Zentrum für Technomathematik

Fachbereich 3 – Mathematik und Informatik

The generalized spectral radius and extremal norms

Fabian Wirth

Report 00–16

Berichte aus der Technomathematik

Report 00–16 Oktober 2000

The generalized spectral radius and extremal norms

Fabian Wirth
Zentrum für Technomathematik
Universität Bremen
D-28344 Bremen
Germany
email: fabian@math.uni-bremen.de

Abstract

The generalized spectral radius, also known under the name of joint spectral radius, or (after taking logarithms) maximal Lyapunov exponent of a discrete inclusion is examined. We present a new proof for a result of Barabanov, which states that for irreducible sets of matrices an extremal norm always exists. This approach lends itself easily to the analysis of further properties of the generalized spectral radius. We prove that the generalized spectral radius is locally Lipschitz continuous on the space of compact irreducible sets of matrices and show a strict monotonicity property of the generalized spectral radius. Sufficient conditions for the existence of extremal norms are obtained.

1 Introduction

In recent years discrete inclusions have attracted the interest of researchers from quite distinct fields. They occur in the theory of wavelets, where discrete inclusions can be used to determine Hoelder exponents of compactly supported wavelets, see Daubechies and Lagarias [7], Heil and Strang [15], and references therein. For discussions of applications in the theory of Markov

Keywords: Discrete inclusions, generalized spectral radius, joint spectral radius, extremal norms, irreducibility
1991 Subject Classification: 34D08, 93C55, 93D09

chains, iterated function systems, hysteresis nonlinearities we refer to references given in the papers [2, 14, 26]. For stability analysis of numerical algorithms using this framework we refer to Guglielmi and Zennaro [13]. And this list is, of course, far from complete.

Given a set of matrices $\mathcal{M} \subset \mathbb{K}^{n\times n}$, where $\mathbb{K} = \mathbb{R}, \mathbb{C}$, we are interested in the asymptotic behavior of solutions of the discrete inclusion

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

This problem has been studied from an abstract point of view in [23, 1, 16, 2, 17, 14, 9, 10, 20, 24, 26, 12]. Infinite dimensional versions of this problem have been studied in [14, 28]. A more general spectral theory for a wide class of discrete inclusions can be found in [27], see also [6] for continuous time analogues.

This author was first interested in stability of discrete inclusions from a control theory point of view. A discrete inclusion of the form (1) may be interpreted as a model for time-varying uncertainty for a nominal system $x(t+1) = Ax(t)$. One problem area in this direction consists in the calculation of *stability radii*. Given an increasing family of sets $\mathcal{U} := \{\mathcal{M}_\gamma \mid \gamma \geq 0\}$ the problem is to determine the smallest $\gamma > 0$ such that (1) defined by $\mathcal{M}_\gamma$ is not exponentially stable, see also [29].

A recurrent problem is the question whether $\mathcal{M}$ has *left convergent products* or is *product bounded*. The first of these properties means that for any sequence $\{A(k)\}_{k\in\mathbb{N}} \in \mathcal{M}^{\mathbb{N}}$ it holds that

$$A(k)A(k-1)\cdots A(0)$$

is convergent for $k \to \infty$. Product boundedness means that there is a constant $C > 0$ such that $\|A(k)A(k-1)\cdots A(0)\| < C$ for all possible products of matrices in $\mathcal{M}$. This property is also called *absolute stability* in [16] and *nondefectiveness* in [13].

The property of left convergent products has been studied in [7, 9, 10]. In particular, this property is characterized in a number of ways for finite sets of matrices by Vladimirov et al. [26], where also results on general sets of matrices are obtained, which are not quite as far-reaching.

One of the main tools in the study of discrete inclusions consists of the generalized (or joint) spectral radius. This approach originates with Rota and Strang [23], who defined the joint spectral radius and Daubechies and Lagarias [7], who did the same for the generalized spectral radius. We now define these two numbers. Associated to the set $\mathcal{M}$ we can consider the sets of products of length t

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

and the semigroup given by

$$\mathcal{S} := \bigcup_{t=1}^{\infty} \mathcal{S}_t\,.$$

Let $\|\cdot\|$ be some operator norm on $\mathbb{K}^{n\times n}$ and define for $t \in \mathbb{N}$

$$\overline{\rho}_t(\mathcal{M}) := \sup\{r(S_t)^{1/t} \mid S_t \in \mathcal{S}_t\}\,, \qquad \hat{\rho}_t(\mathcal{M}) := \sup\{\|S_t\|^{1/t} \mid S_t \in \mathcal{S}_t\}\,. \tag{2}$$

