}$. This is used to express E_n in terms of powers of $\mathbf{C}$. We remark that an analogous statement is well known and commonly used for symmetric operators (see, e.g., (1.2) in the Introduction)

Proposition 3.1. *Let $(\mathbf{A}, \mathcal{D}(\mathbf{A}))$ be normal and $\mathbf{C}$ be defined by* (3.1)*. Then for $n \geq 1$ we have*

$$\mathrm{E}_n(f, g) = (f, \mathbf{C}^n g)_\mu\,,$$

whenever the right-hand side is bounded (i.e. for all $f, g \in \mathcal{D}(\mathbf{A}^n)$).

Proof. We prove that the sequence, $\left\{(f, \mathbf{C}^n g)_\mu\right\}_{n\in\mathbb{N}}$ satisfies the same recursion formula as $\mathrm{E}_n(f, g)$. For $n = 1$, we have already seen that $(f, \mathbf{C}g)_\mu = -(f, \mathbf{A}g)_\mu - (\mathbf{A}f, g)_\mu = \mathrm{E}_1(f, g)$.

For $n \geq 2$, we observe that for $f, g \in \mathcal{D}(\mathbf{A}^{n+1})$:

$$\begin{aligned}\left(f, \mathbf{C}^{n+1} g\right)_\mu &= (f, \mathbf{C}\mathbf{C}^n g)_\mu = -(f, \mathbf{A}\mathbf{C}^n g)_\mu - (\mathbf{A}f, \mathbf{C}^n g)_\mu = -(f, \mathbf{C}^n \mathbf{A}g)_\mu - (\mathbf{A}f, \mathbf{C}^n g)_\mu \\ &= -\mathrm{E}_n(f, \mathbf{A}g) - \mathrm{E}_n(\mathbf{A}f, g) = \mathrm{E}_{n+1}(f, g)\,,\end{aligned}$$

where we have used that $\mathbf{A}$ and $\mathbf{C}^n$ commute on $\mathcal{D}(\mathbf{A}^{n+1})$. □

Remark 3.2. We remark that the above formula makes also sense to define fractional powers of the energies E_α via $\overline{\mathbf{C}}^\alpha$ for $\alpha > 0$ by interpolation. This will be investigated in subsequent work.

3.3 Log-convexity in n and polynomial decay

With the explicit and closed form of the energies $\mathrm{E}_n(f, g) = (f, \mathbf{C}^n g)_\mu$ we are able to show that $t \mapsto \mathfrak{e}_n(t)$ decays with polynomial rate. To see this we show that the sequence $(\mathrm{E}_n(g))_{n\in\mathbb{N}}$ is log-convex for any $g \in \mathcal{D}(\mathbf{A}^\infty)$.

Proposition 3.3. *Let $(\mathbf{A}, \mathcal{D}(\mathbf{A}))$ be a normal operator and let $g \in \mathcal{D}(\mathbf{A}^\infty)$. Then:*

1 *The sequence $(\mathrm{E}_n(g))_{n\in\mathbb{N}}$ is log-convex, i.e. for all $n \in \mathbb{N}$ we have $\mathrm{E}_{n+1}(g)\mathrm{E}_{n-1}(g) \geq \mathrm{E}_n(g)^2$.*

2 *If $\mathrm{E}_0(g) > 0$, then we have $\left(\frac{\mathrm{E}_{n+1}(g)}{\mathrm{E}_0(g)}\right)^{\frac{1}{n+1}} \geq \left(\frac{\mathrm{E}_n(g)}{\mathrm{E}_0(g)}\right)^{\frac{1}{n}}$.*

Proof. Since $\mathbf{A}$ is normal, the operator $(\mathbf{C},\mathcal{D}(\mathbf{C}))$ defined by (3.1) is positive and essentially self-adjoint. Let $\mathbf{B}$ be the positive and self-adjoint square-root of $\overline{\mathbf{C}}$ on $\mathrm{L}^2(\mu)$, i.e. we have $\mathbf{B}^2 g=\overline{\mathbf{C}}g$ for all $g\in\mathcal{D}(\overline{\mathbf{C}})$. Then, for $\alpha,\beta\in\mathbb{N}$ and $g\in\mathcal{D}(\mathbf{B}^{2\max\{\alpha,\beta\}})$ we have that

