

Duality results for the joint spectral radius and transient behavior

Elmar Plischke
Zentrum für Technomathematik
Universität Bremen
Germany
email: elmar.plischke@uni-bremen.de

Fabian Wirth
Institut für Mathematik
Universität Würzburg
Germany
email: wirth@mathematik.uni-wuerzburg.de

Abstract For linear inclusions in discrete or continuous time several quantities characterizing the growth behavior of the corresponding semigroup are analyzed. These quantities are the joint spectral radius, the initial growth rate and (for bounded semigroups) the transient bound. It is discussed how these constants relate to one another and how they are characterized by various norms. A complete duality theory is developed in this framework, relating semigroups and dual semigroups and extremal or transient norms with their respective dual norms.

1 Introduction

In this paper we discuss duality relations between growth rates, i.e., joint spectral radii, initial growth rates and transient bounds of semigroups generated by linear inclusions. Our approach is based on the classical concept of a dual norm, see e.g. [1], as norms play a role in the description of all three quantities.

The joint spectral radius as introduced by Rota and Strang [2] characterizes the exponential growth rate of a linear semigroup generated by a compact set of matrices in discrete or continuous time, see also [3, 4]. One important tool in the study of the joint spectral radius are the extremal norms associated to the semigroup. These norms characterize the joint spectral in finite time. The investigation of extremal norms has begun with results by Barabanov [5, 3] and Kozyakin [6] and has recently lead to the discovery of interesting properties of the joint spectral radius [7, 8].

The initial growth rate is another quantity that describes an aspect of the exponential growth of a semigroup. It was introduced by Dahlquist [9] and Lozinskii [10] under the names “logarithmic norm” or “logarithmic derivative”, Vidyasagar [11] uses the name “matrix measure”. Its study has been motivated by problems in numerics, where the concept is used to obtain estimates for the accuracy of ODE solvers, or calculating the sensitivity of matrix exponentials. Recently, the concept of the initial growth rate has been extended to systems of differential algebraic equations (DAEs), see [12].

Finally, the transient bound characterizes the overshoot of the semigroup. If a semigroup is exponentially stable, it may still happen, that over a short horizon trajectories move far away from the origin. This may be undesirable in practice, so that good bounds for the transients are a useful tool in understanding the dynamics of the semigroup. There are different names associated with this phenomenon. In [13, 14, 15] the authors introduce the overshoot measure of a linear semigroup (which is the transient bound by another name) and derive several estimates for it. In particular in relation to nonlinear control problems the peaking phenomenon has been widely discussed [16]. While the practical problem is classical, sharp bounds for the transient behavior are elusive and there is no complete theory on this issue. A number of results may be found in [17] and [18, Chapter 5]. It is known that the transient behavior may be characterized via appropriate norms. Also in this setting we develop a duality theory.

The main results of the paper concern duality notions between the different constants describing the growth of linear semigroups. To this end, we define for a given linear inclusion the corresponding dual inclusion. It is easy to see that joint spectral radius and transient bound for the dual semigroup are the same as for the original semigroup. Furthermore, two main results are obtained concerning dual norms: In the literature two specific constructions for extremal norms may be found, due to Barabanov [5, 3] and Protasov [19]. We show that these constructions are dual to one another in the sense, that if v is a Barabanov norm for a semigroup, then the dual norm v^* is a Protasov norm for the dual semigroup. A similar situation occurs for transient norms. In the literature there is a standard technique for constructing a norm that characterizes the transient bound, which to the best of our knowledge goes back to Feller, [20]. We show that there is a dual construction to Feller's, which appears to be new in this context. Again a norm is a Feller norm for a bounded semigroup if and only if its dual norm is of this dual type for the dual semigroup.

The paper is organized as follows. After introducing the basic concepts in Section 2, in Section 2.3 we discuss initial growth rates for the continuous time case and give interpretation in terms of subgradients of norms. Section 3 is devoted to some easy bounds for the transient behavior of linear semigroups. In Section 4 several important norms are introduced. Extremal norms are those that characterize the joint spectral radius, in that the initial growth rate is equal to the joint spectral radius with respect to these norms. Similarly, transient norms characterize the transient behavior in terms of their eccentricity, which we define below. Section 5 is devoted to duality results concerning the joint spectral radius. Dual semigroups are introduced and it is shown that some construction procedures for extremal norms are dual. In Section 5.2 similar results are obtained for transient norms.

2 Linear Inclusions

In the following we study linear inclusions in continuous and discrete time. When necessary we specify the time set \mathbb{T} , which is thus either equal to $\mathbb{R}_+ := [0, \infty)$ or to \mathbb{N} .

Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Given a compact set $\emptyset \neq \mathcal{M} \subset \mathbb{K}^{n \times n}$ and the time set $\mathbb{T} = \mathbb{N}$ we consider the discrete inclusion

$$\begin{aligned} x(t+1) &\in \{Ax(t) \mid A \in \mathcal{M}\}, \quad t \in \mathbb{N} \\ x(0) &= x_0 \in \mathbb{K}^n. \end{aligned} \tag{1}$$

A sequence $\{x(t)\}_{t \in \mathbb{N}}$ is called a solution of (1) with initial condition x_0 if $x(0) = x_0$ and if for all $t \in \mathbb{N}$ there exists an $A(t) \in \mathcal{M}$ such that $x(t+1) = A(t)x(t)$. Associated to (1) we consider the sets of products of length t given by

$$\mathcal{S}_t := \{A(t-1) \dots A(0) \mid A(s) \in \mathcal{M}, s = 0, \dots, t-1\},$$

where we set $\mathcal{S}_0 = \{I_n\}$ for $t = 0$, and the semigroup given by $\mathcal{S} := \bigcup_{t=0}^{\infty} \mathcal{S}_t$.

In a similar manner we obtain a semigroup in the continuous time case. Given a compact set $\emptyset \neq \mathcal{M} \subset \mathbb{K}^{n \times n}$ and the time set $\mathbb{T} = \mathbb{R}_+$, we consider the semigroup generated by a differential inclusion

$$\dot{x} \in \{Ax(t) \mid A \in \mathcal{M}\}. \tag{2}$$

A function $x : \mathbb{R}_+ \rightarrow \mathbb{K}^n$ is called solution of (2) if it is absolutely continuous and satisfies $\dot{x}(t) \in \{Ax(t) \mid A \in \mathcal{M}\}$ almost everywhere. Equivalently, $x(\cdot)$ is the solution of a linear time-varying differential equation

$$\dot{x} = A(t)x(t) \tag{3}$$

for an appropriately chosen measurable map $A : \mathbb{R}_+ \rightarrow \mathcal{M}$. We denote the evolution operators of (3) by $\Phi_A(t, s)$. The set of time t transition operators is then given by

$$\mathcal{S}_t := \{\Phi_A(t, 0) \mid A : \mathbb{R}_+ \rightarrow \mathcal{M} \text{ measurable}\}.$$

Again $\mathcal{S} = \bigcup_{t \in \mathbb{T}} \mathcal{S}_t$ defines a semigroup. In the sequel, we will tacitly assume that \mathcal{S} is generated by an inclusion of the form (1) or of the form (2), if we speak of a semigroup $(\mathcal{S}, \mathbb{T})$. Moreover it is assumed, that \mathcal{M} is convex if $\mathbb{T} = \mathbb{R}_+$. Together with our compactness assumption this ensures that the sets $\mathcal{S}_t, t \in \mathbb{T}$ are compact. Note that if $\mathbb{T} = \mathbb{R}_+$ we have by classical relaxation results, that

$$\text{cl } \mathcal{S}_t(\mathcal{M}) = \mathcal{S}_t(\text{conv } \mathcal{M}),$$

so that by going over to the convex hull of \mathcal{M} we do not alter the semigroup significantly.

In the following we wish to introduce several quantities that characterize the growth behavior of a semigroup \mathcal{S} . These are the *joint spectral radius* (or maximal Lyapunov exponent, or Lyapunov indicator), that characterizes the long term exponential growth behavior, the *initial growth rate* and the *transient bound*.