The *joint spectral radius*, respectively the *generalized spectral radius* are now defined as

$$\overline{\rho}(\mathcal{M}) := \limsup_{t\to\infty} \overline{\rho}_t(\mathcal{M})\,, \quad \hat{\rho}(\mathcal{M}) := \lim_{t\to\infty} \hat{\rho}_t(\mathcal{M})\,.$$

However, there is no need to insist on different notation as Theorem 4 in Berger and Wang [2] states that for bounded $\mathcal{M}$ we have $\hat{\rho}(\mathcal{M}) = \overline{\rho}(\mathcal{M})$, so that we will simply use the notation $\rho(\mathcal{M})$. Different proofs for this equality can be found in [8, 24]. Note also that for all $t \geq 1$

$$\overline{\rho}_t(\mathcal{M}) \leq \rho(\mathcal{M}) \leq \hat{\rho}_t(\mathcal{M})\,. \tag{3}$$

In a paper by Lagarias and Wang [17] the by now famous "finiteness conjecture" was formulated, which states that for a finite set of matrices $\mathcal{M}$ there always exists a $t \geq 1$ such that

$$\rho(\mathcal{M}) = \overline{\rho}_t(\mathcal{M})\,.$$

It has recently be shown by Bousch and Mairesse [4], that this conjecture is false. But in special cases it can be shown to hold, see [14, 17].

The calculation of the generalized spectral radius has been treated using different approaches. While Gripenberg [11] and Maesumi [19] reduce the number of matrix products that have to be evaluated to obtain upper, respectively lower bounds given by $\hat{\rho}_t, \overline{\rho}_t$, an optimal control approach is used in [29]. Nice computational results cannot be really expected as Kozyakin [16] has shown that ρ is not an algebraic function on the vector space of k-tuples of $n \times n$ matrices and the determination of ρ is NP-hard by a result of Tsitsiklis and Blondel [25].

In this paper we show two further properties of the generalized spectral radius, namely local Lipschitz continuity on the set of irreducible compact sets of matrices and a monotonicity property. Our approach is based on a further important idea in the analysis of exponential stability of discrete inclusions that was introduced by Barabanov [1]. Recall that $\mathcal{M} \subset \mathbb{K}^{n\times n}$ is called *irreducible* if only the trivial subspaces $\{0\}$ and $\mathbb{K}^n$ are invariant under all matrices $A \in \mathcal{M}$. Otherwise $\mathcal{M}$ is called *reducible.*

An immediate consequence of irreducibility of $\mathcal{M}$ is that $\rho(\mathcal{M}) > -\infty$, because in this case the semigroup $\mathcal{S}$ is irreducible and does therefore not consist of nilpotent elements by the Levitzky theorem [18]. Note that this implies in particular, that we can always normalize an irreducible set of matrices $\mathcal{M}$ to $\rho(\mathcal{M})^{-1}\mathcal{M}$ which is a set with generalized spectral radius equal to 1.

The fundamental contribution of Barabanov consists of the following result.

Theorem 1.1 *If $\mathcal{M}$ is compact and irreducible, then there exists a norm v on $\mathbb{K}^n$ such that*

(i) for all $x \in \mathbb{K}^n, A \in \mathcal{M}$ it holds that

$$v(Ax) \leq \rho(\mathcal{M})v(x)\,,$$

(ii) for all $x \in \mathbb{K}^n$ there exists an $A \in \mathcal{M}$ such that

$$v(Ax) = \rho(\mathcal{M})v(x)\,.$$

We will in particular be interested in the existence of *extremal norms*, that is norms with the property that $\|A\| = \rho(\mathcal{M})$ for all $A \in \mathcal{M}$. It follows from the result in Kozyakin that an extremal norm exists for $\mathcal{M}$ if and only if $\rho(\mathcal{M})^{-1}\mathcal{M}$ is product bounded, [16, Theorem 3]. A further characterization is obtained in [12, Section 3]. As the question whether a pair of matrices is product bounded is undecidable by a recent result of Blondel and Tsitsiklis [3] we do not expect to obtain an easily checkable criterion and so our condition is just sufficient but not necessary.

