Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Lyapunov Functions for Positive Linear Evolution Problems

Holger Stephan[1]

submitted: 31 Oct 2004

[1] Weierstrass Institute
for Applied Analysis and Stochastics
E-Mail: stephan@wias-berlin.de

No. 978
Berlin 2004

2000 *Mathematics Subject Classification.* 47D07, 47A63, 35B40, 35B50, 37A35, 46E10, 82C31.

Key words and phrases. Markov operator, Lyapunov function, Fokker-Planck equation, positive continuous semigroup, positive minimum principle, Radon measures.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

We rigorously investigate the time monotonicity of Lyapunov functions for general positive linear evolution problems, including degenerate problems. This can be done by considering the problem in the convex set of probability measures and finding a general inequality for such Radon measures and Markov operators. For linear evolution problems (with discrete or continuous time), the existence of time monotone Lyapunov functions is not a consequence of any physical properties, but of the positivity and norm conservation of the equation. In some special cases the structure of such equations is given. Moreover, we describe completely the case of time constant Lyapunov functions – a property of deterministic dynamical systems.

1 Introduction

The time asymptotic of the solution to an evolution equation – the decay towards their equilibrium solution – is of fundamental interest in applied mathematics and statistical physics. As usual, a mathematical model is called thermodynamically consistent if it has a time monotone Lyapunov function, which is interpreted as a negative entropy. This Lyapunov function (often generated by some convex function) measures the distance of two solutions. If $W_1(t)$ and $W_2(t)$ are two solutions of the problem (with differing initial data), then the Lyapunov function $H(t) = H[W_1, W_2]$ is desired to be bounded and monotone

$$0 \leq H(t_2) \leq H(t_1), \; 0 \leq t_1 \leq t_2 \; . \tag{1}$$

If we take as one solution the equilibrium solution, then $H(t) \to 0$ for $t \to \infty$ means the second law of thermodynamics.

On the other hand, if the solution of the problem can be understood as a probability (probability density, probability measure, concentration), the equation is physically sensible if the solution remains positive and normalized for all times. In the present paper we thoroughly investigate the connection between these two important properties for general linear evolution problems.

It turns out that the second law of thermodynamics (the time decay of a Lyapunov function) is not a property of the physical background of the problem, but a consequence of the positivity and norm conservation of the equation. In some sense, these two properties are equivalent. Moreover, a Lyapunov function can be generated for any linear equation by a wide class of convex functions – independent of the equation.

In the right mathematical setting, the proof of the monotonicity of $H(t)$ is very simple and transparent, and therefore allows us to describe completely the case when the Lyapunov function is constant in time: Decreasing and time constant Lyapunov functions distinguish between random and deterministic problems and not – as widely believed - between time reversible and irreversible problems.

As an example let us consider the classical Fokker-Planck equation

$$\frac{\partial}{\partial t}W(z,t) = -\sum_{i=1}^{n}\frac{\partial}{\partial z_i}\Big(a_i(z)W(z,t)\Big) + \sum_{i,j=1}^{n}\frac{\partial^2}{\partial z_i \partial z_j}\Big(b_{ij}(z)W(z,t)\Big) \tag{2}$$

with a positive definite matrix (b_{ij}), an initial condition $W_0(z) = W(z,t)$ and suitable boundary conditions. The Fokker-Planck equation describes the evolution of the probability density $W(z,t)$ of a Markov process $z(t)$ in phase space $\mathcal{Z} \subset \mathbb{R}_n$. The solution has the natural properties $W(z,t) \geq 0$ (positivity conservation) and $\int_{\mathcal{Z}} W(z,t)dz = 1$ (norm conservation) for $t \geq 0$ if W_0 is positive and normalized. The famous Boltzmann entropy

$$H(t) = H[W_1, W_2] = \int_{\mathcal{Z}} W_2(z,t)\log\frac{W_2(z,t)}{W_1(z,t)}\,dz \tag{3}$$

is a Lyapunov function for (2), i.e. (1) holds. Indeed, for strong elliptic problems much more is true:

$$\frac{d}{dt}H(t) = -\sum_{i,j=1}^{n}\int_{\mathcal{Z}}\left(\frac{W_2}{W_1}\frac{\partial}{\partial z_i}\frac{W_2}{W_1}\right)\left(\frac{W_2}{W_1}\frac{\partial}{\partial z_j}\frac{W_2}{W_1}\right)b_{ij}(z)W_1(z,t)dz \leq -cH(t)$$

with some constant $c > 0$. So, the Lyapunov function decreases even exponentially. Lyapunov functions of the type

$$H(t) = H[W_1, W_2] = \int_{\mathcal{Z}} W_2(z,t)\psi\left(\frac{W_1(z,t)}{W_2(z,t)}\right)\,dz \tag{4}$$

with convex functions $\psi(x)$ for classical Fokker-Planck equations are considered, for instance, in [2] and [13], assuming that this expression is well defined.

In general, exponential decay of $H(t)$ can not be expected and it is difficult how to understand (3) or (4) if the solution is zero somewhere. This can happen if the coefficients $a_i(z)$ and/or $b_{ij}(z)$ can degenerate and moreover, there can be more than one equilibrium solution, the solution of (2) can become singular even in finite time, the solution has finite support or the solution is a probability measure without a density w.r.t. the Lebesgue measure (or any other). For examples see, e.g., [11]. Therefore, we will look for Lyapunov functions for solutions to most general linear equations with positive and normalized solutions in the weakest monotonicity sense (1).

The mathematical framework (chapter 2) will be the space of continuous functions and its dual, containing probability measures.

The outline of the method is the following: Let p and q be two probability measures, $\mathbf{M}^*$ an operator transforming probability measures to itself – i.e. conserving positivity and norm – and $H[p,q]$ a suitable functional. We will show an inequality $0 \le H[\mathbf{M}^*p, \mathbf{M}^*q] \le H[p,q]$ (chapter 4) based on a version of Jensen's inequality (chapter 3). Then, if p and q are solutions of an equation like (2) at time t_1 and $\mathbf{M}^*$ is a solution operator transforming a solution at time t_1 into a solution at time t_2, we will get inequality (1). This result is true for general linear positive evolution problems with discrete (Markov chains) or continuous time (semigroups) and can be proved assuming only positivity and norm conservation (chapter 5). In chapter 6 we will describe in some typical situations the class of equations with this property. The classical Fokker-Planck equation (2) is a very special case of such equations.

2 Notations

In this chapter we give the used notations. All considered objects have a physical meaning. The typical physical notations are given in italics.

Let $\mathcal{Z}$ be a compact (if necessary, suitably compactified) topological Hausdorff space (*space of states*), $\mathcal{C}(\mathcal{Z})$ the space of continuous real-valued functions on $\mathcal{Z}$ (*space of observables*) and $\mathcal{C}^*(\mathcal{Z})$ (the dual of $\mathcal{C}(\mathcal{Z})$) the space of regular Radon measures on Borel sets $\mathcal{B}(\mathcal{Z})$. $\langle g, p\rangle$ with $g \in \mathcal{C}(\mathcal{Z})$ and $p \in \mathcal{C}^*(\mathcal{Z})$ is the dual pairing. $\mathbb{1} \in \mathcal{C}(\mathcal{Z})$ is the function $\mathbb{1}(z) \equiv 1$. $\mathcal{C}_{\mathbb{1}}(\mathcal{Z})$ is the corresponding one-dimensional subspace. $\mathcal{C}(\mathcal{Z})$ is an algebra by the pointwise multiplication.

$\mathcal{C}(\mathcal{Z})$ and $\mathcal{C}^*(\mathcal{Z})$ are Banach lattices with the order relations $\mathcal{C}(\mathcal{Z}) \ni g \ge 0 \iff g(z) \ge 0,\ \forall\ z \in \mathcal{Z}$ and $\mathcal{C}^*(\mathcal{Z}) \ni p \ge 0 \iff p(B) \ge 0,\ \forall\ B \in \mathcal{B}(\mathcal{Z})$. The order relation in $\mathcal{C}^*(\mathcal{Z})$ coincides with the order relation in dual spaces $p \ge 0 \iff \langle g, p\rangle \ge 0,\ \forall\ 0 \le g \in \mathcal{C}(\mathcal{Z})$. Sometimes we write $\ge_{\mathbb{R}}$, $\ge_{\mathcal{C}}$ or $\ge_{\mathcal{C}^*}$ to explain in which sense an inequality is to be understood. Elements $g \ge_{\mathcal{C}} 0$ and $p \ge_{\mathcal{C}^*} 0$ are called positive. A linear operator on a Banach lattice is called positive if it conserves positivity. This is in contrast to the notation of positivity of bilinear forms in Hilbert spaces. The subset $\mathcal{S}^*(\mathcal{Z}) = \big\{p \in \mathcal{C}^*(\mathcal{Z}) \;\big|\; p \ge 0,\ \|p\| = 1\big\}$ is the set of probability measures (*space of statistical states*). $\mathcal{L}(\mathcal{C})$ and $\mathcal{L}(\mathcal{C}^*)$ are the spaces of linear bounded operators. $\mathbf{I}$ and $\mathbf{I}^*$ are the identities, $\mathbf{O}$ and $\mathbf{O}^*$ are the zero operators.