$$\begin{aligned}0&\le\int_{\mathcal{Z}}\int_{\mathcal{Z}}\Big[\big(\mathbf{B}^\alpha g\big)(z)\big(\mathbf{B}^\beta g\big)(z')-\big(\mathbf{B}^\alpha g\big)(z')\big(\mathbf{B}^\beta g\big)(z)\Big]^2\mu(\mathrm{d}z)\mu(\mathrm{d}z')\\&=\Big[\int\big(\mathbf{B}^\alpha g\big)^2(z)\mu(\mathrm{d}z)\cdot\int\big(\mathbf{B}^\beta g\big)^2(z')\mu(\mathrm{d}z')+\int\big(\mathbf{B}^\alpha g\big)^2(z')\mu(\mathrm{d}z')\cdot\int\big(\mathbf{B}^\beta g\big)^2(z)\mu(\mathrm{d}z)\\&\qquad-2\int\big(\mathbf{B}^\alpha g\big)(z)\big(\mathbf{B}^\beta g\big)(z)\mu(\mathrm{d}z)\cdot\int\big(\mathbf{B}^\alpha g\big)(z')\big(\mathbf{B}^\beta g\big)(z')\mu(\mathrm{d}z')\Big]\\&=(\mathbf{B}^\alpha g,\mathbf{B}^\alpha g)_\mu(\mathbf{B}^\beta g,\mathbf{B}^\beta g)_\mu-(\mathbf{B}^\alpha g,\mathbf{B}^\beta g)_\mu^2=(g,\mathbf{B}^{2\alpha}g)_\mu(g,\mathbf{B}^{2\beta}g)_\mu-(g,\mathbf{B}^{\alpha+\beta}g)_\mu^2\,.\end{aligned}$$

Setting, $\alpha=n+1$ and $\beta=n-1$ and using $\mathbf{B}^2g=\overline{\mathbf{C}}g=\mathbf{C}g$ for $g\in\mathcal{D}(\mathbf{C}^{n+1})$, we obtain

$$\begin{aligned}0&\le(g,\mathbf{B}^{2(n+1)}g)_\mu(g,\mathbf{B}^{2(n-1)}g)_\mu-(g,\mathbf{B}^{2n}g)_\mu^2\\&=(g,\mathbf{C}^{n+1}g)_\mu(g,\mathbf{C}^{n-1}g)_\mu-(g,\mathbf{C}^n g)_\mu^2=\mathrm{E}_{n+1}(g)\mathrm{E}_{n-1}(g)-\mathrm{E}_n(g)^2\,,\end{aligned}$$

which proves the first claim.

For the second claim, we use the fact that for a given non-negative log-convex sequence $(a_n)_{n\in\mathbb{N}}$ with $a_0>0$, we have $\left(\frac{a_{n+1}}{a_0}\right)^{\frac{1}{n+1}}\ge\left(\frac{a_n}{a_0}\right)^{\frac{1}{n}}$ for all $n\ge1$. A proof[8] for this can be found for example in [BeB61], which is given here for completeness.

Assuming that $(a_n)_{n\in\mathbb{N}}$ is a non-negative log-convex sequence with $a_0>0$, we have

$$a_1^2\le a_0a_2,\quad a_2^2\le a_1a_3,\quad a_3^2\le a_2a_4,\quad\ldots,\quad a_n^2\le a_{n-1}a_{n+1}\,.$$

If we take the j-th inequalities to the j-th power and multiply all inequalities, we get

$$a_1^2a_2^4a_3^6\cdots a_{n-1}^{2n-2}a_n^{2n}\le(a_0a_2)^1(a_1a_3)^2(a_2a_4)^3\cdots(a_{n-2}a_n)^{n-1}(a_{n-1}a_{n+1})^n.$$

This can be simplified to $a_n^{2n}\le a_0a_n^{n-1}a_{n+1}^n$, which implies $(a_n/a_0)^{\frac1n}\le(a_{n+1}/a_0)^{\frac{1}{n+1}}$.

□

With the help of the log-convexity of the sequence $(\mathrm{E}_n(g))_{n\in\mathbb{N}}$, we can derive asymptotic polynomial decay for $t\mapsto\mathfrak{e}_n(t)$ as an upper bound.