Remark 2.1 *Whenever discrete time and continuous time systems are considered simultaneously, the dilemma appears that in discrete time it is natural to denote exponential growth in the form r^t , while in continuous time it of interest to consider $e^{\log r t}$. To keep notation short we have opted for a unified notation using the discrete time approach.*

2.1 Joint Spectral Radius

We begin our definitions with the joint spectral radius. Let $r(A)$ denote the spectral radius of A and let $\|\cdot\|$ be some operator norm on $\mathbb{K}^{n \times n}$. Define for $t \in \mathbb{N}$

$$\begin{aligned} \bar{\rho}_t(\mathcal{M}) &:= \sup\{r(S)^{1/t} \mid S \in \mathcal{S}_t\}, \\ \hat{\rho}_t(\mathcal{M}) &:= \sup\{\|S\|^{1/t} \mid S \in \mathcal{S}_t\}. \end{aligned} \tag{4}$$

The *joint spectral radius* is defined by

$$\rho(\mathcal{M}) := \limsup_{t \rightarrow \infty} \bar{\rho}_t(\mathcal{M}) = \lim_{t \rightarrow \infty} \hat{\rho}_t(\mathcal{M}).$$

By the results in [21] the above quantity is well-defined. Note in particular that it does not depend on the choice of the norm $\|\cdot\|$. A further characterization of ρ is given by

$$\rho(\mathcal{M}) = \inf\{\rho \in \mathbb{R} \mid \exists M_\rho : \forall t \in \mathbb{T}, S \in \mathcal{S}_t : \|S\| \leq M_\rho \rho^t\}. \tag{5}$$

2.2 Initial Growth Rate

On the other hand, the *initial growth rate* of \mathcal{M} depends on the norm under consideration. Given a norm $\|\cdot\|$ on \mathbb{K}^n we define for the discrete time case $\mathbb{T} = \mathbb{N}$ the initial growth rate by

$$\mu(\mathcal{M}) := \sup\{\|A\| \mid A \in \mathcal{M}\}. \tag{6}$$

In the continuous time case $\mathbb{T} = \mathbb{R}_+$ we set for an individual matrix $A \in \mathbb{K}^{n \times n}$

$$\mu(A) := \exp\left(\lim_{t \rightarrow 0} \frac{1}{t} \log \|e^{At}\|\right) \tag{7}$$

and for $\mathcal{M} \subset \mathbb{K}^{n \times n}$ we let $\mu(\mathcal{M}) := \sup\{\mu(A) \mid A \in \mathcal{M}\}$.

If $\mu(A) \leq 1$ then A is called *dissipative*. The following result extends statements in [11] and shows that μ is conceptually closely related to ρ , if we compare (5) with (8).

Proposition 2.2 *Let $\mathbb{T} = \mathbb{N}, \mathbb{R}_+$. Suppose $\|\cdot\|$ is an operator norm on $\mathbb{K}^{n \times n}$. For $\mathcal{M} \subset \mathbb{K}^{n \times n}$ the initial growth rate $\mu(\mathcal{M})$ is the least upper exponential bound for $\|S\|$, i.e.,*

$$\mu(\mathcal{M}) = \min\{\mu \in \mathbb{R} \mid \forall t \geq 0, S \in \mathcal{S}_t : \|S\| \leq \mu^t\}. \tag{8}$$

Proof. For the discrete case, this is immediately clear from the definition of μ in (6) and the discrete semigroup $\mathcal{S}_t, t \in \mathbb{N}$. For the continuous case, we begin with the case of a matrix $A \in \mathcal{M}$. Let $\|e^{At}\| \leq \mu^t$ for all $t \in \mathbb{R}_+$. Then $\frac{1}{t} \log \|e^{At}\| \leq \log \mu$. Hence $\mu(A) \leq \mu$. On the other hand, we have $\|e^{At}\| \leq \mu(A)^t$ for all $t \in \mathbb{R}_+$. This follows from the semigroup property as we have for all $t \in \mathbb{R}_+$,

$$\|e^{At}\| \leq \|e^{As}\|^{\frac{t}{s}} \quad \text{for all } s \in \mathbb{R}_+ \text{ such that } \frac{t}{s} \in \mathbb{N}.$$

As $\mu(A) = \lim_{s \rightarrow 0} \|e^{As}\|^{\frac{1}{s}}$ is well-defined, $\|e^{At}\| \leq \mu(A)^t$ holds for all $t \in \mathbb{R}_+$. Now let $S \in \mathcal{S}_t$ be arbitrary. Then for any $\varepsilon > 0$ there exist $0 < t_1 < t_2 < \dots < t_n = t$ and matrices $A_i \in \mathcal{M}, i = 1, \dots, n$ such that

$$\left\| S - e^{A_n(t-t_{n-1})} \dots e^{A_2(t_2-t_1)} e^{A_1 t_1} \right\| < \varepsilon$$

(see e.g. [4]). Hence

$$\|S\| \leq \left\| e^{A_n(t-t_{n-1})} \dots e^{A_2(t_2-t_1)} e^{A_1 t_1} \right\| + \varepsilon \leq \mu(\mathcal{M})^t + \varepsilon.$$

As $\varepsilon > 0$ was arbitrary the assertion follows. \square

Note, that from comparing (5) with (8) it is immediate, that for any norm $\|\cdot\|$ on \mathbb{K}^n we have

$$\rho(\mathcal{M}) \leq \mu_{\|\cdot\|}(\mathcal{M}).$$

For stable linear inclusions (those that generate a bounded semigroup) we define the transient bound by

$$M_0(\mathcal{M}) := \sup\{\|S\| \mid S \in \mathcal{S}(\mathcal{M})\}. \quad (9)$$

In particular, if \mathcal{M} is a set of dissipative matrices then the semigroup $\mathcal{S}(\mathcal{M})$ generated by \mathcal{M} is a contraction semigroup with $M_0(\mathcal{M}) = 1$.

As we have noted in the introduction, the transient bound has received various different names in the literature, such as overshoot measure. It is related to the concepts of overshooting behavior and the peaking effect.

In Section 5 we will discuss the interplay of ρ, μ and M_0 with dual semigroups. Before doing so, we discuss initial growth rates in more detail.

2.3 Initial Growth Rate and Dual Norms

In this section we review some known results on initial growth rates. We also present a characterization of initial growth rates in terms of subgradients of the norm. While this characterization is not difficult or surprising from the point of view of nonsmooth analysis, the remark does not appear to have been made in this context.

The following summarizes some known facts about initial growth rates for the case $\mathbb{T} = \mathbb{R}_+$, [22, 11].

Proposition 2.3 *Let $\mathbb{T} = \mathbb{R}_+$ and fix a norm $\|\cdot\|$ with unit ball \mathcal{B} . Given matrices $A, A' \in \mathbb{K}^{n \times n}$ and scalars $z \in \mathbb{K}, \alpha \in \mathbb{R}$, the initial growth rate $\mu(\cdot)$ satisfies*

$$(i) \quad -\log \mu(-A) \leq \operatorname{Re} \lambda \leq \log \mu(A), \quad \lambda \in \sigma(A)$$

$$(ii) \quad \log \mu(\alpha A) = |\alpha| \log \mu((\operatorname{sgn} \alpha)A)$$

$$(iii) \quad |\log \mu(A)| \leq \|A\|$$

$$(iv) \quad \log \mu(A + zI) = \log \mu(A) + \operatorname{Re} z$$

$$(v) \quad \log \mu(A + A') \leq \log \mu(A) + \log \mu(A')$$

(vi)

$$\log \mu(A) = \lim_{t \downarrow 0} t^{-1} (\|I + tA\| - 1). \quad (10)$$

(vii) the unit ball \mathcal{B} is forward-invariant under the flow of $\dot{x} = Ax$ if and only if $\mu(A) \leq 1$.

The initial growth rate can also be interpreted as a *dissipativity radius*, as the following result implies.