The paper is organized as follows. In Section 2 we present the class of systems that is studied; as our methods work just as well for semigroups generated by continuous time systems we briefly introduce the necessary concepts. In Section 3 we introduce our main technical tool, which we call the *limit semigroup* and which is obtained as the ω-limit set of the semigroup normalized to a generalized spectral radius equal to 1.

In Section 4 we use the result of the previous section to show that ρ is locally Lipschitz continuous on the set of compact irreducible sets of matrices. In Section 5 we show that the generalized spectral radius is a strictly increasing function under a natural growth condition on a function with values in the compact sets of matrices. This result is motivated by the problem of calculating time-varying stability radii and its consequences will be discussed in a forthcoming paper.

Finally, in Section 6 we show the existence of extremal norms under a nondefectiveness condition, which generalizes the corresponding result for the spectral radius of a matrix. Note, that we found it useful to use a slightly different sense of the word nondefective than found in the literature. In [12] "nondefective" just means that an extremal norm exists.

2 Preliminaries

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

$$\begin{aligned} x(t+1) &\in \{Ax(t) \mid A \in \mathcal{M}\}, \quad t \in \mathbb{N} \\ x(0) &= x_0 \quad \in \mathbb{K}^n . \end{aligned} \tag{4}$$

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 for all $t \in \mathbb{N}$ there exists an $A(t) \in \mathcal{M}$ such that $x(t+1) = A(t)x(t)$. We continue to use the notation introduced in Section 1.

As all our arguments are also valid in continuous time, we will just consider an irreducible semigroup $\mathcal{S} \subset \mathbb{K}^{n \times n}$ with an associated time scale $\mathbb{T} = \mathbb{N}, \mathbb{R}_+ := [0, \infty)$. To be concrete, in the case $\mathbb{T} = \mathbb{R}_+$ we assume that the semigroup is generated by a differential inclusion

$$\dot{x} \in \{Ax(t) \mid A \in \mathcal{M}\}, \tag{5}$$

where $\mathcal{M} \subset \mathbb{K}^{n \times n}$ is compact. In the latter case the elements of $\mathcal{S}_t, t \in \mathbb{R}_+$ are the evolution operators $\Phi_{A(\cdot)}(t, 0)$ corresponding to measurable functions $A : \mathbb{R}_+ \to \mathcal{M}$ and the time-varying differential equation

$$\dot{x}(t) = A(t)x(t), \text{ a.e.}.$$

For a semigroup defined by (5) the quantities $\overline{\rho}_t(\mathcal{S}), \hat{\rho}_t(\mathcal{S}), t \in \mathbb{R}_+$ can be defined analogously to (2) and make obviously sense.

We will denote the corresponding limit by $\rho(\mathcal{S})$. We call this quantity the maximal Lyapunov exponent if we consider differential inclusions (although in the literature this name is normally reserved for $\log \rho(\mathcal{S})$). There is abundant literature on the theory of Lyapunov exponents of differential inclusions, see e.g. [5, 6] and references therein.

If we fear that there is a chance of confusion we will denote the generalized spectral radius given by a set $\mathcal{M}$ via the discrete inclusion (1) by $\rho(\mathcal{M}, \mathbb{N})$ and the maximal Lyapunov exponent by $\rho(\mathcal{M}, \mathbb{R}_+)$.

Note that given a semigroup $(\mathcal{S}, \mathbb{R}_+)$ we can always associate a discrete inclusion by defining $\mathcal{M} := \mathcal{S}_1$. Under our assumptions it is an easy exercise to check that $\rho(\mathcal{S}, \mathbb{R}_+) = \rho(\mathcal{M}, \mathbb{N})$. In the sequel, we will always tacitly assume that for each $t \in \mathbb{T}$ the set $\mathcal{S}_t$ is bounded, if we just speak of a semigroup $(\mathcal{S}, \mathbb{T})$.