Let us recall some properties of the mentioned spaces (see, e.g., [1], [10] and [7]):

$\mathcal{S}^*(\mathcal{Z})$ is a convex, weak* compact subset of $\mathcal{C}^*(\mathcal{Z})$. Its extremal elements $\partial_e\mathcal{S}^*(\mathcal{Z})$ are the point (or Dirac) measures δ_z for $z \in \mathcal{Z}$: $\partial_e\mathcal{S}^*(\mathcal{Z}) = \big\{p \in \mathcal{C}^*(\mathcal{Z}) \mid \exists z \in \mathcal{Z} :\ \langle g, p\rangle = g(z),\ \forall g \in \mathcal{C}(\mathcal{Z})\big\}$. We can consider $\mathcal{C}(\mathcal{Z})$ as continuous (nonlinear) functionals on $\mathcal{Z}$. In this sense the embedding $\mathcal{Z} \longleftrightarrow \partial_e\mathcal{S}^*(\mathcal{Z}) \longrightarrow \mathcal{S}^*(\mathcal{Z})$ is the canonical embedding of $\mathcal{Z}$ in its bidual. The point functionals are weak* dense in $\mathcal{C}^*(\mathcal{Z})$ and $\mathcal{S}^*(\mathcal{Z}) = \overline{\text{conv}\ \partial_e\mathcal{S}^*(\mathcal{Z})}^*$. Throughout, we will consider linear operators (and equations) acting admittedly sometimes on nonlinear sets like $\mathcal{S}^*(\mathcal{Z})$.

A deterministic physical problem can be considered in the *space of states* $\mathcal{Z}$, whereas a random problem has to be considered in the *space of statistical states* $\mathcal{S}^*(\mathcal{Z})$. Via the bijection $\mathcal{Z} \longleftrightarrow \partial_e\mathcal{S}^*(\mathcal{Z})$ the extremal elements (point measures) of all probability measures can be interpreted as *pure states*, the others are *mixed states*.

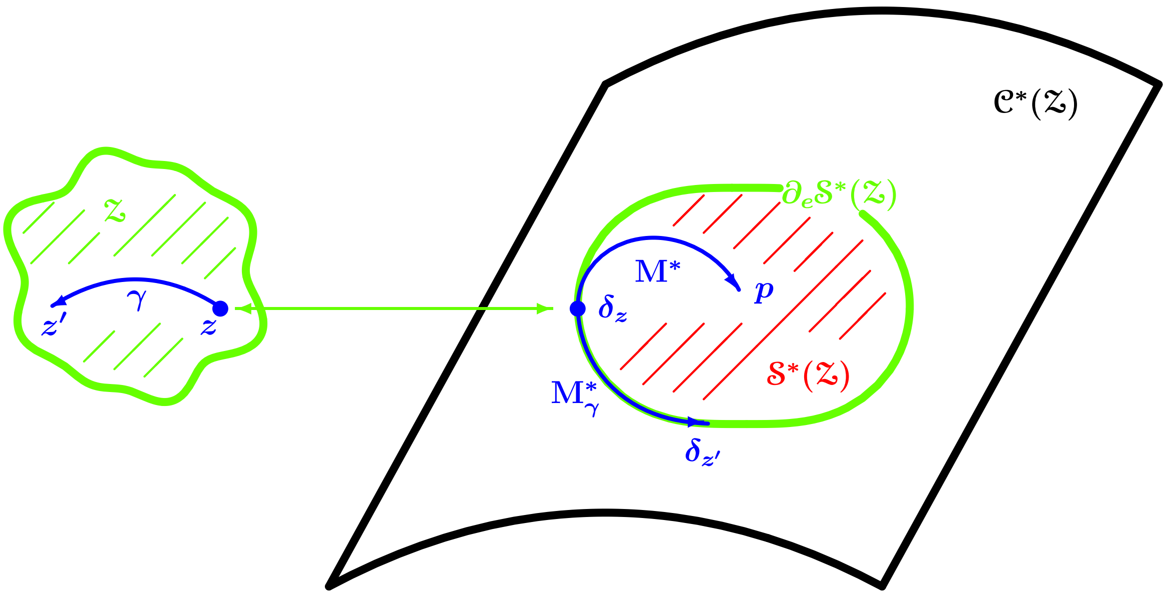
A solution operator of an evolution problem has to transform probability measures into probability measures, i.e. these operators have to map $\mathcal{S}^*(\mathcal{Z})$ into $\mathcal{S}^*(\mathcal{Z})$. This is the case iff the operator is the adjoint of a Markov operator. A linear, bounded operator $\mathbf{M}$ on $\mathcal{C}(\mathcal{Z})$ with $\mathbf{M}\mathbb{1} = \mathbb{1}$ and $\mathbf{M} \geq 0$ is called Markov operator. The set of Markov operators $\mathcal{M} = \{\mathbf{M} \in \mathcal{L}(\mathcal{C}) \mid \mathbf{M} \geq 0,\ \mathbf{M}\mathbb{1} = \mathbb{1}\}$ is convex. Its extremal elements $\partial_e\mathcal{M}$ can be characterized by one of the following equivalent properties:

$$
\begin{aligned}
\mathbf{M} \in \partial_e\mathcal{M} \quad &\Longleftrightarrow \quad \mathbf{M}(f \cdot g) = \mathbf{M}f \cdot \mathbf{M}g \quad \text{(algebra homomorphism)} \\
&\Longleftrightarrow \quad \mathbf{M}|g| = |\mathbf{M}g| \quad \text{(lattice homomorphism)} \\
&\Longleftrightarrow \quad \exists\, \gamma \in \mathcal{C}(\mathcal{Z},\mathcal{Z}) :\ \mathbf{M}g = g \circ \gamma, \ \text{i.e. } \big(\mathbf{M}g\big)(z) = g\big(\gamma(z)\big),\ z \in \mathcal{Z} \\
&\Longleftrightarrow \quad \mathbf{M}^*\partial_e\mathcal{S}^* \subset \partial_e\mathcal{S}^*, \ \text{i.e. } \forall z_1 \in \mathcal{Z}\ \exists z_2 \in \mathcal{Z} :\ \mathbf{M}^*\delta_{z_1} = \delta_{z_2}\ .
\end{aligned}
$$

Because of the last property ($\mathbf{M}^*$ maps *pure states* on *pure states*) we will call the extremal elements of Markov operators deterministic Markov operators. Because of the third property there is a one-to-one correspondence between continuous functions γ on $\mathcal{Z}$ (nonlinear objects) and deterministic Markov operators (linear objects). To point this out, sometimes we will denote deterministic Markov operators by $\mathbf{M}_\gamma$ underlying the corresponding continuous function γ. $\mathbf{M}_\gamma$ is invertible iff γ is so (i.e. γ is a topological automorphism).

Adjoints of Markov operators have the properties $\mathbf{M}^* \geq 0$ and $\|\mathbf{M}^*\| = 1$. The set of adjoints of Markov operators $\mathcal{M}^* = \{\mathbf{M}^* \in L(\mathcal{C}^*) \mid \mathbf{M} \in \mathcal{M}\}$ is convex, too.

The connection between the spaces $\mathcal{Z}$ and $\mathcal{S}^*(\mathcal{Z})$ is illustrated in the following picture.



From $\mathbf{M}\mathbb{1} = \mathbb{1}$ it follows trivially (because 1 is an eigenvalue) that there is a fixed point measure $\mu \in \mathcal{C}^*(\mathcal{Z})$ with $\mathbf{M}^*\mu = \mu$. But even more is true: Since $\mathcal{S}^*(\mathcal{Z})$ is a convex weak* compact set, the Frobenius–Perron–Krein–Rutman theorem tells us that there is a $q_0 \in \mathcal{S}^*(\mathcal{Z})$ with $\mathbf{M}^*q_0 = q_0$. If for some $\gamma \in \mathcal{C}(\mathcal{Z},\mathcal{Z})$ there exists $z_0 \in \mathcal{Z}$ with $\gamma(z_0) = z_0$, then of course $\mathbf{M}^*_\gamma \delta_{z_0} = \delta_{z_0}$. A general $\gamma \in \mathcal{C}(\mathcal{Z},\mathcal{Z})$ does not necessarily have to have a fixed point, whereas $\mathbf{M}^*_\gamma$ always have. If $(\mathbf{M}_\alpha)_{\alpha\in A}$ is a commuting family of Markov operators, then there exists $q_0 \in \mathcal{S}^*(\mathcal{Z})$ not depending on α with $\mathbf{M}^*_\alpha q_0 = q_0$, $\alpha \in A$. This is the Markov–Kakutani theorem (see [4]).