Theorem 3.4. *Let g be a solution of $\dot g=\mathbf{A}g$ with $g(0)=g_0\in\mathcal{D}(\mathbf{A}^\infty)$ such that $\mathfrak{e}_0(0)>0$. Moreover, let the generator $\mathbf{A}$ be normal. If $\mathfrak{e}_n(0)>0$, then we have*

$$\mathfrak{e}_n(t)\le\left(\mathfrak{e}_n(0)^{-1/n}+\frac{t}{n}\mathfrak{e}_0(0)^{-1/n}\right)^{-n}.$$

In particular, $t\mapsto\mathfrak{e}_n(t)$ decays to zero with polynomial rate $O(t^{-n})$.

Proof. Using that $g=g(t)$ is solution, we have that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{e}_n(g)=\dot{\mathfrak{e}}_n(t)=-\mathfrak{e}_{n+1}(t)\,.$$

[8] The historic proof is from 1729 and goes back to C. Maclaurin.

By Proposition 3.3 we have $\mathrm{E}_{n+1}(g) \geq \mathrm{E}_0(g)^{-1/n}\mathrm{E}_n(g)^{\frac{n+1}{n}}$ for all $n \geq 1$. Using $\mathfrak{e}_0(0) \geq \mathfrak{e}_0(t)$, we obtain that

$$\dot{\mathfrak{e}}_n(t) \leq -\mathfrak{e}_0(t)^{-1/n}\mathfrak{e}_n(t)^{\frac{n+1}{n}} \leq -\mathfrak{e}_0(0)^{-1/n}\mathfrak{e}_n(t)^{\frac{n+1}{n}} . \tag{3.2}$$

From this differential inequality we conclude that $\mathfrak{e}_n(t)$ satisfies the desired estimate. Indeed, we may assume that $\mathfrak{e}_n(t) > 0$ for all $t > 0$, otherwise the claim is trivial. Introducing

$$\alpha := \mathfrak{e}_0(0)^{-1/n}, \quad X(\mathfrak{e}) := \frac{n}{\alpha}\mathfrak{e}^{-1/n}, \quad x(t) := X(\mathfrak{e}_n(t)) - X(\mathfrak{e}_n(0)) - t,$$

we get $x(0) = 0$ and

$$\dot{x}(t) = X'(\mathfrak{e}_n(t))\dot{\mathfrak{e}}_n(t) - 1 = -\frac{1}{\alpha}\mathfrak{e}_n(t)^{-1-1/n}\dot{\mathfrak{e}_n}(t) - 1 \geq 0,$$

by (3.2). Hence, we have $x(t) \geq 0$ for $t \geq 0$ which means $X(\mathfrak{e}_n(t)) \geq X(\mathfrak{e}_n(0)) + t$. Inserting the definition of X, we obtain

$$\frac{n}{\alpha}\mathfrak{e}_n(t)^{-1/n} \geq \frac{n}{\alpha}\mathfrak{e}_n(0)^{-1/n} + t \quad \Rightarrow \quad \mathfrak{e}_n(t) \leq \left(\mathfrak{e}_n(0)^{-1/n} + \mathfrak{e}_0(0)^{-1/n}\frac{t}{n}\right)^{-n},$$

which we wanted to show. □

Remark 3.5. Crucial for Theorem 3.4 was the log-convexity from Proposition 3.3 of the form $\frac{\mathrm{E}_{n+1}(g)}{\mathrm{E}_0(g)} \geq \left(\frac{\mathrm{E}_n(g)}{\mathrm{E}_0(g)}\right)^{\frac{n+1}{n}}$. Heuristically, the exponent on the right-hand side converges to 1 as $n \to \infty$, meaning that the above estimate becomes more and more a linear inequality with increasing $n \in \mathbb{N}$. This means that at least in theory it becomes easier to prove a linear inequality for larger $n \in \mathbb{N}$, and to proceed as in Corollary 2.13 to obtain exponential decay.

Remark 3.6. We remark, that for a general Markov generator $\mathbf{A}$ no exponential decay is to be expected, because, in principle, the real parts of the spectrum of $\mathbf{A}$ may have a clustering point in 0, or, in other words, no spectral gap is present.