Definition 2.1 *Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$, $\mathbb{T} = \mathbb{N}, \mathbb{R}_+$ and let $(\mathcal{S}, \mathbb{T})$ be a semigroup in $\mathbb{K}^{n \times n}$. A norm v on $\mathbb{K}^n$ is called* Barabanov norm *corresponding to $\mathcal{S}$ if*

(i) $v(Sx) \le \rho(\mathcal{S})^t v(x)\,, \quad$ for all $x \in \mathbb{K}^n, t \in \mathbb{T}, S \in \mathcal{S}_t$,

(ii) for all $x \in \mathbb{K}^n$, $t \in \mathbb{T}$ there is an $S \in \operatorname{cl} \mathcal{S}_t$ such that

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

A norm v on $\mathbb{K}^n$ is called extremal *for $\mathcal{S}$ if for the corresponding operator norm it holds that*

$$v(S) \le \rho(\mathcal{S})^t\,, \text{ for all } t \in \mathbb{T}, S \in \mathcal{S}_t\,.$$

We will investigate further conditions guaranteeing the existence of extremal norms in Section 6.

We will also consider the behavior of the generalized spectral radius as a function of the set $\mathcal{M}$. As we only have to consider compact sets $\mathcal{M} \subset \mathbb{K}^{n \times n}$, we introduce

$$\mathcal{K}(\mathbb{K}^{n \times n}) := \{\mathcal{M} \subset \mathbb{K}^{n \times n} \mid \mathcal{M} \text{ compact, nonempty}\}\,.$$

The space $\mathcal{K}(\mathbb{K}^{n \times n})$ becomes a complete metric space if it is endowed with the usual Hausdorff metric defined by

$$H(\mathcal{M}, \mathcal{N}) := \max\{\max_{A \in \mathcal{M}} \operatorname{dist}(A, \mathcal{N}), \max_{B \in \mathcal{N}} \operatorname{dist}(B, \mathcal{M})\}\,.$$

Note that with respect to this topology the set

$$I(\mathbb{K}^{n \times n}) := \{\mathcal{M} \in \mathcal{K}(\mathbb{K}^{n \times n}) \mid \mathcal{M} \text{ irreducible}\}$$

is open and dense in $\mathcal{K}(\mathbb{K}^{n \times n})$.

3 The limit semigroup

In this section we present an alternative and we hope less intricate proof of Barabanov's result. We need the following property of irreducible semigroups.

Lemma 3.1 *Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$, $\mathbb{T} = \mathbb{N}, \mathbb{R}_+$ and let $(\mathcal{S}, \mathbb{T})$ be an irreducible semigroup in $\mathbb{K}^{n\times n}$. Then there are $\varepsilon > 0$ and $\tau \in \mathbb{T}$ such that for all $z \in \mathbb{K}^n, A \in \mathbb{K}^{n\times n}$ there is an $S \in \bigcup_{1\le t\le\tau} \mathcal{S}_t$ with*

$$\|ASz\| \ge \varepsilon \|A\| \|z\| \, .$$

Proof: Assume the assertion is false, so that there are $\varepsilon_k \to 0$, $\tau_k \to \infty, \tau_k \in \mathbb{T}, z_k \in \mathbb{K}^n, A_k \in \mathbb{K}^{n\times n}$ such that for all $S \in \bigcup_{1\le t\le\tau_k} \mathcal{S}_t$ we have

$$\|A_k S z_k\| < \varepsilon_k \|A_k\| \|z_k\| \, . \tag{6}$$

Without loss of generality we may assume that $\|z_k\| = \|A_k\| = 1$. Thus we may assume $z_k \to z, A_k \to A$ with $\|z\| = \|A\| = 1$. Then irreducibility of $\mathcal{S}$ implies that there exists an $S^* \in \mathcal{S}$ with

$$\|AS^* z\| = \varepsilon^* > 0 \, ,$$

otherwise $\{Sz \mid S \in \mathcal{S}\}$ is contained in the kernel of A. This, however, contradicts irreducibility of $\mathcal{S}$ as $\mathbb{K}^n \neq \ker A$ due to $\|A\| = 1$. For all k large enough we have $S^* \in \bigcup_{1\le t\le\tau_k} \mathcal{S}_t$ and

$$\|A_k S^* z_k\| \ge \varepsilon^*/2 \, ,$$

which contradicts (6). This concludes the proof. □

Given our irreducible semigroup $(\mathcal{S}, \mathbb{T})$ we define the *limit semigroup* $\mathcal{S}_\infty$ by

$$\mathcal{S}_\infty := \{S \in \mathbb{K}^{n\times n} \mid \exists t_k \to \infty, S_{t_k} \in \mathcal{S}_{t_k} \text{ such that } \rho(\mathcal{S})^{-t_k} S_{t_k} \to S\} \, . \tag{7}$$

We note the following properties of $\mathcal{S}_\infty$.