An important role for evolution problems play representations of additive semigroups on $\mathcal{C}$. A Markov chain $(q_n)_{n\geq 0}$ is the image of the adjoint $\mathbf{P}^{*n}$ of a representation $\mathbf{P}^n$ of the additive semigroup $\mathbb{N}$ in $\mathcal{M}$ on a measure $q \in \mathcal{S}^*(\mathcal{Z})$: $q_n = \mathbf{P}^{*n}q$. $\mathbf{P}^n$ has the properties $\mathbf{P}^0 = \mathbf{I}$, $\mathbf{P}^{n+m} = \mathbf{P}^n\mathbf{P}^m$, $n, m \in \mathbb{N}$. A Markov semigroup is a representation $\mathbf{T}(t)$ of the additive semigroup $\mathbb{R}_+$ in $\mathcal{M}$. It has the properties $\mathbf{T}(0) = \mathbf{I}$, $\mathbf{T}(t_1 + t_2) = \mathbf{T}(t_2)\mathbf{T}(t_1)$, $t \geq 0$. A Markovian process $(q(t))_{t\geq 0}$ is the image of the adjoint $\mathbf{T}^*(t)$ of a Markov semigroup $\mathbf{T}(t)$ on a measure $q \in \mathcal{S}^*(\mathcal{Z})$: $q(t) = \mathbf{T}^*(t)q$.

Because $\mathcal{Z}$ is fixed, we will often omit $\mathcal{Z}$, writing $\mathcal{C}$, $\mathcal{S}^*$ and so on instead of $\mathcal{C}(\mathcal{Z})$, $\mathcal{S}^*(\mathcal{Z})$, ...

3 Some variants of Jensen's inequality

We will consider convex functionals on $\mathcal{C}(\mathcal{Z})$ and $\mathcal{S}^*(\mathcal{Z})$. For this purpose we use some variants of Jensen's inequality. Throughout, $F(x)$ is a real-valued convex function, defined everywhere on $\mathbb{R}$ with values in $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$. If $F(x)$ is naturally defined on $x \in [a, b]$, we set $F(x) = +\infty$ otherwise. In this sense, an inequality $+\infty \geq c$ is allowed. Let us recall the classical Jensen inequality for sequences: Let $i = 1, ..., n$, $\alpha_i \geq 0$ with $\alpha_1 + ... + \alpha_n = 1$. Then, for $x_i \in \mathbb{R}$, the inequality

$$\sum\nolimits_{i=1}^{n} F(x_i)\alpha_i \geq F\left(\sum\nolimits_{i=1}^{n} x_i\alpha_i\right) \tag{5}$$

holds. Equality holds if F is linear, $x_i = x_j$ or $\alpha_1 = 1$. The interesting case for us is the third.

Lemma 1 *Let $g \in \mathcal{C}(\mathcal{Z})$ and $p \in \mathcal{S}^*(\mathcal{Z})$. Then inequality*

$$\big\langle F(g), q\big\rangle \quad \geq \quad F\big(\langle g, q\rangle\big) \tag{6}$$

holds. Equality holds for $q \in \partial_e\mathcal{S}^*(\mathcal{Z})$.

Proof: (6) is equivalent to the usual notation $\int_{\mathcal{Z}} F(g(z))dq(z) \geq F\big(\int_{\mathcal{Z}} g(z)dq(z)\big)$ and a simple consequence of Jensen's inequality for sequences (5) by the weak* density of point measures in $\mathcal{S}^*(\mathcal{Z})$. The case of equality is obvious. ■

Lemma 2 *Let $g \in \mathcal{C}(\mathcal{Z})$, $\mathbf{M} \in \mathcal{M}$ and $F(x)$ be continuously differentiable. Then inequality*

$$\mathbf{M}F(g) \geq F(\mathbf{M}g) \tag{7}$$

holds.

Proof: Since F is convex, for $x_1, x_2 \in \mathbb{R}$, we have

$$F(x_2) - F(x_1) \geq F'(x_1)x_2 - F'(x_1)x_1 \ .$$

Setting $x_2 = g(z_2)$, $x_1 = (\mathbf{M}g)(z_1)$, $z_1, z_2 \in \mathcal{Z}$, we get the following inequalities in $\mathbb{R}$ and $\mathcal{C}$

$$F\big(g(z_2)\big) - F\big((\mathbf{M}g)(z_1)\big) \geq_{\mathbb{R}} F'\big((\mathbf{M}g)(z_1)\big)g(z_2) - F'\big((\mathbf{M}g)(z_1)\big)(\mathbf{M}g)(z_1),\ z_1, z_2 \in \mathcal{Z}$$
$$F\big(g\big) - F\big((\mathbf{M}g)(z_1)\big)\mathbb{1} \geq_{\mathcal{C}} F'\big((\mathbf{M}g)(z_1)\big)g - F'\big((\mathbf{M}g)(z_1)\big)(\mathbf{M}g)(z_1)\mathbb{1},\ z_1 \in \mathcal{Z} \ .$$

Applying $\mathbf{M}$, we get, with $\mathbf{M}\mathbb{1} = \mathbb{1}$,

$$\mathbf{M}F\big(g\big) - F\big((\mathbf{M}g)(z_1)\big)\mathbb{1} \geq_{\mathcal{C}} F'\big((\mathbf{M}g)(z_1)\big)\mathbf{M}g - F'\big((\mathbf{M}g)(z_1)\big)(\mathbf{M}g)(z_1)\mathbb{1},\ z_1 \in \mathcal{Z} \ ,$$

and finally for $z_1 \in \mathcal{Z}$

$$\mathbf{M}F(g) - F(\mathbf{M}g) \quad \geq_{\mathcal{C}} \quad F'(\mathbf{M}g)\mathbf{M}g - F'(\mathbf{M}g)\mathbf{M}g = 0 \ . \quad \blacksquare$$

Corollary 1 *Inequality*

$$\big\langle \mathbf{M}F(g), q \big\rangle \geq \big\langle F(\mathbf{M}g), q \big\rangle \geq F\big(\langle \mathbf{M}g, q \rangle\big) \ , \quad g \in \mathcal{C} \ , \quad q \in \mathcal{S}^* \ , \quad \mathbf{M} \in \mathcal{M} \tag{8}$$

holds.

The **Proof** of the first inequality follows from (7) and the positivity of $q \in \mathcal{S}^*$. The second inequality follows from (6). $\blacksquare$

Equality in (7) holds for linear functions $F(x)$ or $g = \mathbb{1}$. The important third case of equality we put in a separate

Lemma 3 *Let $F(x)$ be strictly convex and $\mathbf{M} \in \mathcal{M}$. Then $\mathbf{M}F(g) = F(\mathbf{M}g)$ for all $g \in \mathcal{C}(\mathcal{Z})$ iff $\mathbf{M} \in \partial_e\mathcal{M}$.*

Proof:

$\Longleftarrow$ If $\mathbf{M} \in \partial_e\mathcal{M}$, there is a continuous function $\gamma : \mathcal{Z} \longrightarrow \mathcal{Z}$ with $\mathbf{M}g = g \circ \gamma$. It follows

$$\mathbf{M}F(g) = \mathbf{M}(F \circ g) = F \circ g \circ \gamma = F(g \circ \gamma) = F(\mathbf{M}g) \ .$$

$\Longrightarrow$ Let $z \in \Gamma$ and $\eta = \mathbf{M}^*\delta_z$. Of course, $\eta \in \mathcal{S}^*(\Gamma)$. We get from $\mathbf{M}F(g) = F(\mathbf{M}g)$

$$\begin{aligned} F\big(\langle g, \eta\rangle\big) &= F\big(\langle g, \mathbf{M}^*\delta_z\rangle\big) = F\big(\langle \mathbf{M}g, \delta_z\rangle\big) = F(\mathbf{M}g)(z) = (\mathbf{M}F(g))(z) = \\ &= \big\langle \mathbf{M}F(g), \delta_z\big\rangle = \big\langle F(g), \mathbf{M}^*\delta_z\big\rangle = \big\langle F(g), \eta\big\rangle \ . \end{aligned} \quad (9)$$

We will show that $\eta \in \partial_e \mathcal{S}^*(\Gamma)$. This means $\mathbf{M}^*$ maps point measures into point measures and therefore $\mathbf{M} \in \partial_e \mathcal{M}$. Assuming the opposite, $\eta \notin \partial_e \mathcal{S}^*(\Gamma)$. Then η can be represented as a convex combination, i.e. there are $\eta_1, \eta_2 \in \mathcal{S}^*(\Gamma)$ with $\eta_1 \neq \eta_2$ and $\eta = \frac{1}{2}\eta_1 + \frac{1}{2}\eta_2$. We have from inequality (6) that

$$\big\langle F(g), \eta_1\big\rangle \geq F\big(\langle g, \eta_1\rangle\big),\ \big\langle F(g), \eta_2\big\rangle \geq F\big(\langle g, \eta_2\rangle\big) \ . \quad (10)$$

Using (9) and (10), we get

$$\begin{aligned} F\left(\frac{1}{2}\langle g, \eta_1\rangle + \frac{1}{2}\langle g, \eta_2\rangle\right) &= F\big(\langle g, \frac{1}{2}\eta_1 + \frac{1}{2}\eta_2\rangle\big) = F\big(\langle g, \eta\rangle\big) = \big\langle F(g), \eta\big\rangle = \\ &= \frac{1}{2}\big\langle F(g), \eta_1\big\rangle + \frac{1}{2}\big\langle F(g), \eta_2\big\rangle \geq \frac{1}{2}F\big(\langle g, \eta_1\rangle\big) + \frac{1}{2}F\big(\langle g, \eta_2\rangle\big), \end{aligned}$$

i.e., with $x = \langle g, \eta_1\rangle$ and $y = \langle g, \eta_2\rangle$, we get

$$F\left(\frac{1}{2}(x+y)\right) \geq \frac{1}{2}F(x) + \frac{1}{2}F(y) \ .$$

But F is strictly convex. Therefore, $x = y$. It follows that $\langle g, \eta_1\rangle = \langle g, \eta_2\rangle$, $g \in \mathcal{C}(\mathcal{Z})$, hence $\eta_1 = \eta_2$, a contradiction. ■

This lemma shows that the equality $\mathbf{M}F(g) = F(\mathbf{M}g)$ is a further equivalent characterization of deterministic Markov operators.