Although it is not easy to find examples for this and to make the above polynomial estimates explicit, we provide a well-known example, namely diffusion on the real line. However, we note that the example does not fit exactly into the presented theory since it considers a non-compact domain. Considering the diffusion equation $u_t = \frac{D}{2}u_{xx}$ on $\mathbb{R}$, the fundamental solution is given by the Gaussian kernel $u = \frac{1}{\sqrt{2\pi Dt}}\mathrm{e}^{-\frac{x^2}{2Dt}}$, where the stationary measure is the Lebesgue measure $\mu = \mathrm{d}x$. It is well-known that whole negative real axis belongs to the continuous spectrum. It is easy to calculate that

$$\begin{aligned}
&\mathfrak{e}_0(t) = 1, && \mathfrak{e}_1(t) = -2(u, \mathbf{A}u)_\mu = \frac{1}{4\sqrt{\pi D t^3}}, \\
&\mathfrak{e}_2(t) = 4(u, \mathbf{A}^2 u))_\mu = \frac{3}{8\sqrt{\pi D t^5}}, && \mathfrak{e}_3(t) = -8(u, \mathbf{A}^3 u))_\mu = \frac{3 \cdot 5}{16\sqrt{\pi D t^7}}, \\
&\mathfrak{e}_4(t) = 16(u, \mathbf{A}^4 u))_\mu = \frac{3 \cdot 5 \cdot 7}{32\sqrt{\pi D t^9}}, && \mathfrak{e}_5(t) = -32(u, \mathbf{A}^5 u))_\mu = \frac{3 \cdot 5 \cdot 7 \cdot 9}{64\sqrt{\pi D t^{11}}} .
\end{aligned}$$

3.4 Log-convexity in time

The log-convexity of the sequence $(\mathrm{E}_n(g))_n \in \mathbb{N}$ provides also log-convexity of the trajectory of energies $t \mapsto \mathfrak{e}_n(t)$, which is a stronger statement than the convexity of $t \mapsto \mathfrak{e}_n(t)$ from Theorem

2.12. Moreover, log-convexity provides another interesting feature, namely that $t \mapsto \big(\mathfrak{e}_n(t)/\mathfrak{e}_n(0)\big)^{1/t}$ is increasing in time. This can be understood an a complementary information to the asymptotic decay of $t \mapsto \big(\mathfrak{e}_n(t)/\mathfrak{e}_n(0)\big)$ as in Theorem 3.4.

Theorem 3.7. *Let g be a solution of $\dot{g} = \mathbf{A}g$ with $g(0) = g_0 \in \mathcal{D}(\mathbf{A}^\infty)$ such that for fixed $n \in \mathbb{N}$ we have $\mathfrak{e}_n(t) > 0$ for all $t > 0.$ Moreover, let $\mathbf{A}$ be normal. Then we have the following:*

1. *The energy trajectory $t \mapsto \mathfrak{e}_n(t)$ is log-convex in time, i.e., $\frac{\mathrm{d}^2}{\mathrm{d}t^2} \log \mathfrak{e}_n(t) \geq 0.$*
2. *The trajectory $t \mapsto \Big(\frac{\mathfrak{e}_n(t)}{\mathfrak{e}_n 0}\Big)^{1/t}$ is increasing in time, and, moreover, it converges.*

Proof. By log-convexity with respect to n we have

$$\ddot{\mathfrak{e}}_n(t) = -\dot{\mathfrak{e}}_{n+1}(t) = \mathfrak{e}_{n+2}(t) \geq \mathfrak{e}_n^{-1}(t)\mathfrak{e}_{n+1}^2(t) = \mathfrak{e}_n^{-1}(t)\dot{\mathfrak{e}}_n^2(t),$$

where we have used that $\mathfrak{e}_n(t)$ is strictly positive for all $t > 0$. Hence, we conclude

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \log \mathfrak{e}_n(t) = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\dot{\mathfrak{e}}_n(t)}{\mathfrak{e}_n(t)} = \frac{\mathfrak{e}_n(t)\ddot{\mathfrak{e}}_n(t) - \dot{\mathfrak{e}}_n^2(t)}{\mathfrak{e}_n^2(t)} \geq 0.$$

This proves the first claim.

For the second claim, we define the function $f_n(t) = \frac{\mathfrak{e}_n(t)}{\mathfrak{e}_n(0)}$. Then, f_n is also log-convex in time, positive, and we have $f_n(0) = 1$. Using the following identity

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg(\frac{1}{t}\log f_n(t)\bigg) = \frac{1}{t^2}\int_0^t s\,\frac{\mathrm{d}^2}{\mathrm{d}s^2}\bigg(\log f_n(s)\bigg)\,\mathrm{d}s$$

which can be easily checked to hold for all $t > 0$ and any positive C^1-function f_n satisfying $f_n(0) = 1$, we have that

$$0 \leq \frac{\mathrm{d}}{\mathrm{d}t}\bigg(\frac{1}{t}\log f_n(t)\bigg) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg(\log f_n(t)^{1/t}\bigg).$$

Hence, $t \mapsto \log f_n(t)^{1/t}$ is increasing which implies that also $t \mapsto f_n(t)^{1/t} = \big(\frac{\mathfrak{e}_n(t)}{\mathfrak{e}_n(0)}\big)^{1/t}$ is increasing.