Proposition 3.2 *Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$, $\mathbb{T} = \mathbb{N}, \mathbb{R}_+$ and let $(\mathcal{S}, \mathbb{T})$ be an irreducible semigroup in $\mathbb{K}^{n\times n}$. The set $\mathcal{S}_\infty$ defined by (7) satisfies*

(i) $\mathcal{S}_\infty$ is compact and nonempty,

(ii) $\mathcal{S}_\infty$ is a semigroup,

(iii) for $T \in \mathcal{S}_t, S \in \mathcal{S}_\infty$ we have

$$\rho(\mathcal{S})^{-t} TS\,, \quad \rho(\mathcal{S})^{-t} ST \in \mathcal{S}_\infty\,,$$

(iv) for all $t \in \mathbb{T}$, $S \in \mathcal{S}_\infty$ there exist $T \in \mathcal{S}_\infty, A \in \operatorname{cl}\mathcal{S}_t$ as well as $R \in \mathcal{S}_\infty, B \in \operatorname{cl}\mathcal{S}_t$ such that

$$S = \rho(\mathcal{S})^{-t} TA = \rho(\mathcal{S})^{-t} BR\,,$$

(v) $\mathcal{S}_\infty$ is irreducible.

Proof: Without loss of generality we may assume $\rho(\mathcal{S}) = 1$ in this proof.

(i) For $A \in \mathcal{S}_t$ it holds that $r(A) \leq \rho(\mathcal{S})^t = 1$, hence $\{A^t\}$ is a bounded sequence which has an accumulation point S. By definition $S \in \mathcal{S}_\infty$. To see that $\mathcal{S}_\infty$ is closed it suffices to use a standard argument from the construction of ω-limit sets.

It remains to show that $\mathcal{S}_\infty$ is bounded. Assume this is not the case and let $\varepsilon > 0$ and $\tau \in \mathbb{T}$ be the constants given by Lemma 3.1. Unboundedness of $\mathcal{S}_\infty$ implies that there exists some $t \in \mathbb{T}$, $S \in \mathcal{S}_t$ with $\|S\| > 2/\varepsilon$. Thus for $x_0, \|x_0\| = 1$ arbitrary, there is a $T \in \bigcup_{1\leq t\leq\tau} \mathcal{S}_t$ with

$$\|STx_0\| > 2$$

and applying this argument repeatedly we obtain a sequence $\{T_k\}_{k\in\mathbb{N}} \subset \bigcup_{1\leq t\leq\tau} \mathcal{S}_t$ such that

$$\|ST_k \ldots ST_1 x_0\| > 2^k\,, \quad k \in \mathbb{N}\,.$$

This implies $\hat\rho_{kt+\tau_k}(\mathcal{S}) \geq 2^{1/(t+\tau)}$, where $k \leq \tau_k \leq k\tau$, a contradiction.

(ii) Let $S, T \in \mathcal{S}_\infty$ and consider sequences $s_k, t_k \to \infty$, $S_k \in \mathcal{S}_{s_k}, S_k \to S$ and $T_k \in \mathcal{S}_{t_k}, T_k \to T$. Then

$$\|ST - S_kT_k\| \leq \|S - S_k\|\|T\| + \|S_k\|\|T - T_k\|\,,$$

which goes to zero as both terms go to zero for $k \to \infty$. Hence $ST \in \mathcal{S}_\infty$.

(iii) This is clear, as approximation of S by a sequence S_k implies approximation of TS and ST by TS_k, respectively S_kT.

(iv) Let $t_k \to \infty, S_k \in \mathcal{S}_{t_k}$ be sequences such that $S_k \to S$. We can write $S_k = T_k A_k$ with $T_k \in \mathcal{S}_{t_k - t}, A_k \in \mathcal{S}_t$. Without loss of generality $A_k \to A \in \mathrm{cl}\, \mathcal{S}_t$ and $T_k \to T \in \mathcal{S}_\infty$. This implies $S = TA$, as required. The argument for the left factorization is exactly the same.