Remark:
The results can be extended to arbitrary convex functions by suitable limit processes using, e.g., the Yoshida approximation $F_\lambda(x) = \inf_{y\in\mathbb{R}}\left(\frac{1}{2\lambda}(x-y)^2 + F(x)\right)$ (see [3]).

4 A convex functional on probability measures

We are looking for convex functionals on $\mathcal{S}^* \times \mathcal{S}^*$. Formula (3) suggests a convex function of a quotient of two measures, what, of course, is difficult to understand in general. A natural convex functional generated by a convex function $F : \mathbb{R} \to \mathbb{R}$ and defined on $\mathcal{C} \times \mathcal{S}^*$ is $\big\langle F(g), q\big\rangle$, for $g \in \mathcal{C}$ and $p \in \mathcal{S}^*$. To get from this a functional on $\mathcal{S}^* \times \mathcal{S}^*$, we use the Legendre transform. Let $p, q \in \mathcal{S}^*$. We define a functional on the convex set $\mathcal{S}^* \times \mathcal{S}^*$ by

$$H[p, q] = \sup_{g\in\mathcal{C}}\Big(\langle g, p\rangle - \big\langle F(g), q\big\rangle\Big) \ . \quad (11)$$

If $F(x) = +\infty$ for $x \notin [a, b]$ we take the supremum over the set $\mathcal{C}_{[a,b]} = \{g \in \mathcal{C} \mid a\mathbb{1} \leq_{\mathcal{C}} g \leq_{\mathcal{C}} b\mathbb{1}\}$.

This functional is well defined, but the supremum is taken over a large set, so it is not clear whether the functional is $+\infty$ everywhere. We will show that on the diagonal $H[q,q] = F^*(1)$, where $F^*(y) = \sup_{x\in\mathbb{R}}\big(xy - F(x)\big)$ is the convex conjugate of $F(x)$. To exclude the uninteresting case $H \equiv +\infty$, we will normalize H, setting $F^*(1) = 0$. The following theorem holds:

Theorem 1 *Let* $\mathbf{M} \in \mathcal{M}$. *The functional* $H[p,q]$ *defined on* $\mathcal{S}^* \times \mathcal{S}^*$ *by (11) has the following properties:*

i) (convexity) H *is convex on* $\mathcal{S}^* \times \mathcal{S}^*$ *.*

ii) (boundedness) $H[p,q] \geq H[q,q] = 0$.

iii) (monotonicity) $0 \leq H[\mathbf{M}^*p, \mathbf{M}^*q] \leq H[p,q]$.

Proof: At first, let us note, that the proof does not change if we take in the following the supremum over $\mathcal{C}_{[a,b]}$ instead of $\mathcal{C}$, because $F^*(y) = \sup_{x\in\mathbb{R}}\big(xy - F(x)\big) = \sup_{x\in[a,b]}\big(xy - F(x)\big)$, and from $a\mathbb{1} \leq_{\mathcal{C}} g \leq_{\mathcal{C}} b\mathbb{1}$ we conclude $a\mathbb{1} \leq_{\mathcal{C}} \mathbf{M}g \leq_{\mathcal{C}} b\mathbb{1}$, thus $\mathbf{M}\mathcal{C}_{[a,b]} \subset \mathcal{C}_{[a,b]}$.

Convexity of H: Let $\binom{p_1}{q_1} \in \mathcal{S}^* \times \mathcal{S}^*$, $\binom{p_2}{q_2} \in \mathcal{S}^* \times \mathcal{S}^*$, $\alpha_1, \alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$ and $\binom{p}{q} = \alpha_1\binom{p_1}{q_1} + \alpha_2\binom{p_2}{q_2}$. From the definition (11) we get

$$\begin{aligned} \alpha_1 H[p_1,q_1] &\geq \alpha_1\langle g,p_1\rangle - \alpha_1\big\langle F(g),q_1\big\rangle = \langle g,\alpha_1 p_1\rangle - \big\langle F(g),\alpha_1 q_1\big\rangle\ , \quad g \in \mathcal{C}\ , \\ \alpha_2 H[p_2,q_2] &\geq \alpha_2\langle g,p_2\rangle - \alpha_2\big\langle F(g),q_2\big\rangle = \langle g,\alpha_2 p_2\rangle - \big\langle F(g),\alpha_2 q_2\big\rangle\ , \quad g \in \mathcal{C} \end{aligned}$$

hence, adding this inequalities,

$$\alpha_1 H[p_1,q_1] + \alpha_2 H[p_2,q_2] \;\geq\; \Big(\langle g,p\rangle - \big\langle F(g),q\big\rangle\Big)\ , \quad g \in \mathcal{C}\ .$$

It follows

$$\alpha_1 H[p_1,q_1] + \alpha_2 H[p_2,q_2] \;\geq\; \sup_{g\in\mathcal{C}}\Big(\langle g,p\rangle - \big\langle F(g),q\big\rangle\Big) = H[p,q]\ .$$

Equality and Boundedness: We have

$$\begin{aligned} H[p,q] &= \sup_{g\in\mathcal{C}}\Big(\langle q,p\rangle - \big\langle F(g),q\big\rangle\Big) \geq \sup_{g\in\mathcal{C}_1}\big(\langle g,p\rangle - \langle F(g),q\rangle\big) = \\ &= \sup_{g_0\in\mathbb{R}}\big(g_0\langle\mathbb{1},p\rangle - F(g_0)\langle\mathbb{1},q\rangle\big) = \sup_{g_0\in\mathbb{R}}\big(g_0 - F(g_0)\big) = F^*(1) = 0\ . \end{aligned}$$

In particular, $H[q,q] \geq F^*(1) = 0$. On the other hand, we have

$$F^*(1) \;=\; \sup_{x\in\mathbb{R}}\big(x - F(x)\big) \;\geq_{\mathbb{R}}\; x - F(x)\ , \quad x \in \mathbb{R}\ .$$

Setting $x = g(z)$, $g \in \mathcal{C}$, $z \in \mathcal{Z}$ we get

$$\begin{aligned} F^*(1) &\geq_{\mathbb{R}} g(z) - F(g(z))\ , \quad g \in \mathcal{C}\ , \quad z \in \mathcal{Z}\ , \\ F^*(1)\mathbb{1} &\geq_{\mathcal{C}} g - F(g)\ , \quad g \in \mathcal{C}\ , \\ F^*(1) &\geq_{\mathbb{R}} \langle g,q\rangle - \langle F(g),q\rangle\ , \quad g \in \mathcal{C}\ , \quad q \in \mathcal{S}^*\ , \\ F^*(1) &\geq_{\mathbb{R}} \sup_{g\in\mathcal{C}}\Big(\langle g,q\rangle - \langle F(g),q\rangle\Big) = H[q,q]\ , \quad q \in \mathcal{S}^*\ . \end{aligned}$$

Together we can conclude $H[p,q] \geq H[q,q] = F^*(1) = 0$.

Monotonicity: Denoting by $R(\mathbf{M})$ the range of $\mathbf{M}$ and using Lemma 2, we get

$$H[\mathbf{M}^*p, \mathbf{M}^*q] = \sup_{g\in\mathcal{C}} \Big(\langle g, \mathbf{M}^*p\rangle - \big\langle F(g), \mathbf{M}^*q\big\rangle \Big) = \sup_{g\in\mathcal{C}} \Big(\langle \mathbf{M}g, p\rangle - \big\langle \mathbf{M}F(g), q\big\rangle \Big) \leq$$
$$\leq \sup_{g\in\mathcal{C}} \Big(\langle \mathbf{M}g, p\rangle - \big\langle F(\mathbf{M}g), q\big\rangle \Big) = \sup_{h\in R(\mathbf{M})} \Big(\langle h, p\rangle - \big\langle F(h), q\big\rangle \Big) \leq$$
$$\leq \sup_{h\in\mathcal{C}} \Big(\langle h, p\rangle - \big\langle F(h), q\big\rangle \Big) = H[p,q] \ .$$

Together with the boundedness we get

$$0 \leq H[\mathbf{M}^*p, \mathbf{M}^*q] \leq H[p,q] \ , \quad p,q \in \mathcal{S}^* \ , \quad \mathbf{M}^* \in \mathcal{M}^* \ . \quad \blacksquare \tag{12}$$

Let us point out that equality $H[\mathbf{M}^*p, \mathbf{M}^*q] = H[p,q]$ holds if

(1) $\mathbf{M}F(g) \ = \ F(\mathbf{M}g)$, and

(2) $\overline{R(\mathbf{M})} = \mathcal{C}(\mathcal{Z})$,

i.e. by Lemma 3 if $\mathbf{M}$ is a deterministic Markov operator with weakly dense range.