Lemma 2.4 *Let $\mathbb{T} = \mathbb{R}_+$, $\mathcal{M} \subset \mathbb{K}^{n \times n}$, $N \in \mathbb{K}^{n \times n}$. Suppose that N is dissipative with $\mu(N) \leq 1$. If $\text{dist}(\mathcal{M}, N) \leq \delta$ then*

$$\log \mu(\mathcal{M}) \leq \log \mu(N) + \delta.$$

Proof. Let $t \in \mathbb{R}_+$, $S \in \mathcal{S}_t$. By Proposition 2.2 we have to show that $\|S\| \leq (\mu(N)e^\delta)^t$. For $A \in \mathcal{M}$ we have $\|a - N\| \leq \delta$ and so Proposition 2.3 (iii) and (v) yield the estimate

$$\|e^{At}\| = \left\| e^{(A-N+N)t} \right\| \leq \mu(A-N)^t \mu(N)^t \leq (\mu(N)e^\delta)^t.$$

□

Hence if N is dissipative, matrix perturbations $\Delta \in \mathbb{K}^{n \times n}$ with norm $\|\Delta\| < |\log \mu(N)|$ will not destroy dissipativity. Note that the corresponding statement for the discrete time case is trivial.

We now discuss the characterization of initial growth rates using subgradients. To this end we need subgradients of norms, which are given through the dual norm. Recall that for a fixed norm v on \mathbb{K}^n the *dual norm* is defined by

$$v^*(x) := \max\{|\langle l, x \rangle| \mid v(l) \leq 1\}. \quad (11)$$

Note that $v = (v^*)^*$. A pair of vectors $l, x \in \mathbb{K}^n$ is called *dual pair*, if $\langle l, x \rangle = v(x)v^*(l)$. See [1] for further details.

Let us denote the unit sphere of the vector norm v by $\mathcal{T}_v = \{x \in \mathbb{K}^n \mid v(x) = 1\}$ and the unit sphere of its dual norm by \mathcal{T}_v^* . We want to relate the initial growth rate to pairs of dual vectors. To this end we need to recall the concept of a subdifferential of a convex function. Given a convex function $f : \mathbb{K}^n \rightarrow \mathbb{R}$ the subdifferential $\partial_P f(x)$ is defined by

$$\partial_P f(x) := \{l \in \mathbb{K}^n \mid f(y) \geq f(x) + \text{Re} \langle l, y - x \rangle, \forall y \in \mathbb{K}^n\}.$$

We see that for a norm v on \mathbb{K}^n that the elements of $\partial_P v(x)$ are normals for supporting hyperplanes in x of the convex set $v(x)\mathcal{B}_v := \{y \in \mathbb{K}^n \mid v(y) \leq v(x)\}$. It follows from [23, Corollary 23.5.3] that

$$\partial_P v(x) = \{l \in \mathbb{K}^n \mid v^*(l) = 1, \langle l, x \rangle = v(x)\}. \quad (12)$$

Proposition 2.5 *Let $\mathbb{T} = \mathbb{R}_+$. Given a norm v on \mathbb{K}^n , the associated initial growth rate μ_v of a matrix A is given by*

$$\log \mu_v(A) = \max\{\text{Re} \langle l, Ax \rangle \mid (l, x) \in \mathcal{T}_v^* \times \mathcal{T}_v \text{ is a dual pair of } v\}.$$

Proof. Let us consider the alternative characterization of the initial growth rate given by (10), i.e.

$$\log \mu(A) = \lim_{t \downarrow 0} t^{-1} (\|I + tA\| - 1).$$

For a fixed $t \in \mathbb{R}_+$ there are unit vectors $x \in \mathcal{T}_v$ and $y \in \mathcal{T}_v^*$ such that $v(I + At) = \text{Re} \langle y, (I + At)x \rangle$. Hence

$$\frac{1}{t}(v(I + At) - 1) = \text{Re} \langle y, Ax \rangle + \frac{1}{t}(\text{Re} \langle y, x \rangle - 1) \leq \text{Re} \langle y, Ax \rangle$$

where equality holds if y and x form a dual pair. On the other hand, for all $t \in \mathbb{R}_+$ and all dual pairs l, x with $l \in \mathcal{T}_v^*$ the subdifferential inequality is satisfied,

$$v((I + At)x) \geq \text{Re} \langle l, x \rangle + t \text{Re} \langle l, Ax \rangle = v(x) + t \langle l, Ax \rangle.$$

Passing to the limit over all $x \in \mathcal{T}_v$ and to $t \rightarrow 0$ yields the desired formula for $\mu_v(A)$. □

For $A \in \mathbb{K}^{n \times n}$ we denote by A^* the dual matrix with respect to the standard scalar product, that is, $A^* = \overline{A}^T$. Proposition 2.5 can be used to derive, in a simple manner, a (well-known) formula for the initial growth rate μ_2 with respect to the Euclidean norm.

Lemma 2.6 *The Euclidean initial growth rate of $A \in \mathbb{K}^{n \times n}$ is given by $\log \mu_2(A) = \frac{1}{2} \lambda_{\max}(A + A^*)$.*

Proof. For the Euclidean norm all vectors are self-dual, as $\langle x, x \rangle = \|x\|_2^2$. Hence

$$\begin{aligned} \log \mu_2(A) &= \max\{\operatorname{Re} \langle x, Ax \rangle \mid \|x\|_2 = 1\} \\ &= \frac{1}{2} \max_{x \neq 0} \frac{\langle x, (A+A^*)x \rangle}{\langle x, x \rangle} = \frac{1}{2} \lambda_{\max}(A + A^*) \end{aligned}$$

where the last equality follows from the Rayleigh-Ritz theorem, see [1]. \square

3 Bounds for the Transient Behavior

In this section we discuss how to obtain transient bounds via the initial growth rate. To this end we introduce the eccentricity to be able to compare two norms.

Definition 3.1 *Suppose that v and $\|\cdot\|$ are norms on \mathbb{K}^n . The eccentricity of $v(\cdot)$ with respect to $\|\cdot\|$ is given by*

$$\operatorname{ecc}(v) = \operatorname{ecc}(v, \|\cdot\|) := \frac{\max_{\|x\|=1} v(x)}{\min_{\|x\|=1} v(x)}. \quad (13)$$

The eccentricity measures the deformation of the unit balls of two norms w.r.t. each other. It is easy to see that

$$\operatorname{ecc}(v, \|\cdot\|) = \frac{\max_{x \neq 0} \frac{v(x)}{\|x\|}}{\min_{x \neq 0} \frac{v(x)}{\|x\|}} = \frac{\max_{x \neq 0} \frac{\|x\|}{v(x)}}{\min_{x \neq 0} \frac{\|x\|}{v(x)}} = \operatorname{ecc}(\|\cdot\|, v). \quad (14)$$

This notion can be used to compare the transient behavior under different norms.

Corollary 3.2 *Let $\mathcal{M} \subset \mathbb{K}^{n \times n}$ and consider two norms $\|\cdot\|, v(\cdot)$ on \mathbb{K}^n . Then for all $S \in \mathcal{S}_t(\mathcal{M}, \mathbb{T})$ we have*

$$\|S\| \leq \operatorname{ecc}(v, \|\cdot\|) \mu_v(\mathcal{M})^t, \quad t \in \mathbb{T}. \quad (15)$$

Proof. By Proposition 2.2 we have the exponential estimate $v(S) \leq \mu_v(\mathcal{M})^t$ for all $S \in \mathcal{S}_t$, $t \geq 0$. For all $y \in \mathbb{K}^n$, $y \neq 0$ we get

$$\min_{\|x\|=1} v(x) \leq v\left(\frac{y}{\|y\|}\right) \leq \max_{\|x\|=1} v(x). \quad (16)$$

This implies for the associated operator norms that for all $T \in \mathbb{K}^{n \times n}$,

$$\begin{aligned} \|T\| &= \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} \leq \sup_{x \neq 0} \frac{(\min_{\|x\|=1} v(x))^{-1} v(Tx)}{(\max_{\|x\|=1} v(x))^{-1} v(x)} \\ &= \operatorname{ecc}(v)v(T). \end{aligned} \quad (17)$$

Setting $T = S$, $S \in \mathcal{S}_t$ yields the desired result. \square

4 Extremal and Transient Norms

4.1 Extremal Norms

Both joint spectral radius and transient bound are closely related to specific norms.

Definition 4.1 Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$, $\mathbb{T} = \mathbb{N}, \mathbb{R}_+$ and let $(\mathcal{S}, \mathbb{T})$ be a semigroup in $\mathbb{K}^{n \times n}$.