(v) By (ii) and (iii) we know that

$$\overline{\mathcal{S}} := \mathcal{S}_\infty \cup \bigcup_{t \in \mathbb{T}} \rho(\mathcal{S})^{-t} \mathcal{S}_t$$

is an irreducible semigroup of which $\mathcal{S}_\infty$ is a closed irreducible semigroup ideal. Now $\mathcal{S}_\infty$ is irreducible by [22, Lemma 1].

□

We give an easy example for the above construction, that will turn out to be of use in the remainder of the article.

Example 3.3 Consider the set

$$\mathcal{M} := \left\{ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}.$$

It is easy to see that

$$\mathcal{S}_{2k} = \left\{ 0, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\},$$

whereas $\mathcal{S}_{2k+1} = \mathcal{M} \cup \{0\}$. Hence $\mathcal{S}_\infty = \mathcal{M} \cup \mathcal{S}_2$.

Given our irreducible semigroup $(\mathcal{S}, \mathbb{T})$ and the associated limit semigroup $\mathcal{S}_\infty$ we now define the function

$$v(x) := \max_{S \in \mathcal{S}_\infty} \|Sx\| \tag{8}$$

and note that this defines the norm we are looking for.

Lemma 3.4 *Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$, $\mathbb{T} = \mathbb{N}, \mathbb{R}_+$ and let $(\mathcal{S}, \mathbb{T})$ be an irreducible semigroup in $\mathbb{K}^{n \times n}$. Then v is a Barabanov norm for $\mathcal{S}$.*

Proof:

(i) We first show that v is a norm. The properties $v(0) = 0, v(\lambda x) = |\lambda| v(x)$ are clear. If $x \neq 0$ then $v(x) \neq 0$ as otherwise $\mathrm{span}\,\{x\}$ would be in the kernel of all $S \in \mathcal{S}_\infty$ contradicting irreducibility. The function $v(x)$ is finite as $\mathcal{S}_\infty$ is compact and finally

$$v(x+y) \leq \max_{S \in \mathcal{S}_\infty} \|Sx\| + \|Sy\| \leq \max_{S \in \mathcal{S}_\infty} \|Sx\| + \max_{S \in \mathcal{S}_\infty} \|Sy\|.$$

(ii) Without loss of generality let $\rho(\mathcal{S}) = 1$. Let $x \in \mathbb{K}^n, S \in \mathcal{S}$ be arbitrary, then

$$v(Sx) = \max_{T \in \mathcal{S}_\infty} \|TSx\| \leq \max_{T \in \mathcal{S}_\infty} \|Tx\| = v(x)\,, \tag{9}$$

as $TS \in \mathcal{S}_\infty$ for all $T \in \mathcal{S}_\infty$. To prove the second statement assume that $S_x \in \mathcal{S}_\infty$ is such that $v(x) = \|S_x x\|$, then by Proposition 3.2 (iv) S_x factors into $S_x = TA$ with $T \in \mathcal{S}_\infty, A \in \operatorname{cl} \mathcal{S}_t$. Hence

$$v(Ax) = \max_{S \in \mathcal{S}_\infty} \|SAx\| \geq \|TAx\| = v(x)\,,$$

and so by (9) we have $v(Ax) = v(x)$.

□

The existence of a Barabanov norm has many consequences as already noted in [1]. For instance, it is immediate that $\rho(\mathcal{M}) = \rho(\operatorname{cl} \mathcal{M})$ and $\rho(\mathcal{M}) = \rho(\operatorname{conv} \mathcal{M})$. In particular, we cite the following continuity result from [1] which will be of use for us in the sequel. Alternatively, it has been noted by Heil and Strang [15] that the continuity of the generalized spectral radius is a direct consequence of the equality $\rho(\mathcal{M}) = \overline{\rho}(\mathcal{M}) = \hat{\rho}(\mathcal{M})$. (The argument is given for the case of pairs of matrices, but is easily seen to extend to general bounded sets of matrices.)

Lemma 3.5 *The map $\mathcal{M} \to \rho(\mathcal{M})$ is continuous from $\mathcal{K}(\mathbb{K}^{n\times n})$ to $\mathbb{R}_+$.*

4 Lipschitz continuity of the generalized spectral radius

In this section we intend to show that the generalized spectral radius is locally Lipschitz continuous on the set of irreducible compact sets of matrices.