To illustrate functional (11), we consider

4.1 Some examples.

Theorem 1 shows that $H[q,q] = 0$, so $H[p,q]$ is not identical $+\infty$. It is interesting to describe the subset of $\mathcal{S}^* \times \mathcal{S}^*$, where $H[p,q] < \infty$.

1) Let us assume that there exists the Radon-Nikodym derivative of p by q: $h_{p/q} = \frac{p}{q}_{\text{RN}} \in L_1(dq)$. Then we have

$$\begin{aligned} H[p,q] &= \sup_{g\in\mathcal{C}} \Big(\langle g, p\rangle - \big\langle F(g), q\big\rangle \Big) = \\ &= \sup_{g\in\mathcal{C}} \left(\int_{\mathcal{Z}} g(z)dp(z) - \int_{\mathcal{Z}} F\big(g(z)\big)dq(z) \right) = \\ &= \sup_{g\in\mathcal{C}} \left(\int_{\mathcal{Z}} h_{p/q}(z)g(z)dq(z) - \int_{\mathcal{Z}} F\big(g(z)\big)dq(z) \right) = \\ &= \sup_{g\in\mathcal{C}} \int_{\mathcal{Z}} \Big(h_{p/q}(z)g(z) - F\big(g(z)\big)\Big)dq(z) = \int_{\mathcal{Z}} F^*\Big(h_{p/q}(z)\Big)dq(z) \end{aligned}$$

because the supremum can be taken pointwise under the integral. If, moreover, q is differentiable w.r.t. the Lebesgue measure, then so is p. Let $W_p(z)dz = dp(z)$ and $W_q(z)dz = dq(z)$. Then $h_{p/q}(z) = W_p(z)/W_q(z)$, and functional (11) reads

$$H[p,q] = \int_{\mathcal{Z}} F^* \left(\frac{W_p(z)}{W_q(z)} \right) W_q(z)dz \ . \tag{13}$$

This is the entropy functional used in [2] and [13]. For $F(x) = -1 - \log(-x),\ x < 0$, and $F(x) = \infty$, otherwise, we get $F^*(x) = -\log x,\ x > 0$ — Boltzmann's entropy (3).

The following example is in some sense opposite:

2) Let $A, B \in \mathcal{B}(\mathcal{Z})$ be two closed disjunct Borel subsets, $p(A) = 1$, $q(B) = 1$ and $F(1) < +\infty$. Then – because of $\mathcal{Z}$ being Hausdorff – for every real c there exists $g_c \in \mathcal{C}$ with $g_c(A) = c$ and $g_c(B) = 1$,

$$\begin{aligned} H[p,q] &= \sup_{g\in\mathcal{C}} \Big(\langle g, p\rangle - \big\langle F(g), q \big\rangle \Big) \geq \\ &\geq \sup_{g_c\in\mathcal{C}} \Big(\langle g, p\rangle - \big\langle F(g), q \big\rangle \Big) = \sup_{c\in\mathbb{R}} \ (c - F(1)) = \infty \ . \end{aligned}$$

The following example shows that for some F, H is bounded everywhere on $\mathcal{S}^* \times \mathcal{S}^*$.

3) Let $-\infty < a < b < \infty$ and $F(x) = +\infty$ for $x \notin [a, b]$. Then, $F(x)$ is bounded from below, $F^*(y)$ is defined everywhere and $\inf_{x\in[a,b]} F(x) = F^*(0)$. From (6) and $g \leq b$ for $g \in \mathcal{C}_{[a,b]}$ follows

$$H[p,q] \ = \ \sup_{g\in\mathcal{C}_{[a,b]}} \Big(\langle g, p\rangle - \big\langle F(g), q \big\rangle \Big) \leq \sup_{g\in\mathcal{C}_{[a,b]}} \Big(\langle g, p\rangle - F\big(\langle g, q\rangle\big) \Big) \leq b - F^*(0) \ .$$

5 Lyapunov functions for evolution problems

If p and q are two solutions of some evolution problem at time t_1 and $\mathbf{M}^*p$ resp. $\mathbf{M}^*q$ are two solutions of the same problem at time $t_2 > t_1$, then inequality (12) gives us a monotonicity condition of a Lyapunov function of this problem. For discrete times we get Lyapunov functions for Markov chains, for continuous time we get Lyapunov functions for Markovian processes, i.e. for positive semigroups on $\mathcal{S}^*(\mathcal{Z})$.

5.1 Lyapunov functions for Markov chains

Corollary 2 *Let* $\mathbf{P} \in \mathcal{M}$ *be the generator of a Markov chain.* $p_0, p_1, p_2, ...$ *and* $q_0, q_1, q_2, ...$ *are the corresponding Markov chains with the initial measures* p_0 *and* q_0*, resp. Then the inequality*

$$0 \leq H[p_{n+1}, q_{n+1}] \leq H[p_n, q_n]$$

holds.

Proof: Since $\mathbf{P}$ is the generator, we have $p_{n+1} = \mathbf{P}^* p_n$ and $q_{n+1} = \mathbf{P}^* q_n$. Take in (12) $p = p_n$, $q = q_n$ and $\mathbf{M} = \mathbf{P}$, the claim follows. ■

Remark: A Markov chain $p_n = p(t_n)$ can be considered as a time approximation of a time continuous problem having Lyapunov functions. If the approximation $p(t_{n+1}) = \mathbf{P}^* p(t_n)$ conserves positivity and norm, then by the corollary the time discrete problem has Lyapunov functions, too. See the example in 6.1 for further remarks.

5.2 Lyapunov functions for positive semigroups

For continuous time, it is natural to take for $\mathbf{M}^*$ a semigroup $\mathbf{T}^*(t)$. If p_0 is the initial value, of some evolution equation and $p = p(t_1) = \mathbf{T}^*(t_1)p_0$ is the solution at time t_1, taking $\mathbf{M}^* = \mathbf{T}^*(t_2 - t_1)$ we get by the semigroup property $\mathbf{M}^* p = \mathbf{T}^*(t_2 - t_1)\mathbf{T}^*(t_1)p_0 = \mathbf{T}^*(t_2)p_0 = p(t_2)$ – the solution at time t_2, and inequality (12) reads

$$0 \leq H[p(t_2), q(t_2)] \leq H[p(t_1), q(t_1)] \ . \tag{14}$$

Now we have to answer the question whether an evolution equation has a semigroup solution in $\mathcal{S}^*(\mathcal{Z})$ and, of course, it is more interesting to characterize these semigroups by properties of their generators $\mathbf{A}$. For strongly continuous semigroups $\mathbf{T}(t)$ it is well known that $g(t) = \mathbf{T}(t)g_0$ is the solution of the equation $\dot{g}(t) = \mathbf{A}g(t)$ (here and in the following $\dot{g}(t)$ denotes the time derivative). Dealing in $\mathcal{C}^*(\mathcal{Z})$, the problem is that there are no strong or weak continuous semigroups (except those with bounded generator). Therefore, we can analyze evolution equations in $\mathcal{C}^*(\mathcal{Z})$ only in a weak* sense, starting with continuous semigroups in $\mathcal{C}(\mathcal{Z})$ (for continuous semigroups in Banach spaces and their adjoints, see, e.g., [9]).

Looking for positive semigroups in $\mathcal{C}^*(\mathcal{Z})$, we have to look for positive continuous semigroups in $\mathcal{C}(\mathcal{Z})$ at first. There is a necessary and sufficient condition for a generator of a continuous semigroup to be a generator of a positive one in the space of continuous functions on compact topological spaces. Following [1], we will say that an operator $\mathbf{A}$ with dense domain $D(\mathbf{A}) \subset \mathcal{C}(\mathcal{Z})$ satisfies the positive minimum principle, if

$$(\mathbf{A}g)(z_+) \leq 0, \ \ g \in D(\mathbf{A}) \ , \tag{15}$$

where z_+ is the point where g contains its maximum: $g(z_+) = \sup_{z \in \mathcal{Z}} g(z)$. A generator of a continuous semigroup in $\mathcal{C}(\mathcal{Z})$ is a generator of a positive continuous semigroup iff it satisfies (15) (for the proof, see [1]). If additionally $\mathbf{A}\,\mathbb{1} = 0$, then the solution $g(t) = \mathbf{T}(t)g_0$ of the equation

$$\dot{g}(t) = \mathbf{A}g(t) \tag{16}$$

is $g(t) = \mathbf{T}(t)g_0$, where $\mathbf{T}(t)$ is a semigroup of Markov operators. In probability theory, equation (16) is called Kolmogorov backwards equation. From a physical point of view, equation (16) describes the evolution of an observable and therefore is called *Heisenberg representation* or *Heisenberg picture*.