(i) A norm v on \mathbb{K}^n is called extremal for \mathcal{S} if

$$\mu_v(\mathcal{S}) = \rho(\mathcal{S}). \quad (18)$$

(ii) an extremal norm v on \mathbb{K}^n is called Barabanov norm corresponding to \mathcal{S} if for all $x \in \mathbb{K}^n$, $t \in \mathbb{T}$ there is an $S \in \text{cl } \mathcal{S}_t$ such that

$$v(Sx) = \rho(\mathcal{S})^t v(x). \quad (19)$$

(iii) a norm v on \mathbb{K}^n is called Protasov norm corresponding to \mathcal{S} if the unit ball \mathcal{B}_v of v satisfies

$$\rho(\mathcal{S})^t \mathcal{B}_v = \text{conv cl } \mathcal{S}_t \mathcal{B}_v, \quad \forall t \in \mathbb{T}. \quad (20)$$

Remark 4.2 It has become common to use the name Barabanov norms because they have been introduced in [5], [3]. A sufficient criterion for their existence is that \mathcal{M} is irreducible, i.e., only the trivial subspaces $\{0\}$ and \mathbb{K}^n are invariant under all matrices $A \in \mathcal{M}^1$. It is clear that Barabanov norms are extremal. The converse is false.

In a similar vein, we use the name Protasov norm, because these norms have been introduced in [19]. It is easy to see directly that they are extremal. This also follows from the duality result Theorem 5.2.

The next lemma provides a construction for Protasov norms.

Lemma 4.3 Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$, $\mathbb{T} = \mathbb{N}, \mathbb{R}_+$ and let $(\mathcal{S}, \mathbb{T})$ be an irreducible semigroup in $\mathbb{K}^{n \times n}$. If v is an extremal norm on \mathbb{K}^n with unit ball \mathcal{B} then

$$\tilde{\mathcal{B}} = \bigcap_{t \in \mathbb{T}} \text{conv cl } \bigcup_{S \in \text{cl } \mathcal{S}_t} \rho(\mathcal{S})^{-t} S \mathcal{B} \quad (21)$$

is the unit ball of a Protasov norm for \mathcal{S} .

Proof. We may assume $\rho(\mathcal{S}) = 1$, because of the normalizing factor in the right hand side of (21). As v is extremal we have $S\mathcal{B} \subset \mathcal{B}$ for all $t \in \mathbb{T}$, $S \in \mathcal{S}_t$. This implies for $t_1 < t_2 \in \mathbb{T}$ that

$$\bigcup_{S \in \text{cl } \mathcal{S}_{t_2}} S \mathcal{B} \subset \bigcup_{S \in \text{cl } \mathcal{S}_{t_1}} S \mathcal{B} \subset \mathcal{B}.$$

Thus $\tilde{\mathcal{B}}$ is nonempty as a descending intersection of compact sets. By construction it is convex and balanced. Also $\tilde{\mathcal{B}} \neq \{0\}$ by extremality of v and irreducibility of \mathcal{S} . (Recall that for an irreducible semigroup we have $\rho(\mathcal{S}) < 0$ if all solutions of (1), resp. (2) converge to 0, see [5, 3].) From the definition it follows that $\tilde{\mathcal{B}} = \text{conv cl } \mathcal{S}_t \tilde{\mathcal{B}}$, $\forall t \in \mathbb{T}$. This implies in particular, that $\tilde{\mathcal{B}}$ is not contained in any hyperplane, because the semigroup is irreducible and otherwise $\text{span } \tilde{\mathcal{B}}$ would be a nontrivial invariant subspace. Hence, $\tilde{\mathcal{B}}$ is a level set of a norm which is a Protasov norm by construction. \square

¹The name irreducibility has been taken on from representation theory. This concept has been called nonsingularity in [5], quasicontrollability in [13] and has received further names in the literature.

4.2 Transient Norms

Just as in the case of the joint spectral radius, the transient behavior can be characterized using norms. The following definition follows the same ideas as Definition 4.1.

Definition 4.4 Let $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{T} = \mathbb{N}, \mathbb{R}_+$ and let $(\mathcal{S}, \mathbb{T})$ be a bounded semigroup in $\mathbb{K}^{n \times n}$ generated by $\mathcal{M} \subset \mathbb{K}^{n \times n}$. Consider a fixed vector norm $\|\cdot\|$ with unit ball \mathcal{B} .

(i) A norm v on \mathbb{K}^n is called transient for \mathcal{S} if

$$\mu_v(\mathcal{M}) \leq 1, \quad \text{and } \text{ecc}(v, \|\cdot\|) = M_0(\mathcal{M}). \quad (22)$$

(ii) The Feller norm corresponding to \mathcal{M} and $\|\cdot\|$ is defined by

$$\|x\|_{\mathcal{M}} = \sup_{S \in \mathcal{S}} \|Sx\|. \quad (23)$$

(iii) The convex-transient norm corresponding to \mathcal{M} and $\|\cdot\|$ is defined by its unit ball given by

$$\tilde{\mathcal{B}}_{\mathcal{M}} = \text{conv cl} \{Sx \mid S \in \mathcal{S}, x \in \mathcal{B}\} = \text{conv cl} \bigcup_{S \in \mathcal{S}} SB. \quad (24)$$

Remark 4.5 The name transient norm comes from the fact that the eccentricity of the norm characterizes the transient bound. We opted for the name Feller norm as the norm $x \mapsto \sup_{t \geq 0} \|T_t x\|$ for a one-parameter semigroup (T, \mathbb{R}_+) has been introduced in [20]. Note that $\mu_v(\mathcal{M}) \leq 1$ implies for any norm v that $\text{ecc}(v, \|\cdot\|) \geq M_0(\mathcal{M})$. We will now show that the Feller norm $\|\cdot\|_{\mathcal{M}}$ is a transient norm. The same statement for the convex-transient norm follows from our duality results.

It is easily verified that $\|\cdot\|_{\mathcal{M}}$ is indeed a norm. We claim that $\text{ecc}(\|\cdot\|_{\mathcal{M}}) = M_0(\mathcal{M})$. This may be seen as follows:

$$\text{ecc}(\|\cdot\|_{\mathcal{M}}, \|\cdot\|) = \frac{\sup_{\|x\|=1} \sup_{S \in \mathcal{S}} \|Sx\|}{\inf_{\|x\|=1} \sup_{S \in \mathcal{S}} \|Sx\|} = \frac{M_0(\mathcal{M})}{\inf_{\|x\|=1} \sup_{S \in \mathcal{S}} \|Sx\|},$$

but $\inf_{\|x\|=1} \sup_{S \in \mathcal{S}} \|Sx\| = 1$ because otherwise \mathcal{S} is not bounded. Furthermore, $\mu_{\mathcal{M}}(\mathcal{M}) \leq 1$ because

$$\|Sx\|_{\mathcal{M}} = \sup_{T \in \mathcal{S}} \|TSx\| \leq \sup_{T \in \mathcal{S}} \|Tx\| = \|x\|_{\mathcal{M}}. \quad (25)$$

More precisely, we have the following result for the initial growth rate with respect to the transient norm.

Lemma 4.6 Let $\mathcal{M} \subset \mathbb{K}^{n \times n}$ generate a bounded semigroup \mathcal{S} . Then the initial growth rate associated with the Feller norm satisfies $\mu_{\mathcal{M}}(\mathcal{M}) = \min\{\mu(\mathcal{M}), 1\}$.

Proof. If for the original norm $\mu(\mathcal{M}) \leq 1$, then, by Proposition 2.2, $\|Sx\| \leq \mu(\mathcal{M})^t \|x\| \leq \|x\|$ for all $x \in \mathbb{K}^n$ and all $S \in \mathcal{S}_t$. Hence $\|\cdot\|_{\mathcal{M}} = \sup_{S \in \mathcal{S}} \|S\cdot\| = \|\cdot\|$ and so $\mu_{\mathcal{M}}(\mathcal{M}) = \mu(\mathcal{M})$.

Now, if $\mu(\mathcal{M}) > 1$ there exist $x_0 \in \mathbb{K}^n$ and $S \in \mathcal{S}_{t_0}$, $t_0 > 0$ such that $\|Sx_0\| = \|x_0\|_{\mathcal{M}} > \|x_0\|$. By Proposition 2.2 this shows $\mu_{\mathcal{M}}(\mathcal{M}) \geq 1$ and so $\mu_{\mathcal{M}}(\mathcal{M}) = 1$ by (25). \square

We note the following property of the unit ball of the Feller norm for further reference.