To see that f_n converges, we use the polynomial bound from Theorem 3.4. We have

$$f_n(t) = \frac{\mathfrak{e}_n(t)}{\mathfrak{e}_n(0)} \leq \left(1 + \left(\frac{\mathfrak{e}_0(0)}{\mathfrak{e}_n(0)}\right)^{-1/n}\frac{t}{n}\right)^{-n}$$

$$\Rightarrow f_n(t)^{1/t} = \left(\frac{\mathfrak{e}_n(t)}{\mathfrak{e}_n(0)}\right)^{1/t} \leq \left(1 + \left(\frac{\mathfrak{e}_0(0)}{\mathfrak{e}_n(0)}\right)^{-1/n}\frac{t}{n}\right)^{-n/t} \leq 1.$$

In the last estimate we used that $(1+\alpha s)^{-1/s} \leq 1$ for all $s > 0$ and $\alpha > 0$. This completes the proof of the second claim. ☐

Since $\mathfrak{e}_n(0)^{1/t}$ converges to 1, we obviously obtain that $\mathfrak{e}_n(t)^{1/t}$ converges as $t \to \infty$. We finally remark that the above formulas show clearly the connection of the limit $\lim_{t\to\infty} f_n(t)^{1/t} =: \mathrm{e}^{-\lambda}$, $\lambda \in [0,\infty[$ to the spectral gap of $\mathbf{A}$.

3.5 Discussion of normality

In this section we discuss the assumption that the $\mathrm{L}^2(\mu)$-version of $\mathbf{A}$ is normal. In particular, we are interested how normality transfers to operators on the space $\mathcal{C}(\mathcal{Z})$. Moreover, we will see that if $\mathbf{A}$ is not normal in $\mathrm{L}^2(\mu)$ then the sequence $(\mathrm{E}_n(g))_{n\in\mathbb{N}}$ is not log-convex in general.

From operator-theoretic perspective, normality for an unbounded operator is the canonical property of an operator in Hilbert space. Roughly speaking, a normal operator is just as diagonalizable as a self-adjoint operator, but may have complex spectrum. Normal operators behave to self-adjoint ones like complex numbers to real ones. One can show that for a generic Markov operator there exists a Hilbert space in which its extension is normal (see [StS22] where for general Markov operators appropriate Hilbert spaces are constructed). The Hilbert space is not necessarily $\mathrm{L}^2(\mu)$ suggesting that lack of normality of an operator indicates an incorrectly chosen Hilbert space (compare also with the footnote on p. 14).

To understand what $\mathrm{L}^2(\mu)$-normality means on $\mathcal{C}(\mathcal{Z})$, we make the technical assumption that all involved operators are bounded (i.e. neglecting non-trivial domain issues), although much reasoning generalizes to unbounded operators. We define the multiplication operator $\mathbf{Q}_\mu$ given by

$$\mathbf{Q}_\mu : \mathcal{C}(\mathcal{Z}) \to \mathcal{C}^*(\mathcal{Z}), \quad \forall f, g \in \mathcal{C}(\mathcal{Z}) \ : \ \langle f, \mathbf{Q}_\mu g\rangle = \langle f \cdot g, \mu\rangle \, .$$

Then $\mathbf{Q}_\mu$ is symmetric and positive. With this we may express the $\mathrm{L}^2(\mu)$-adjoint $\mathbf{A}^\star$ of $\mathbf{A}$ in term of the Banach-space dual $\mathbf{A}^*$: The $\mathrm{L}^2(\mu)$-adjoint is given by $(f, \mathbf{A}^\star g)_\mu = (\mathbf{A}f, g)_\mu$, which is, by definition, equivalent to

$$\langle f \cdot \mathbf{A}^\star g, \mu\rangle = \langle \mathbf{A}f \cdot g, \mu\rangle \ \Leftrightarrow \ \langle f, \mathbf{Q}_\mu \mathbf{A}^\star g\rangle = \langle \mathbf{A}f, \mathbf{Q}_\mu g\rangle = \langle f, \mathbf{A}^* \mathbf{Q}_\mu g\rangle.$$