To this end we begin by an investigation of the variation of Barabanov norms under changes of $\mathcal{M}$. For irreducible $\mathcal{M}$ we will need to know how much the original norm is deformed under the definition (8). Therefore we introduce the quantities

$$c^-(\mathcal{M}) \ := \ \min\{v(x) \mid \|x\| = 1\}\,, \tag{10}$$

$$c^+(\mathcal{M}) \ := \ \max\{v(x) \mid \|x\| = 1\}\,. \tag{11}$$

Of course, these constant also depend on the choice $\mathbb{T} = \mathbb{N}$ or $\mathbb{T} = \mathbb{R}_+$, but we suppress this dependence. Note that for any $A \in \mathbb{K}^{n\times n}$ we have for the induced operator norm that

$$\frac{c^-(\mathcal{M})}{c^+(\mathcal{M})}\|A\| \leq v(A) \leq \frac{c^+(\mathcal{M})}{c^-(\mathcal{M})}\|A\|\,.$$

Theorem 4.1 *Let $P \subset I(\mathbb{K}^{n\times n})$ be compact and let $\mathbb{T} = \mathbb{N}$ or $\mathbb{T} = \mathbb{R}_+$. Then there is a constant $C > 0$ such that*

$$1 \leq \frac{c^+(\mathcal{M})}{c^-(\mathcal{M})} \leq C\,, \text{ for all } \mathcal{M} \in P\,.$$

Proof: Fix a timeset $\mathbb{T} \in \{\mathbb{N}, \mathbb{R}_+\}$ and consider the corresponding semigroups generated by the sets $\mathcal{M} \in P$. Assume to the contrary that there exists a sequence $\{\mathcal{M}_k\} \subset P$ such that

$$\frac{c^+(\mathcal{M}_k)}{c^-(\mathcal{M}_k)} \to \infty\,.$$

Without loss of generality we may assume that $\mathcal{M}_k \to \mathcal{M} \in P$.

For every k choose a $S_k \in \mathcal{S}_{\infty,k}$ (the limit semigroup corresponding to $(\mathcal{M}_k, \mathbb{T})$) such that $\|S_k\| = c^+(\mathcal{M}_k)$ and denote

$$\tilde{S}_k := \frac{S_k}{\|S_k\|}\,.$$

Then we may assume that $\tilde{S}_k \to \tilde{S}$ with $\|\tilde{S}\| = 1$.

Now let $x_0 \in \mathbb{K}^n, \|x_0\| = 1$ be arbitrary. We will show that $c^+(\mathcal{M}_k)/v_k(x_0)$ is bounded by a constant independent of x_0, which proves the assertion.

Let $\varepsilon > 0, \tau \in \mathbb{T}$ be the constants for $\mathcal{S}$ (the semigroup generated by $(\mathcal{M}, \mathbb{T})$ guaranteed by Lemma 3.1. Then by convergence of the sets $\mathcal{M}_k$ there exists a $k_0 \in \mathbb{N}$ such that for every $k \geq k_0$ and some $R_k \in \mathcal{S}_{t_k,k}, 1 \leq t_k \leq \tau$ we have

$$\|\tilde{S}R_k x_0\| \geq \frac{\varepsilon}{2}\,.$$

Note, that k_0 is chosen independently of x_0. For all $k \geq k_0$ we now define

$$T_k := \rho(\mathcal{S}_k)^{-t_k} S_k R_k \in \mathcal{S}_{\infty,k}\,,$$

and obtain for the norm v_k defined through $\mathcal{S}_{\infty,k}$ that

$$v_k(x_0) \geq \|T_k x_0\| = \rho(\mathcal{S}_k)^{-t_k}\|S_k R_k x_0\| = \frac{c^+(\mathcal{M}_k)}{\rho(\mathcal{S}_k)^{t_k}}\|\tilde{S}_k R_k x_0\|$$

$$\geq \frac{c^+(\mathcal{M}_k)}{\rho(\mathcal{S}_k)^{t_k}}\left(\|\tilde{S}R_k x_0\| - \|\tilde{S} - \tilde{S}_k\|\|R_k x_0\|\right) \geq \frac{c^+(\mathcal{M}_k)}{\rho(\mathcal{S}_k)^{t_k}}\left(\frac{\varepsilon}{2} - \|\tilde{S} - \tilde{S}_k\|\|R_k\|\right)\,.$$