The adjoint semigroup $\mathbf{T}^*(t)$ acts in $\mathcal{S}^*(\mathcal{Z})$. Pairing (16) with some $p_0 \in \mathcal{S}^*(\mathcal{Z})$, we get

$$\frac{d}{dt}\Big\langle \mathbf{T}(t)g_0, p_0 \Big\rangle = \Big\langle \mathbf{A}\mathbf{T}(t)g_0, p_0 \Big\rangle,\ g_0 \in D(\mathbf{A})\ .$$

Since $\mathbf{A}$ and $\mathbf{T}(t)$ commute, setting $p(t) = \mathbf{T}^*(t)p_0$, we get

$$\frac{d}{dt}\Big\langle g_0, p(t) \Big\rangle = \Big\langle \mathbf{A}g_0, p(t) \Big\rangle,\ g_0 \in D(\mathbf{A})\ . \tag{17}$$

$\mathbf{A}^*$ exists since $D(\mathbf{A})$ is dense in $\mathcal{C}(\mathcal{Z})$, so we can write

$$\frac{d}{dt}\Big\langle g_0, p(t) \Big\rangle = \Big\langle g_0, \mathbf{A}^*p(t) \Big\rangle,\ g_0 \in D(\mathbf{A})\ ,$$

i.e. $p(t)$ is the solution with $p(0) = p_0$ of some equation in weak* sense. We will write this in the following way

$$\dot{p}(t) \stackrel{*}{=} \mathbf{A}^*p(t),\ p(0) = p_0\ . \tag{18}$$

This is the Kolmogorov forward equation.

In general, it is very difficult to write down $\mathbf{A}^*$ explicitly or even to describe the domain $D(\mathbf{A}^*)$. Therefore, from a practical point of view, it is better to solve equation (16) and calculate the adjoint $\mathbf{T}^*(t)$ than to solve (18).

Now we can state the following

Theorem 2 *Let* $\mathbf{A}$ *be a generator of a continuous semigroup satisfying* $\mathbf{A}\,\mathbb{1} = 0$ *and the positive minimum principle (15). Then for two initial probability measures* $p_0, q_0 \in \mathcal{S}^*(\mathcal{Z})$*, for* $t \geq 0$*, the weak* solutions* $p(t) \subset \mathcal{S}^*(\mathcal{Z})$ *and* $q(t) \subset \mathcal{S}^*(\mathcal{Z})$ *of equation (18) satisfy for any functional* H *defined by (11) the inequality*

$$0 \leq H\big[p(t_2), q(t_2)\big] \quad \leq \quad H\big[p(t_1), q(t_1)\big]\ ,\quad t_2 \geq t_1 \geq 0\ . \tag{19}$$

Proof: Since $\mathbf{A}$ is a generator of a continuous semigroup satisfying the positive minimum principle, the unique solution of equation $\dot{g}(t) = \mathbf{A}g(t)$, $g(0) = g_0$ is given by a positive continuous semigroup $\mathbf{T}(t)$: $g(t) = \mathbf{T}(t)g_0$. From $\mathbf{A}\mathbb{1} = 0$ and $\mathbf{T}(t)g_0 - g_0 = \int_0^t \mathbf{A}g_0 dt$, we conclude $\mathbf{T}(t)\mathbb{1} = \mathbb{1}$. Therefore, $\mathbf{T}(t) \subset \mathcal{M}$ and the dual semigroup $\mathbf{T}^*(t)$ maps $\mathcal{S}^*(\mathcal{Z})$ into $\mathcal{S}^*(\mathcal{Z})$. From the strong equation $\dot{g}(t) = \mathbf{A}g(t)$, we get for any $p_0 \in \mathcal{S}^*(\mathcal{Z})$ the weak* equation (18). Now, as mentioned above, taking two initial data $p_0, q_0 \in \mathcal{S}^*(\mathcal{Z})$, we get (19) for the corresponding solutions $p(t)$ and $q(t)$. ■

The theorem shows that any linear evolution equation of type (18) conserving positivity and norm has a large family of Lyapunov functions (11) parameterized by more or less arbitrary convex functions F. Moreover, the operator $\mathbf{A}^*$ and the function F do not depend on each other.

5.3 Time constant Lyapunov functions

We have shown that $H(t)$ is a function monotone in time. Considering semigroups $\mathbf{T}^*(t)$, it is interesting to investigate the case when the functional $H[p,q]$ does not depend on time. In this case we have

$$H[p(t), q(t)] = H[p_0, q_0],\ t \geq 0\ . \tag{20}$$

If the problem is time-reversible, i.e. if $\mathbf{T}^{-1}(t)$ exists, then $\mathbf{T}^*(t)$ is a group and, of course, (20) holds: If $\mathbf{M}^{-1}$ exists, then $\mathbf{M}^{*-1}$ is a Markov operator, too, and setting $p := \mathbf{M}^{*-1}p$ and $q := \mathbf{M}^{*-1}q$, from $H[\mathbf{M}^*p, \mathbf{M}^*q] \leq H[p,q]$ and $H[\mathbf{M}^{*-1}p, \mathbf{M}^{*-1}q] \leq H[p,q]$, we get $H[p,q] \leq H[\mathbf{M}^*p, \mathbf{M}^*q] \leq H[p,q]$. But this is not the only case.

A linear operator $\mathbf{A}$ in $\mathcal{C}(\mathcal{Z})$ is called a derivation (see [1]), if $D(\mathbf{A})$ is a sub-algebra in $\mathcal{C}(\mathcal{Z})$ and $\mathbf{A}(f \cdot g) = g \cdot \mathbf{A}f + f \cdot \mathbf{A}g$.

Theorem 3 *Let* $\mathbf{A}$ *be a generator of a continuous semigroup and a derivation. Then, for any two solutions* $p(t) \subset \mathcal{S}^*(\mathcal{Z})$ *and* $q(t) \subset \mathcal{S}^*(\mathcal{Z})$ *of equation* $\dot{p}(t) = \mathbf{A}^*p(t)$ *with* $p(0) = p_0$ *and* $q(0) = q_0$*, the following identity holds*

$$H\big[p(t), q(t)\big] = H[p_0, q_0],\ t \geq 0\ . \tag{21}$$

Proof:
Since $\mathbf{A}$ is a generator of a continuous semigroup $\mathbf{T}(t)$, the range of $\mathbf{T}(t)$ is dense in $\mathcal{C}(\mathcal{Z})$. Moreover, because $\mathbf{A}$ is a derivation, for every $t \geq 0$, the operator $\mathbf{T}(t)$ is a deterministic Markov operator (see [1]). Now, by Lemma 3, (21) follows. ■

5.4 Stationary solutions and solvability in L_p

Since the family $\mathbf{T}^*(t)$ commutes, by the Markov–Kakutani theorem, there exists a $q_\infty \in \mathcal{S}^*(\mathcal{Z})$ with $\mathbf{T}^*(t)q_\infty = q_\infty$. From this we can conclude $q_\infty \in D(\mathbf{A}^*)$ and $\mathbf{A}^*q_\infty = 0$. Now in Theorem 2 we can take q_∞ as one solution and get

$$0 \leq H[p(t_2), q_\infty] \leq H[p(t_1), q_\infty] \tag{22}$$

and

$$0 \leq H[q_\infty, p(t_2)] \leq H[q_\infty, p(t_1)]\ . \tag{23}$$

But we can not conclude from this that a solution $p(t)$ tends to q_∞. Of course, there exist a $q \in \mathcal{S}^*(\mathcal{Z})$ and a subsequence $\{t_k\}$ with $p_{t_k} \overset{*}{\rightharpoonup} q$, but in general $\mathbf{A}^*q \neq 0$. Moreover, in general q_∞ is not unique. A general result is the following (see [11]): If $p_0 \in D(\mathbf{A}^*)$ and $p_t \overset{*}{\rightharpoonup} q_\infty$, then $\mathbf{A}^*q_\infty = 0$ and $p_\infty = p_0 + \int_0^\infty \mathbf{T}^*(t)\mathbf{A}^*p_0\ dt$.

Of interest is whether an equation of type $\dot{p}(t) = \mathbf{A}^*p(t)$ can be solved in a more customary function space, say L_p.

Let $q_\infty \in \mathcal{S}^*(\mathcal{Z})$ be a stationary solution ($\mathbf{A}^* q_\infty = 0$) and q_0 be such that the Radon-Nikodym derivative $u_0 = \frac{q_0}{q_\infty}_{\mathrm{RN}}$ exists. Taking $F(x) = |x|^{\frac{p}{p-1}}$ with some $1 < p < \infty$, we get

$$H[q_0, q_\infty] = \int_{\mathcal{Z}} u_0^p(z) dq_\infty(z) = \|u_0\|^p_{L_p(dq_\infty)} \,.$$

Now, from (22) we conclude

$$0 \le H[q(t), q_\infty] = \|u(t)\|^p_{L_p(dq_\infty)} \le H[q(0), q_\infty] = \|u_0\|^p_{L_p(dq_\infty)} \,.$$

So, for $u_0 \in L_p(dq_\infty)$, the equation $\dot{u}(t) = \mathbf{A}^* u(t)$ is solvable in $L_p(dq_\infty)$. Starting with an arbitrary $q_0 \in \mathcal{S}^*$, at some time the solution belongs to $L_p(dq_\infty)$, i.e. there is a time t_0 with $0 \le t_0 \le \infty$ with $\frac{q(t)}{q_\infty}_{\mathrm{RN}}$ existing for $t \ge t_0$.