Lemma 4.7 Suppose that $\mathcal{M} \subset \mathbb{K}^{n \times n}$ generates a bounded semigroup. Then the unit ball $\mathcal{B}_{\mathcal{M}}$ of the associated Feller norm $\|\cdot\|_{\mathcal{M}}$ satisfies

$$\mathcal{B}_{\mathcal{M}} = \bigcap_{S \in \mathcal{S}} S^{-1}\mathcal{B} \quad (26)$$

where \mathcal{B} is the unit ball of $\|\cdot\|$ and where we use the convention $S^{-1}\mathcal{B} = \{y \in \mathbb{K}^n \mid Sy \in \mathcal{B}\}$ in case that S is not invertible.

Proof. The set $\bigcap_{S \in \mathcal{S}} S^{-1}\mathcal{B}$ is a closed convex balanced set. Note that $I \in \mathcal{S}$, so that the intersection is bounded. Also, 0 is contained in the interior of this set by the boundedness of \mathcal{S} . Therefore $\bigcap_{S \in \mathcal{S}} S^{-1}\mathcal{B}$ is the unit ball of a norm. By definition, $x \in \mathcal{B}_{\mathcal{M}}$ holds if and only if for all $S \in \mathcal{S}$, $Sx \in \mathcal{B}$, or equivalently, $x \in S^{-1}\mathcal{B}$ which yields (26). \square

Finally, we note that if $\rho(\mathcal{M}) = 1$ and \mathcal{M} is bounded, then the Feller norm and the convex-transient norm are also extremal for \mathcal{M} , but in general different from a Barabanov or Protasov norm. Thus another way to construct extremal norms for irreducible inclusions, is to define the Feller or convex-transient norm for the normalized semigroup given by the finite time sets $\rho(\mathcal{M})^{-t}\mathcal{S}_t(\mathcal{M})$.

5 Duality

In this section we investigate the dual of semigroups. Also duality properties of Barabanov and Protasov norms as well as Feller and convex-transient norms are shown.

5.1 Duality and Extremal Norms

Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $\mathbb{T} = \mathbb{N}, \mathbb{R}_+$. Given a semigroup $(\mathcal{S}, \mathbb{T}) \subset \mathbb{K}^{n \times n}$, we define the dual semigroup to be

$$\mathcal{S}^* := \{S^* \mid S \in \mathcal{S}\},$$

where we assume in particular that we have

$$(\mathcal{S}^*)_t = \{S^* \mid S \in \mathcal{S}_t\}, \quad t \in \mathbb{T}.$$

It is then immediate that $\rho(\mathcal{S}) = \rho(\mathcal{S}^*)$. Let us briefly discuss how the generating sets can be constructed.

In the case $\mathbb{T} = \mathbb{N}$ let \mathcal{S} be the semigroup generated by $(\mathcal{M}, \mathbb{N})$. Then \mathcal{S}^* is generated by $(\mathcal{M}^*, \mathbb{N})$, where we define

$$\mathcal{M}^* := \{A^* \mid A \in \mathcal{M}\}.$$

In the continuous time case $\mathbb{T} = \mathbb{R}_+$, on the other hand, we consider the differential inclusion

$$\dot{x}(t) \in \{A^*(t) \mid A \in \mathcal{M}\}. \quad (27)$$

It is well known that for every $t \geq 0$ the evolution operators $\Phi(t, 0)$ of (3) of the form

$$\Phi(t, 0) = e^{A_k t_k} e^{A_{k-1} t_{k-1}} \dots e^{A_1 t_1},$$

where $A_j \in \mathcal{M}$, $j = 1, \dots, k$, $\sum_{j=1}^k t_j = t$, lie dense in $\mathcal{S}_t(\mathcal{M}, \mathbb{R}_+)$. It is obvious, that the dual of these operators lies dense in $\mathcal{S}_t(\mathcal{M}^*, \mathbb{R}_+)$, so $\text{cl } \mathcal{S}_t(\mathcal{M}, \mathbb{R}_+)^* = \text{cl } \mathcal{S}_t(\mathcal{M}^*, \mathbb{R}_+)$.

Thus we see that in both discrete and continuous time we may consider \mathcal{M}^* to be the generator of \mathcal{S}^* .

Our first duality result is the following, see also [3]. For the sake of completeness we include a proof.

Lemma 5.1 *Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$, $\mathbb{T} = \mathbb{N}, \mathbb{R}_+$. Let w be an extremal norm for $(\mathcal{S}, \mathbb{T})$. Then w^* is extremal for $(\mathcal{S}^*, \mathbb{T})$.*

Proof. Let $x \in \mathbb{K}^n$ and $S^* \in \mathcal{S}_t^*$ then

$$\begin{aligned} w^*(S^*x) &= \max\{\text{Re} \langle l, S^*x \rangle \mid w(l) \leq 1\} \\ &\leq \max\{\text{Re} \langle l, x \rangle \mid w(l) \leq \rho(\mathcal{S})^t\} \\ &= \rho(\mathcal{S})^t w^*(x). \end{aligned}$$

\square

We now show that Barabanov norms and Protasov norms are dual concepts.

Theorem 5.2 *Let $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{T} = \mathbb{N}, \mathbb{R}_+$ and let $(\mathcal{S}, \mathbb{T})$ be a semigroup. The norm v is a Barabanov norm for $(\mathcal{S}, \mathbb{T})$ if and only if the dual norm v^* is a Protasov norm for $(\mathcal{S}^*, \mathbb{T})$.*

Proof. We may assume that $\rho(\mathcal{S}) = 1$. Assume furthermore that v is a Barabanov norm and let $t \in \mathbb{T}$ and $x \in \mathbb{K}^n, v(x) = 1$ be arbitrary. By assumption there exists an $S \in \mathcal{S}_t$ such that $v(Sx) = v(x) = 1$. As $v = v^*$ it follows that

$$\begin{aligned} v(Sx) &= v(x) = \max \{ \operatorname{Re} \langle l, x \rangle \mid v^*(l) \leq 1 \} \\ &= \max \{ \operatorname{Re} \langle S^*l, x \rangle \mid v^*(l) \leq 1 \} \\ &\leq \max \{ \operatorname{Re} \langle l, x \rangle \mid l \in \operatorname{conv} \mathcal{S}_t^* \mathcal{B}_{v^*} \}. \end{aligned}$$

On the other hand by the extremality of the norm v^* we have

$$v(x) \geq \max \{ \operatorname{Re} \langle l, x \rangle \mid l \in \operatorname{conv} \mathcal{S}_t^* \mathcal{B}_{v^*} \}.$$

This implies equality throughout the calculation. As x was arbitrary this means that v is the dual norm of the norm with the unit ball $\operatorname{conv} \mathcal{S}_t^* \mathcal{B}_{v^*}$. As the dual norm is uniquely defined this implies that $\operatorname{conv} \mathcal{S}_t^* \mathcal{B}_{v^*}$ is the unit ball of v^* and as $t \in \mathbb{T}$ was arbitrary, condition (20) is satisfied. Hence, v^* is a Protasov norm.

Conversely, assume that (20) holds for v^* and let $t \in \mathbb{T}$ and $x \in \mathbb{K}^n, v(x) = 1$ be arbitrary. Then

$$\begin{aligned} v(x) &= \max \{ \operatorname{Re} \langle l, x \rangle \mid v^*(l) \leq 1 \} \\ &= \max \{ \operatorname{Re} \langle l, x \rangle \mid l \in \operatorname{conv} \mathcal{S}_t^* \mathcal{B}_{v^*} \} \\ &= \max \{ \operatorname{Re} \langle S^*l, x \rangle \mid v^*(l) \leq 1, S^* \in \mathcal{S}_t^* \} \\ &= \max \{ \operatorname{Re} \langle l, Sx \rangle \mid v^*(l) \leq 1, S \in \mathcal{S}_t \} \\ &= \max \{ v(Sx) \mid S \in \mathcal{S}_t \}, \end{aligned}$$

which shows that v is a Barabanov norm. □

A particularly satisfying situation occurs in the case that $\mathcal{S} = \mathcal{S}^*$ as in this case there is a natural candidate for an extremal norm. Assumptions concerning irreducibility are not necessary in this case. It has been noted before, that if \mathcal{M} consists of self-adjoint matrices then $\rho(\mathcal{M}, \mathbb{N}) = \max \{ \rho(A) \mid A \in \mathcal{M} \}$, see e.g. [5]. The following result extends this observation.