Since f, g are arbitrary, we conclude that the $\mathrm{L}^2(\mu)$-adjoint $\mathbf{A}^\star$ has to satisfy $\mathbf{Q}_\mu\mathbf{A}^\star = \mathbf{A}^*\mathbf{Q}_\mu$. Assuming that there is a Markov generator $\mathbf{X}$ solving the operator equation $\mathbf{Q}_\mu\mathbf{X} = \mathbf{A}^*\mathbf{Q}_\mu$ as an equation for operators $\mathcal{C}(\mathcal{Z}) \to \mathcal{C}^*(\mathcal{Z})$, we observe that $\mathbf{A}^\star$ is the $\mathrm{L}^2(\mu)$-extension of $\mathbf{X}$. If, moreover, μ is positive (i.e. $\mu(U) > 0$ for any open set $U \subset \mathcal{Z}$), then on the range of $\mathbf{A}^*\mathbf{Q}_\mu$, $\mathbf{Q}_\mu$ is one-to-one[9] and we have $\mathbf{X} = \mathbf{Q}_\mu^{-1}\mathbf{A}^*\mathbf{Q}_\mu$. For details, we refer to [Ste22].

Regarding symmetry, we observe that $\mathbf{A}$ is $\mathrm{L}^2(\mu)$-self-adjoint if and only if $\mathbf{Q}_\mu\mathbf{A} = \mathbf{A}^*\mathbf{Q}_\mu$, which for stochastic processes is usually called *detailed balance*, or that $\mathbf{A}^*$ generates a reversible process. The normality condition can be expressed by $\mathbf{XA} = \mathbf{AX}$, or equivalently, by $\mathbf{A}\mathbf{Q}_\mu^{-1}\mathbf{A}^*\mathbf{Q}_\mu = \mathbf{Q}_\mu^{-1}\mathbf{A}^*\mathbf{Q}_\mu\mathbf{A}$. We discuss these two conditions with a finite dimensional example.

Let us consider on $\mathcal{Z} = \{1, 2, 3\}$ a Markov generator given by

$$\mathbf{A} = \begin{pmatrix} -a & a & 0 \\ 0 & -b & b \\ c & 0 & -c \end{pmatrix},$$

which describes the exchange of mass along a loop with rates $a, b, c > 0$. The stationary measure is proportional to $\mu = \left(\frac{1}{a}, \frac{1}{b}, \frac{1}{c}\right)$, i.e. $\mathbf{A}^*\mu = 0$. The multiplication operator $\mathbf{Q}_\mu$ is a diagonal operator given by $\mathrm{diag}(\mu)$. An easy calculation shows that the $\mathrm{L}^2(\mu)$-adjoint $\mathbf{A}^\star$ is then given by

$$\mathbf{A}^\star = \begin{pmatrix} -a & 0 & a \\ b & -b & 0 \\ 0 & c & -c \end{pmatrix},$$

[9] The function $\mathbf{Q}_\mu^{-1}p$ can be understood as a Radon-Nikodym derivative of $p \in \mathcal{C}^*(\mathcal{Z})$ with respect to μ.

describing the exchange of mass along the reverse loop. We observe that $\mathbf{A} \neq \mathbf{A}^\star$, i.e. $\mathbf{A}$ never satisfies detailed balance.

Moreover, we can easily verify that $\mathbf{A}$ and $\mathbf{A}^\star$ commute (i.e. $\mathbf{A}$ is normal) if and only if $a = b = c$. In this situation, Theorem 3.4 is applicable and provides the polynomial decay. Of course, along solutions $\mathrm{e}^{t\mathbf{A}}$ the energies $\mathfrak{e}_n$ will decay also exponentially fast with rate given by the real part of the first non-trivial eigenvalue of $\mathbf{A}$, which is $-\frac{3}{2}a$.