6 Examples

In the following, we will consider special cases of state spaces $\mathcal{Z}$. We have to investigate the shape of Markov operators, generators of Markov semigroups and their adjoints.

6.1 The finite dimensional case

The simplest but nevertheless important case is a state space consisting of a finite number of states: $\mathcal{Z} = \{z_1, ..., z_n\}$. Taking the discrete topology, then $\mathcal{Z}$ is compact and Hausdorff and the space of continuous functions on $\mathcal{Z}$ is $\mathcal{C}(\mathcal{Z}) = \mathbb{R}_n$, the n-dimensional vector space with the max-norm. The dual is $\mathcal{C}(\mathcal{Z}) = \mathbb{R}_n^*$, the n-dimensional vector space with the l_1-norm. In this spaces we fixed the canonical base $e_i = e_i^*$, consisting of vectors with 1 at the i-th component and 0 otherwise. $\mathbb{1}$ is the vector with 1 at every component. $\mathcal{S}^*(\mathcal{Z})$ is the $(n-1)$-dimensional simplex, the convex hull of the base. All linear operators are bounded and generators of strong continuous semigroups. Thus, corresponding evolution equations are valid in a strong sense.

In the canonical base a Markov operator is a matrix $\mathbf{M} = (m_{ij})$ with $m_{ij} \ge 0$ and $\sum_{j=1}^n m_{ij} = 1$. The adjoint can be written as

$$\mathbf{M}^* = \begin{pmatrix} 1 - m_{21} - \ldots - m_{n1} & m_{12} & \cdots & m_{1n} \\ m_{21} & 1 - m_{12} - \ldots - m_{n2} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & 1 - m_{1n} - \ldots - m_{n-1,n} \end{pmatrix}, \; m_{ij} \ge 0.$$

The set of extremal elements $\partial_e \mathcal{M}$ of $\mathcal{M}$ consists of n^n matrices with one 1 in each row. The set of invertible extremal elements consist of $n!$ matrices with one 1 in each row and each column – the representation of the permutation group on $\mathcal{Z}$.

A semigroup of adjoints of Markov operators in $\mathcal{C}^*(\mathcal{Z})$ is an operator family $\mathbf{T}^*(t) = (\omega_{ij}(t))$ with $\omega_{ij}(t) \geq 0$ and can be written as

$$\mathbf{T}^*(t) = \begin{pmatrix} 1-\omega_{21}(t)-\ldots-\omega_{n1}(t) & \cdots & \omega_{1n}(t) \\ \vdots & \ddots & \vdots \\ \omega_{n1}(t) & \cdots & 1-\omega_{1n}(t)-\ldots-\omega_{n-1,n}(t) \end{pmatrix}, \ \omega_{ij}(t) \geq 0.$$

The $\omega_{ij}(t)$ are continuously differentiable functions for $t \geq 0$. From $\mathbf{T}^*(0) = \mathbf{I}^*$ follows $\omega_{ij}(0) = 0$ and therefore $\omega'_{ij}(0) \geq 0$. Setting $a_{ij} = \omega'_{ij}(0)$, we get for the generator $\mathbf{A}^* = \mathbf{T}^{*\prime}(0)$

$$\mathbf{A}^* = \begin{pmatrix} -a_{21}-\ldots-a_{n1} & a_{12} & \cdots & a_{1n} \\ a_{21} & -a_{12}-\ldots-a_{n2} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & -a_{1n}-\ldots-a_{n-1,n} \end{pmatrix}, \ a_{ij} \geq 0. \quad (24)$$

The corresponding evolution equations are

$$\dot{g}_i(t) = \big(\mathbf{A}g(t)\big)_i = \sum_{i\neq j=1}^{n} a_{ji}\big(g_i(t) - g_j(t)\big) , \quad (25)$$

$$\dot{p}_j(t) = \big(\mathbf{A}^*p(t)\big)_j = \sum_{j\neq i=1}^{n} \big(a_{ji}p_i(t) - a_{ij}p_j(t)\big) . \quad (26)$$

If $\mathbf{A}$ is a generator of a Markov semigroup, then it must have the form (24). The reverse is true, too: For some $g \in \mathcal{C}(\mathcal{Z})$, let j_+ be the index of the maximal element: $g_{j_+} = \max_{i=1,\ldots,n} g_i$, then $g_i \leq g_{j_+}$, for $i = 1, ..., n$, hence

$$\big(\mathbf{A}g\big)_{j_+} = \sum_{j_+\neq i=1}^{n} a_{ij_+}(g_i - g_{j_+}) \leq 0 .$$

Thus, $\mathbf{A}$ satisfies the positive minimum principle. Moreover, $\mathbf{A}\mathbb{1} = 0$. Hence, $\mathbf{A}$ is the generator of a semigroup of Markov operators.

We can conclude: Every functional (13) is a Lyapunov function for equation (26).

If $q_i > 0$, $i = 1, ..., n$, the Radon-Nikodym derivative of p by q exists and therefore for any convex $F^* : \mathbb{R}_+ \to \mathbb{R}$ with $F^*(1) = 0$, we have the Lyapunov function

$$H(t) = H[p(t), q(t)] = \sum_{i=1}^{n} q_i(t) F^*\left(\frac{p_i(t)}{q_i(t)}\right) \geq 0 .$$

In the finite dimensional case, $\mathbf{T}^*(t)$ is strongly continuous, so $H(t)$ is differentiable and a simple calculation shows (assuming F^* continuously differentiable)

$$\frac{d}{dt}H(t) = -\sum_{i,j=1}^{n} a_{ji}q_j(t)\left[F^*\left(\frac{p_j}{q_j}\right) - F^*\left(\frac{p_i}{q_i}\right) - \left(\frac{p_j}{q_j} - \frac{p_i}{q_i}\right)F^{*\prime}\left(\frac{p_i}{q_i}\right)\right] \leq 0 .$$

The finite dimensional case can be considered as the approximation of a continuous problem. If the approximated problem conserves positivity and norm like the continuous one, then it has a Lyapunov function like that one.

Now let us consider the case when $H(t)$ is constant in time. This is the case if $\mathbf{T}(t) \in \partial_e \mathcal{M}$, i.e. $\mathbf{T}(t)$ is a continuous representation of the semigroup $\mathbb{R}_+$ in $\partial_e \mathcal{M}$. But this is a discrete set. So the only continuous representation is the trivial one $\mathbf{T}(t) = \mathbf{I}$ and the only derivation is $\mathbf{A} = \mathbf{O}$. This is important for approximation theory. If we want to approximate the equation $\dot{p}(t) = \mathbf{A}^* p(t)$ with $\mathbf{A}$ a derivation in a finite dimensional space, it is not possible to conserve both positivity and constant Lyapunov function. Of course, if we discretize the time too, it is possible to conserve a constant Lyapunov function, because continuous representations of the discrete semigroup $\mathbb{N}$ into the discrete set $\partial_e \mathcal{M}$ are possible. Note that only the $n!$ invertible elements have dense range (in finite dimensional space, dense range means full range).

We can conclude: If $\mathcal{Z}$ is finite, a constant Lyapunov function of type (11) is equivalent to reversibility in discrete time.

6.2 A manifold in $\mathbb{R}_n$

Let $\mathcal{Z}$ be a compact (if necessary suitably compactified) manifold in $\mathbb{R}_n$. This is the typical situation for evolution equations like (2). A fully rigorous description of a generator, satisfying the positive minimum principle, seems to be unknown so far, but the structure of such operators on inner points is known. It can be shown (see, e.g., [5, 12]) that a generator of a positive semigroup has the following structure ($g \in D(\mathbf{A})$ is twice continuously differentiable and z is an inner point of $\mathcal{Z}$):

$$\big(\mathbf{A}g\big)(z) \;=\; \sum_{i=1}^{n} a_i(z) \frac{\partial}{\partial z_i} g + \sum_{i,j=1}^{n} b_{ij}(z) \frac{\partial^2}{\partial z_i \partial z_j} g + \fint_{\mathcal{Z}} \Big(g(z') - g(z)\Big) Q(z, dz') \;.$$

Here, $a_i(z)$, $b_{ij}(z)$ and $Q(z, A) \geq 0$ are suitable coefficient functions on $\mathcal{Z}$ and $\mathcal{Z} \times \mathcal{B}(\mathcal{Z})$, resp., with a nonnegative matrix $(b_{ij}(z))$ (in the sense of bilinear forms in Hilbert spaces). The integral is to be understood as a principle value integral

$$\fint_{\mathcal{Z}} \Big(g(z') - g(z)\Big) Q(z, dz') = \lim_{\varepsilon \to 0} \int_{\mathcal{Z} \setminus B_\varepsilon(z)} \Big(g(z') - g(z)\Big) Q(z, dz')$$

if $Q(z, A)$ is singular for $z \in A$, where $B_\varepsilon(z)$ is a ball with radius ε and center at z. Sometimes, such operators are called Waldenfels operators and can be written as regularised integral operators, pseudo differential operators or operators with fractional derivatives (cf. [12, 6, 8]).