Proposition 5.3 *Let $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{T} = \mathbb{N}, \mathbb{R}_+$ and let $(\mathcal{S}, \mathbb{T})$ be a semigroup with $\mathcal{S} = \mathcal{S}^*$ then the Euclidean norm $\|\cdot\|_2$ is an extremal norm for \mathcal{S} .*

Proof. The case $\rho(\mathcal{S}) = 0$ can only occur in the case $\mathbb{T} = \mathbb{N}$ and is trivial because then $\mathcal{M} = \{0\}$ due to our assumption $\mathcal{S} = \mathcal{S}^*$. Thus without loss of generality we may assume $\rho(\mathcal{S}) = 1$. Assume the assertion is false, then for some $x \in \mathbb{K}^n, \|x\|_2 = 1$ and some $t \in \mathbb{T}$ there is an $S \in \mathcal{S}_t$ such that $\|Sx\|_2 > 1$. As $S^* \in \mathcal{S}$ we have $S^*S \in \mathcal{S}$ and $r(S^*S) \geq \langle S^*Sx, x \rangle = \|Sx\|_2^2 > 1$. This contradicts $\rho(\mathcal{S}) = 1$ and the contradiction proves the assertion. □

In the continuous time case $\mathbb{T} = \mathbb{R}_+$ we also point out the following reformulation of a result by Barabanov that will turn out to be useful later. Here we relate for the continuous time case extremality properties of a norm with the infinitesimal growth of the trajectories of the system. Furthermore by [23, Theorem 24.4] the set-valued map $x \mapsto \partial_P v(x)$ is upper semicontinuous and by the representation (12) it is clear that its values are convex and compact subsets of \mathbb{K}^n .

Proposition 5.4 *Let $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{T} = \mathbb{R}_+$ and let $(\mathcal{S}, \mathbb{R}_+)$ be an irreducible semigroup generated by an irreducible, convex, compact set \mathcal{M} .*

(i) A norm v is an extremal norm for \mathcal{S} if and only if for all dual pairs $l, x \in \mathbb{K}^n$ and all $A \in \mathcal{M}$ it holds that

$$\operatorname{Re} \langle l, Ax \rangle \leq \log \rho(\mathcal{M}) v(x) v^*(l). \quad (28)$$

In this case, there exist dual pairs $l, x \in \mathbb{K}^n$ and $A \in \mathcal{M}$ where equality in (28) is attained.

(ii) A norm v is a Barabanov norm for \mathcal{S} if and only if for all $x \in \mathbb{K}^n$ there exists an $l \in \mathbb{K}^n$ such that l, x is a dual pair and an $A \in \mathcal{M}$ such that

$$\operatorname{Re} \langle l, Ax \rangle = \log \rho(\mathcal{M}) v(x) v^*(l). \quad (29)$$

Proof. Without loss of generality we may assume $\rho(\mathcal{M}) = 1$ because we may normalize to the set $\mathcal{M} - \log \rho(\mathcal{M}) I$ and we have for dual pairs that

$$\langle l, (A - \log \rho(\mathcal{M}) I)x \rangle = \langle l, Ax \rangle - \log \rho(\mathcal{M}) v(x) v^*(l).$$

(i) Extremality of a norm v is equivalent to the statement that for all $u \in L^\infty(\mathbb{R}_+, \mathcal{M})$ we have that the map $t \mapsto v(\Phi_u(t, 0)x)$ is nonincreasing. By [24, Theorem 4.6.3] this is equivalent to the statement that $\operatorname{Re} \langle l, Ax \rangle \leq 0$ for all $x \in \mathbb{K}^n, l \in \partial_P v(x), A \in \mathcal{M}$. The assertion now follows from (12).

If equality in (28) is not attained then using a compactness argument, there is a constant $c > 0$ such that

$$\operatorname{Re} \langle l, Ax \rangle \leq -cv(x) v^*(l). \quad (30)$$

or equivalently $\operatorname{Re} \langle l, (A + cI)x \rangle \leq 0$, for all dual pairs $l, x \in \mathbb{K}^n$ and all $A \in \mathcal{M}$. From this we obtain by the same arguments as before that $t \mapsto v(e^{ct}\Phi_u(t, 0)x)$ is nonincreasing, so that $\rho(\mathcal{M}) \leq e^{-ct}$, a contradiction.

(ii) An extremal norm v is a Barabanov norm, if and only if for all $x \in \partial \mathcal{B}_v := \{x \in \mathbb{K}^n \mid v(x) = 1\}$ there is a $u \in L^\infty(\mathbb{R}_+, \mathcal{M})$ such that $v(\Phi_u(t, 0)x) = 1$ for all $t \geq 0$. By [24, Theorem 4.2.10] this is equivalent to the statement that for all $x \in \partial \mathcal{B}_v$ we have $\{Ax \mid A \in \mathcal{M}\} \cap H \neq \emptyset$ for some supporting hyperplane H of \mathcal{B}_v in x . This implies that $\langle l, Ax \rangle = 0$ for some $l \in \partial_P v(x), A \in \mathcal{M}$. (Here the assumption of convexity of \mathcal{M} is vital.) By (12) l, x is then a dual pair. \square

The following result is a counterpart of Proposition 5.4 for the discrete-time case, which we quote from [5].

Proposition 5.5 *Let $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{T} = \mathbb{N}$ and let $(\mathcal{S}, \mathbb{N})$ be an irreducible semigroup generated by \mathcal{M} .*

(i) A norm v is an extremal norm for \mathcal{S} if and only if for all pairs $l, x \in \mathbb{K}^n$ and all $A \in \mathcal{M}$ it holds that

$$\operatorname{Re} \langle l, Ax \rangle \leq \rho(\mathcal{M}) v(x) v^*(l). \quad (31)$$

In this case, there exist pairs $l, x \in \mathbb{K}^n$ and $A \in \mathcal{M}$ where equality in (31) is attained.

(ii) A norm v is a Barabanov norm for \mathcal{S} if and only if for all $x \in \mathbb{K}^n$ there exists an $l \in \mathbb{K}^n$ and an $A \in \mathcal{M}$ such that

$$\operatorname{Re} \langle l, Ax \rangle = \rho(\mathcal{M}) v(x) v^*(l). \quad (32)$$

Let us now visualize the results related to the dual norm.

Example 5.6 *Consider the case $\mathbb{T} = \mathbb{N}$. Let $A_1 = \operatorname{diag}(0.5, 1)$, $A_2 = SA_1S^{-1}$, where $S = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$ is a rotation by angle φ . Let us take $\varphi = \frac{\pi}{4}$ for convenience. A_1 is an orthogonal contraction with respect to the y -axis, while $A_2 = \frac{1}{4} \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$ is an orthogonal contraction with respect to the axis $y = -x$.*

The set $\mathcal{M} = \{A_1, A_2\}$ is irreducible. Note also that $\mathcal{M} = \mathcal{M}^$, so that by Theorem 5.2 the dual of a Barabanov norm for \mathcal{M} is a Protasov norm for $\mathcal{M}^* = \mathcal{M}$. It is easy to see that the joint spectral radius satisfies $\rho(\mathcal{M}) = 1$. The unit ball \mathcal{B} of a Barabanov norm is given by a parallelogram spanned by the vertices $\left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \end{pmatrix} \right\}$. Its dual ball $\tilde{\mathcal{B}}$ is a parallelogram spanned by the vertices $\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\}$. As $A_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $A_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ we have $\operatorname{conv} \{A_1 \tilde{\mathcal{B}}, A_2 \tilde{\mathcal{B}}\} = \tilde{\mathcal{B}}$, hence it is the norm ball of a Protasov norm for \mathcal{M} , see Fig. 5.6 for an illustration.*

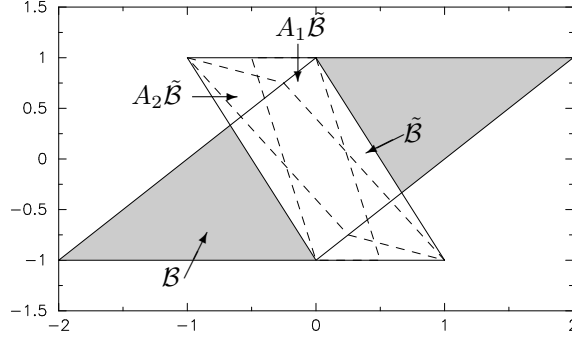


Figure 1: Extremal norms of Example 5.6

5.2 Duality and Transient Norms

Let us now investigate duality issues for transient norms and initial growth rates. We introduce one further notation. If $M_0(\mathcal{M})$ is the transient bound of $\mathcal{S}(\mathcal{M})$ with respect to the norm $\|\cdot\|$, then we denote by $M_0^*(\mathcal{M})$ the transient bound with respect to the dual norm, i.e.,

$$M_0^*(\mathcal{M}) = \sup\{\|S\|^* \mid S \in \mathcal{S}\}.$$

For dual norms we obtain the following result.