Finally, we observe that log-convexity for the sequence $(\mathrm{E}_n(g))_{n\in\mathbb{N}}$ does not hold if $\mathbf{A}$ and $\mathbf{A}^\star$ do not commute. To see this, we set $a = 4$, $b = c = 1$, compute $\mathrm{E}_0, \mathrm{E}_1, \mathrm{E}_2$ and evaluate $\mathrm{E}_2\mathrm{E}_0 - \mathrm{E}_1^2$ for $g_\alpha = (1, 2\alpha, 4\alpha)$ which is

$$\mathrm{E}_2(g_\alpha)\mathrm{E}_0(g_\alpha) - \mathrm{E}_1^2(g_\alpha) = \frac{8}{3}(1 - 3\alpha)\alpha.$$

We see that this is either positive or negative depending on the choice of the parameter α. Thus, we do not have log-convexity. (Although we have exponential decay, of course.)

References

[AM*01] A. ARNOLD, P. MARKOWICH, G. TOSCANI, and A. UNTERREITER. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Comm. Partial Differential Equations*, 26(1-2), 43–100, 2001.

[AG*86] W. ARENDT, A. GRABOSCH, G. GREINER, U. GROH, H.P. LOTZ, U. MOUSTAKAS, R. NAGEL, F. NEUBRANDER, U. SCHLOTTERBECK. *One-parameter Semigroups of Positive Operators*. Springer, 1986.

[ArT20] Y. ARLINSKIJ and C. TRETTER. Everything is possible for the domain intersection $\mathrm{dom}T \cap \mathrm{dom}T^*$. *Advances in Math.*, 374, 2020.

[BaÉ85] D. BAKRY and M. ÉMERY. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[Bak85] D. BAKRY. Transformation de riesz pour les lois gaussiennes, second partie: Étude sous $\Gamma_2 \geq 0$. *Séminaire de Probab. XIX, Lecture Notes in Math., Springer-Verlag, New York/Berlin*, 1123, 145–174, 1985.

[BeB61] E. BECKENBACH and R. BELLMANN. *Inequalities*. Springer-Verlag, 1961.

[BGL14] D. BAKRY, I. GENTIL, and M. LEDOUX. *Analysis and Geometry of Markov Diffusion Operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften*. Springer, 2014.

[DuS59] N. DUNFORD and J. T. SCHWARTZ. *Linear Operator. Part I.* Interscience, John Wiley & Sons, 1959.

[GeI08] I. GENTIL and C. IMBERT. The Lévy–Fokker–Planck equation: Phi-entropies and convergence to equilibrium. *Asymptotic Analysis*, 59(3-4):125–138, 2008.

[Kat95] T. KATO. *Perturbation Theory for Linear Operators*, volume 132 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag Berlin Heidelberg, 2nd edition, 1995.

[Led00] M. LEDOUX. The geometry of Markov diffusion generators. *Ann. Fac. Sci. Toulouse Math.*, 9(2), 305–366, 2000.

[MaV00] P. A. MARKOWICH and C. VILLANI. On the trend to equilibrium for the fokker–planck equation: an interplay between physics and functional analysis. *Math. Contemp.*, 19, 1–29, 2000.

[Mey82] P. A. MEYER. Note sur les proessus d'Ornstein-Uhlenbeck. *Séminaire de Probab. XVI, Lecture Notes in Math., Springer-Verlag, New York/Berlin*, 920, 95–132, 1982.

[Mey84] P. A. MEYER. Transformation de riesz pour les lois gaussiennes. *Séminaire de Probab. XVIII, Lecture Notes in Math., Springer-Verlag, New York/Berlin*, 1059, 179–193, 1984.

[Paz83] A. PAZY. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[Sch12] K. SCHMÜDGEN. *Unbounded Self-adjoint Opertators on Hilbert Spaces*, volume 265 of *Graduate Texts in Mathematics*. Springer-Verlag, 2012.

[Sem71] Z. SEMADENI. *Banach spaces of continuous functions. Vol. I.*, volume Monografie Matematyczne. Polish Scientific Publishers, Warsaw, 1971.

[Ste05] H. STEPHAN. Lyapunov functions for positive linear evolution problems. *Z. angew. Math. Mech. (ZAMM)*, 85, 766–777, 2005.

[Ste22] H. STEPHAN. The physical inverse of a Markov operator. *in preparation*, 2022.

[StS22] A. STEPHAN and H. STEPHAN. Symmetry for Markov operators via a general Hilbert space theory. *in preparation*, 2022.

[Wu00] L. WU. Two inequalities for symmetric diffusion semigroups under $\Gamma_3 \geq 0$. *J. Funct. Analysis*, 175, 383–414, 2000.