So far, necessary and sufficient conditions for the regularity of the coefficients a, b and Q and a general description of the value of $\mathbf{A}$ at the boundary are unknown. Special cases can be found in [12].

On the other hand, operators of this kind satisfy the positive minimum principle: If z_+ – the point of the sup of some $g(z) \in D(\mathbf{A})$ – is an inner point of $\mathcal{Z}$, then the first term of $(\mathbf{A}g)(z_+)$ is =0 because of $\frac{\partial}{\partial z_i} g = 0$, the second term is ≤ 0, because it is the trace of the product of a nonnegative matrix $(b_{ij}(z))$ and a non positive definite symmetric matrix $\frac{\partial^2}{\partial z_i \partial z_j} g$, and the third term is ≤ 0, because of $g(z') \leq g(z_+)$ and $Q(z_+, A) \geq 0$.

Formally, assuming $Q(z, dz') = Q(z, z')dz'$, the adjoint operator $\mathbf{A}^*$ acting on the derivative of measures w.r.t. the Lebesgue measure $W(z,t) = dp(t)/dz$ can be calculated (in general it is very difficult to describe $\mathbf{A}^*$ rigorously). The corresponding evolution equation reads as

$$\frac{\partial}{\partial t} W(z,t) = -\sum_{i=1}^{n} \frac{\partial}{\partial z_i}\Big(a_i(z) W(z,t)\Big) + \sum_{i,j=1}^{n} \frac{\partial^2}{\partial z_i \partial z_j}\Big(b_{ij}(z) W(z,t)\Big) + \quad (27)$$

$$+ \fint_{\mathcal{Z}} \Big(Q(z', z) W(z', t) - Q(z, z') W(z,t)\Big) dz' \,. \quad (28)$$

Defining a Lyapunov function $H(t)$ with a twice differentiable convex function, $F^* : \mathbb{R}_+ \to \mathbb{R}$, $F^*(1) = 0$ by

$$H(t) = H[W_1, W_2] = \int_{\mathcal{Z}} W_2(z,t) F^*\left(\frac{W_1(z,t)}{W_2(z,t)}\right) dz \geq 0 \,,$$

we can formally calculate

$$\frac{d}{dt} H(t) = -\int_{\mathcal{Z}} \sum_{i,j=1}^{n} F^{*\prime\prime}\left(\frac{W_1(z,t)}{W_2(z,t)}\right)\left(\frac{W_2}{W_1}\frac{\partial}{\partial z_i}\frac{W_1}{W_2}\right)\left(\frac{W_2}{W_1}\frac{\partial}{\partial z_j}\frac{W_1}{W_2}\right) b_{ij}(z) W_2(z,t) dz -$$

$$- \iint_{\mathcal{Z}\mathcal{Z}} \Big(F^*(\theta) - F^*(\theta') - F^{*\prime}(\theta')(\theta - \theta')\Big) W_2(z) Q(z, z') dz\, dz' \leq 0 \,, \quad (29)$$

with $\theta = \frac{W_1(z)}{W_2(z)}$, $\theta' = \frac{W_1(z')}{W_2(z')}$.

The classical Fokker-Planck equation (2) is the special case $Q \equiv 0$ of (28). In which sense this is a special case can be understood if we write $\mathbf{A}$ formally as a pseudo-differential operator (see, e.g., [11]):

$$\big(\mathbf{A}g\big)(z) = \frac{1}{2\pi} \int_{\mathcal{Z}} \int_{\mathbb{R}_n} e^{i\langle\lambda, z - z'\rangle} \alpha(z, \lambda) g(z') d\lambda dz'$$

with the symbol

$$\alpha(z, \lambda) = i\lambda a(z) - |\lambda|^2 b(z) + \fint_{\mathcal{Z}} Q(z, z')\big(e^{i\langle\lambda, z - z'\rangle} - 1\big) dz'$$

(here $\langle\cdot,\cdot\rangle$ and $|\cdot|$ are the usual Euclidean scalar product and norm in $\mathbb{R}_n$). Real and imaginary part of the symbol $\alpha(z,\lambda) = \alpha_+(z,\lambda) + i\alpha_-(z,\lambda)$ are odd and even functions of λ, $\alpha_+(z,-\lambda) = \alpha_+(z,\lambda)$, $\alpha_-(z,-\lambda) = -\alpha_-(z,\lambda)$. They have the growth properties $|\alpha_-(z,\lambda)| \lesssim |\lambda|$ and $|\alpha_+(z,\lambda)| \lesssim |\lambda|^2$ for $\lambda \to \infty$. So the integral operator can be understood as a sum of fractional differential operators $\frac{\partial^\alpha}{\partial z^\alpha}$ and $\frac{\partial^\beta}{\partial |z|^\beta}$ with $0 \leq \alpha < 1$ and $0 \leq \beta < 2$. Such operators for evolution problems investigated in the last years [6, 8]. The classical Fokker-Planck equation is the limiting case $\alpha \to 1$ and $\beta \to 2$ of the integral operator.

Now let us consider the case when $H(t)$ is constant in time. This is the case if $\mathbf{T}(t) \in \partial_e \mathcal{M}$, i.e. for every $t \geq 0$, there is a continuous function φ_t with $\mathbf{T}(t)g = g \circ \varphi_t$. From the semigroup property $\mathbf{T}(t_1 + t_2) = \mathbf{T}(t_1)\mathbf{T}(t_2)$, we conclude $\varphi_{t_1+t_2} = \varphi_{t_2} \circ \varphi_{t_1}$ i.e. φ_t is a continuous semi-flow. This semi-flow is the solution $z(t) = \varphi_t(z(0))$ of some dynamical system

$$\begin{array}{rcl} \dot{z}_1 & = & \Phi_1(z_1, ..., z_n) \\ \cdots & \cdot & \cdots \\ \dot{z}_n & = & \Phi_n(z_1, ..., z_n) \end{array}$$

in $\mathcal{Z}$ with a suitable function $\Phi : \mathcal{Z} \to \mathcal{Z}$. The corresponding Liouville equation reads

$$\frac{\partial}{\partial t} W(z,t) = -\sum_{i=1}^{n} \frac{\partial}{\partial z_i} \big(\Phi_i(z) W(z,t)\big) .$$

So, the generator of $\mathbf{T}(t)$

$$\big(\mathbf{A}g\big)(z) = \sum_{i=1}^{n} \Phi_i(z) \frac{\partial}{\partial z_i} g(z)$$

consists only of first derivatives and is a derivation, as expected. $H(t)$ =const can be formally derived from (29), setting $b_{ij} \equiv 0$ and $Q \equiv 0$.

From the one-to-one correspondence between deterministic Markov semigroups and semi-flows we can conclude: For any dynamical system, the Lyapunov function $H(t) = H[p(t), q(t)]$ is constant in time, where $p(t)$ and $q(t)$ are two solutions of the corresponding Liouville equation. Of course, a dynamical system has only a finite number of time-invariant integrals, whereas for the Liouville equation any function $H(t)$ generated by any convex $F(x)$ is a time-invariant integral. An arbitrary dynamical system does not necessarily have to be time reversible, whereas the Lyapunov functions of type (11) are always constant in time. Decreasing and time constant Lyapunov functions distinguish between random and deterministic problems and not between time-irreversible and time-reversible problems, as may be expected.

References

[1] 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.

[2] A. Arnold, P. Markowich, G. Toscani, A. Unterreiter. On logarithmic Sobolev inequalities, Csiszar-Kullback inequalities, and the rate of convergence to equilibrium for Fokker-Planck type equations. Preprint. Berlin, Techn. Univ., Fachbereich Mathematik (3), Preprint-Reihe Mathematik, 592, 1998.

[3] H. Brezis.*Operateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert.* North-Holland Pub Co. 1973.

[4] N. Dunford, J.T. Schwartz. *Theory of linear operators. Part I. General Theory.* Interscience Publishers 1958.

[5] C.W. Gardiner. *Handbook of Stochastic Methods.* Springer Series in Synergetics. Volume 13. Second Edition. 1985.

[6] R. Gorenflo, F. Mainardi. Fractional calculus and stable probability distributions. Arch. Mech., 50(3), 377-388 (1998).

[7] S. Kaplan. *The Bidual of $C(X)$ I.* North Holland, 1985.

[8] R. Metzler. Non-homogeneous random walks, generalised master equations, fractional Fokker-Planck equations, and the generalised Kramers-Moyal expansion. Eur. Phys. J. B 19, 249-258 (2001).

[9] J. van Neerven. *The Adjoint of a Semigroup of Linear Operators.* Springer, 1992.

[10] Z. Semadeni. *Banach Spaces of Continuous Functions.* Warszawa 1971.

[11] H. Stephan. *Nichtgleichgewichtsprozesse. Direkte und inverse Probleme.* Aachen: Shaker 1996.

[12] K. Taira. *Semigroups, boundary value problems and Markov processes.*, Springer, 2004

[13] C. Tzanakis, A.C. Grecos. Classical Markovian Kinetic Equations: Explicit Form and H-Theorem. arXiv:physics/9708031.