Theorem 5.7 *Suppose that $\|\cdot\|$ is a vector norm on \mathbb{K}^n with associated initial growth rate $\mu(\cdot)$ and let $\mu^*(\cdot)$ denote the initial growth rate with respect to the dual norm $\|\cdot\|^*$ on \mathbb{K}^n . Then for a set of matrices $\mathcal{M} \subset \mathbb{K}^{n \times n}$ the following statements hold*

1. $M_0(\mathcal{M}) = M_0^*(\mathcal{M}^*)$,
2. $\mu(\mathcal{M}) = \mu^*(\mathcal{M}^*)$.
3. $\log \mu_2(\mathcal{M}) \leq \frac{1}{2}(\log \mu(\mathcal{M}) + \log \mu^*(\mathcal{M}))$.

Proof. The first statement is obvious. We prove the remaining statements for the case $\mathbb{T} = \mathbb{R}_+$, the discrete time case follows in a similar manner. It follows from Proposition 2.5 that for all $A \in \mathcal{M}$

$$\begin{aligned} \log \mu(A) &= \max_{\|x\|=1} \max_{\|l\|^*=1, \langle l, x \rangle = 1} \operatorname{Re} \langle l, Ax \rangle, \\ \log \mu^*(A^*) &= \max_{\|l\|^*=1} \max_{\|x\|=1, \langle l, x \rangle = 1} \operatorname{Re} \langle x, A^*l \rangle. \end{aligned}$$

Now as $\operatorname{Re} \langle l, Ax \rangle = \operatorname{Re} \langle x, A^*l \rangle$ the equality $\mu^*(A) = \mu(A^*)$ is proved. The final statement follows from the second, because

$$\begin{aligned} \log \mu(A) + \log \mu^*(A) &= \log \mu(A) + \log \mu(A^*) \\ &\geq \log \mu(A + A^*) \geq \alpha(A + A^*) \\ &= \lambda_{\max}(A + A^*) = 2 \log \mu_2(A) \end{aligned}$$

where we used that $\log \mu(B)$ is a convex function, which is bounded from below by the spectral abscissa $\alpha(B) = \max \operatorname{Re} \sigma(B)$, see [11]. In case of a Hermitian matrix $B = A + A^*$ this abscissa is an eigenvalue. By Lemma 2.6 this eigenvalue equals $\log \mu_2(A)$. Clearly, $\log \mu(A) + \log \mu(A^*) \leq \log \mu(\mathcal{M}) + \log \mu(\mathcal{M}^*)$ for all $A \in \mathcal{M}$. So the inequality implies the assertion. \square

This theorem shows that the initial growth rate for the spectral norm is the best lower bound for all mean values of dual initial growth rates. Especially, for the the dual 1- and ∞ -norms we

immediately obtain $\mu_1^*(\mathcal{M}) = \mu_\infty(\mathcal{M})$ for all $\mathcal{M} \subset \mathbb{K}^{n \times n}$. Formulas for these initial growth rates are well known, see [11]. They are given by

$$\begin{aligned}\log \mu_1(A) &= \max_j a_{jj} + \sum_{i \neq j} |a_{ij}|, \\ \log \mu_\infty(A) &= \max_i a_{ii} + \sum_{j \neq i} |a_{ij}|,\end{aligned}$$

Using these formulas part 2 of Theorem 5.7 implies

Corollary 5.8 *Suppose that $\mathcal{M} \subset \mathbb{K}^{n \times n}$ is a set of column or row diagonally dominant matrices with $\log \mu_1(A) + \log \mu_\infty(A) < 0$ for all $A \in \mathcal{M}$. Then $\mathcal{S}(\mathcal{M})$ satisfies $\log \mu_2(\mathcal{M}) \leq 0$.*

Now that we have treated the initial growth of dual norms we proceed to the second main result of this paper. Feller norms and convex-transient norms are dual concepts. To this end we first need the following property of the eccentricity.

Proposition 5.9 *For norms $v, \|\cdot\|$ on \mathbb{K}^n it holds that*

$$\text{ecc}(v, \|\cdot\|) = \text{ecc}(v^*, \|\cdot\|^*).$$

Proof. By symmetry and using (14) it is sufficient to show that

$$\max_{\|y\|^*=1} v^*(y) = \max_{v(x)=1} \|x\| \quad \text{and} \quad \min_{\|y\|^*=1} v^*(y) \geq \min_{v(x)=1} \|x\|.$$

To show the first of these claims note that by definition,

$$\max_{\|y\|^*=1} v^*(y) = \max_{v(x)=1} \max_{\|y\|^*=1} |y^*x| = \max_{v(x)=1} \|x\|. \quad (33)$$

To show the second claim, assume that $\alpha \in \mathbb{R}_+$ is maximal with the property $v(\alpha z) \leq 1$ for all $\|z\| = 1$. Setting $u = \alpha z$ we have $\alpha = \min_{v(u)=1} \|u\|$. Then

$$\begin{aligned}\min_{\|y\|^*=1} v^*(y) &= \min_{\|y\|^*=1} \max_{v(x) \leq 1} |y^*x| \geq \min_{\|y\|^*=1} \alpha |y^*z| \\ &= \alpha = \min_{v(u)=1} \|u\|\end{aligned} \quad (34)$$

where we replaced x by αz such that z satisfies $v(\alpha z) \leq 1$, $\|z\| = 1$ and $y^*z = \|y\|^*$. Combining (33), (34) and (14), we obtain

$$\text{ecc}(v^*, \|\cdot\|^*) \leq \frac{\max_{v(y)=1} \|y\|}{\min_{v(y)=1} \|y\|} = \text{ecc}(\|\cdot\|, v) = \text{ecc}(v, \|\cdot\|).$$

By symmetry we obtain equality throughout. \square

By Proposition 5.9 the dual norm of a transient norm satisfies $\text{ecc}(v^*) = \text{ecc}(v)$ when $\|\cdot\|$ is the Euclidean norm. Let us therefore consider the norm $w(\cdot) := (v(\cdot)_{\mathcal{M}^*})^*$ which has the same eccentricity as the Feller norm $v(\cdot)_{\mathcal{M}^*}$. Then w coincides with the convex-transient norm.

Theorem 5.10 *Let $\mathbb{T} = \mathbb{R}_+, \mathbb{N}, \mathbb{K} = \mathbb{R}, \mathbb{C}$. Assume that $\mathcal{M} \subset \mathbb{K}^{n \times n}$ generates a bounded semi-group $(\mathcal{S}, \mathbb{T})$ and let $\|\cdot\|$ be some vector norm with unit ball \mathcal{B} . Then v is the Feller norm for $\mathcal{M}, \|\cdot\|$ if and only if the dual norm v^* is a convex-transient norm for $\mathcal{M}^*, \|\cdot\|^*$.*

Proof. Assume that v is a Feller norm for $\mathcal{M}, \|\cdot\|$ with unit ball \mathcal{B}_v . By Lemma 4.7 we have

$$\mathcal{B}_v = \bigcap_{S \in \mathcal{S}} S^{-1}\mathcal{B}. \quad (35)$$

Recall that a dual set of a convex set K is given by

$$K^* = \{y \in \mathbb{K}^n \mid |\langle y, x \rangle| \leq 1, \forall x \in K\}.$$

By Corollary 16.5.2 of [23] the dual of an intersection of convex sets C_i is given by the closed convex hull of the union of the convex sets C_i^* , and therefore from (35) we have

$$(\mathcal{B}_v)^* = \text{conv cl} \bigcup_{S \in \mathcal{S}} (S^{-1}\mathcal{B})^*. \quad (36)$$

It is easy to see that $(S^{-1}\mathcal{B})^* = S^*\mathcal{B}^*$ as for $x \in S^*\mathcal{B}^*$, $y \in S^{-1}\mathcal{B}$ we may choose $x_2 \in \mathcal{B}^*$, $x = S^*x_2$ and have $Sy =: y_2 \in \mathcal{B}$. Hence $|\langle y, x \rangle| = |\langle y, S^*x_2 \rangle| = |\langle Sy, x_2 \rangle| = |\langle y_2, x_2 \rangle| \leq 1$ by duality of \mathcal{B} and \mathcal{B}^* . Summarizing, we have

$$(\mathcal{B}_v)^* = \text{conv cl} \bigcup_{S \in \mathcal{S}} S^*\mathcal{B}^*. \quad (37)$$

As $(\mathcal{B}_v)^*$ is the unit ball of the dual norm v^* , this shows that v^* is a convex-transient norm for $\mathcal{M}^*, \|\cdot\|^*$.

If v^* is a convex-transient norm for $\mathcal{M}^*, \|\cdot\|^*$, then (37) holds. This implies (36) and in turn (35), so that the converse direction also holds. \square

Example 5.11 For the case $\mathbb{T} = \mathbb{R}_+$ we consider the differential equation $\dot{x} = Ax$ for the matrix $A = \begin{pmatrix} -5 & 36 \\ 0 & -20 \end{pmatrix}$. Its initial growth rate with respect to the Euclidean norm is given by $\mu_2(A) = \frac{1}{2}\lambda_{\max}(A + A^*) = 7$ hence the Euclidean norm is not a transient norm. On the left, Figure 2 shows the unit balls of the Feller norm, \mathcal{B} , and of the convex-transient norm, $\tilde{\mathcal{B}}$. Both unit balls are invariant under the flow of the system. On the right, the same situation is shown for the dual matrix $A^* = \begin{pmatrix} -5 & 0 \\ 36 & -20 \end{pmatrix}$. Thus the unit balls that are shaded dark on the right hand side and shaded lightly on the left correspond to dual norms (and vice versa).

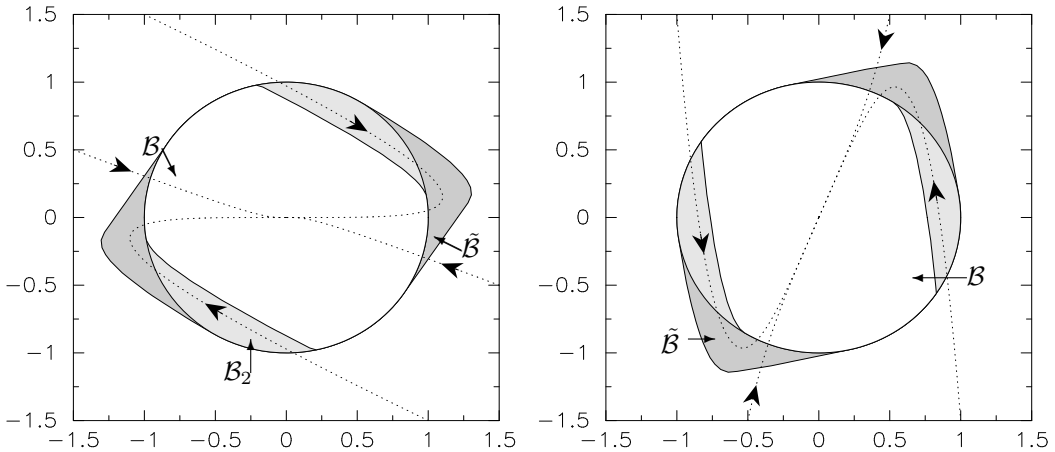


Figure 2: Transient norms, the dual case is shown on the right hand side.

Acknowledgement

This work was partially supported by the Enterprise Ireland grant 00/PI.1/C067. Elmar Plischke gratefully acknowledges support by the DFG grant HI 372/4. The authors thank Nikita Barabanov for useful discussions and comments.

References

- [1] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1990.
- [2] G.-C. Rota, G. Strang, A note on the joint spectral radius, *Indag. Math.* 22 (1960) 379–381.
- [3] N. E. Barabanov, Absolute characteristic exponent of a class of linear nonstationary systems of differential equations, *Siberian Mathematical Journal* 29 (4) (1988) 521–530.
- [4] F. Colonijs, W. Kliemann, Maximal and minimal Lyapunov exponents of bilinear control systems, *J. Diff. Eqns.* 101 (1993) 232–275.
- [5] N. E. Barabanov, Lyapunov indicator of discrete inclusions. I, *Autom. Remote Control* 49 (2) (1988) 152–157.
- [6] V. S. Kozyakin, Algebraic unsolvability of problem of absolute stability of desynchronized systems, *Autom. Rem. Control* 51 (6) (1990) 754–759.
- [7] F. Wirth, The generalized spectral radius and extremal norms, *Lin. Alg. Appl.* 342 (2002) 17–40.
- [8] F. Wirth, The generalized spectral radius is strictly increasing, *Linear Algebra Appl.* 395 (2005) 141–153.
- [9] G. Dahlquist, Stability and error bounds in the numerical integration of ordinary differential equations, *Transactions Royal Inst. of Technology* 130.
- [10] S. M. Lozinskiĭ, Error estimates for numerical integration of ordinary differential equations, *Izv. Vysš. Učebn. Zaved Matematika* 5 (1958) 52–90, errata 5:222, 1959.
- [11] M. Vidyasagar, *Nonlinear Systems Analysis*, 2nd Edition, Prentice Hall, Englewood Cliffs, NJ, 1993.
- [12] I. Higuera, G. Söderlind, Logarithmic norms and nonlinear DAE stability, *BIT* 42 (4) (2002) 823–841.
- [13] V. Kozyakin, A. Pokrovskij, The role of controllability type properties in the study of asynchronous dynamic systems., *Sov. Phys., Dokl.* 37 (5) (1992) 213–215.
- [14] V. Kozyakin, A. Pokrovskij, Quasi-controllability and estimation of the amplitudes of transient regimes in discrete systems., *Izv., Ross. Akad. Estestv. Nauk, Mat. Mat. Model. Inform. Upr.* 1 (3) (1997) 128–150.
- [15] V. Kozyakin, N. Kuznetsov, P. A. V., Quasi-controllability and estimates of amplitudes of transient regimes in discrete systems, in: A. K. P. Borne, M. Ksouri (Ed.), *Proc. of IMACS-IEEE Multiconference on Computational Engineering in Systems Applications CESA98*, Hammamet, Tunisia, 1998, pp. 266–271, (CD-ROM).
- [16] H. Sussmann, P. Kokotovic, The peaking phenomenon and the global stabilization of nonlinear systems., *IEEE Trans. Autom. Control* 36 (4) (1991) 424–440.
- [17] M. I. Gil', *Stability of Finite and Infinite Dimensional Systems*, Kluwer Academic Publishers, Boston/Dordrecht/London, 1998.
- [18] D. Hinrichsen, A. J. Pritchard, *Mathematical Systems Theory I. Modelling, State Space Analysis, Stability and Robustness*, Vol. 48 of *Texts in Applied Mathematics*, Springer-Verlag, Berlin, 2005.
- [19] V. Y. Protasov, The joint spectral radius and invariant sets of linear operators, *Fundam. Prikl. Mat.* 2 (1) (1996) 205–231, in Russian.

- [20] W. Feller, On the generation of unbounded semi-groups of bounded linear operators, *Ann. of Math.*, Ser. 2 58 (1953) 166–174.
- [21] M. A. Berger, Y. Wang, Bounded semigroups of matrices, *Lin. Alg. Appl.* 166 (1992) 21–27.
- [22] T. Ström, On logarithmic norms, *SIAM J. Numer. Anal.* 12 (5) (1975) 741–753.
- [23] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
- [24] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Vol. 178 of Graduate Texts in Mathematics, Springer, New York, 1998.