The last term converges to zero be the definition of $\tilde{S}$ and as the set of all products of length at most τ is uniformly bounded over P. Furthermore, by continuity of ρ we have that $\rho(\mathcal{S}_k) \leq \rho(\mathcal{S}) + \varepsilon$ for $k \geq k_1 \geq k_0$, k_1 sufficiently large. This implies that for all k large enough we have

$$\frac{c^+(\mathcal{M}_k)}{v_k(x_0)} \leq \frac{4}{\varepsilon} \max\{1, \rho(\mathcal{S}) + \varepsilon\}^\tau .$$

This shows the assertion because again we have chosen k_1 independently of x_0. $\square$

As an application of Theorem 4.1 we can sharpen Lemma 3.5. We first just treat the discrete time case.

Corollary 4.2 *The generalized spectral radius is locally Lipschitz continuous on $I(\mathbb{K}^{n\times n})$.*

Proof: Let $P \subset I(\mathbb{K}^{n\times n})$ be compact with respect to the Hausdorff metric. Fix $\mathcal{M}, \mathcal{N} \in P$ arbitrary and let v denote the Barabanov norm with respect to $\mathcal{M}$. In the Hausdorff metric induced by our original norm $\|\cdot\|$ we have

$$H(\mathcal{M}, \mathcal{N}) =: a ,$$

which implies the in the Hausdorff metric H_v induced by v it holds that

$$H_v(\mathcal{M}, \mathcal{N}) \leq \frac{c^+(\mathcal{M})}{c^-(\mathcal{M})} a \leq Ca ,$$

where C is a constant only depending on P which exists by Theorem 4.1. Hence for all $x \in \mathbb{K}^n, A \in \mathcal{N}$ it holds that there exists a $B \in \mathcal{M}$ with $v(A - B) \leq Ca$ and thus

$$v(Ax) \leq v(Bx) + v((A - B)x) \leq (\rho(\mathcal{M}) + Ca)\, v(x) .$$

Hence $\rho(\mathcal{N}) \leq \rho(\mathcal{M}) + Ca$ and by symmetry we obtain

$$|\rho(\mathcal{N}) - \rho(\mathcal{M})| \leq CH(\mathcal{M}, \mathcal{N}) ,$$

as desired. $\square$

We cannot expect that the generalized spectral radius $\rho(\cdot)$ is Lipschitz continuous on $\mathcal{K}(\mathbb{K}^{n\times n})$ as already standard perturbation theory of eigenvalues tells us that generally if an eigenvalue splitting occurs at an eigenvalue with modulus equal to the spectral radius then the spectral radius will behave like a Puiseux series, that is, not Lipschitzean at the splitting point. An example for this phenomenon is given by

$$A_\varepsilon := \begin{bmatrix} 1 & 1 \\ \varepsilon & 1 \end{bmatrix},$$

the spectral radius of which for $\varepsilon > 0$ is given by $r(A_\varepsilon) = 1 + \sqrt{\varepsilon}$.

We note that the result translates immediately to continuous time.

Corollary 4.3 *The maximal Lyapunov exponent is locally Lipschitz continuous on $I(\mathbb{K}^{n\times n})$.*

Proof: The map

$$\mathcal{M} \mapsto \mathcal{S}_1(\mathcal{M}, \mathbb{R}_+)$$

is locally Lipschitz continuous on $\mathbb{K}^{n\times n}$. We have already noted that

$$\rho(\mathcal{M}, \mathbb{R}_+) = \rho(\mathcal{S}_1(\mathcal{M}), \mathbb{N}).$$

Now the assertion is immediate from Corollary 4.2. □

5 Strict monotonicity of the generalized spectral radius

In this section we will consider a further aspect of the generalized spectral radius under variation of the generating set $\mathcal{M}$. The methods we use here are restricted to the discrete time case, so that all results in this section are to understood with respect to the discrete inclusion (1). Whenever we treat different set of matrices $\mathcal{M}_1, \mathcal{M}_2$ in this section, we denote the semigroups and limit semigroups generated by $\mathcal{M}_1$ and $\mathcal{M}_2$ by $\mathcal{S}_1, \mathcal{S}_{\infty,1}$, respectively $\mathcal{S}_2, \mathcal{S}_{\infty,2}$. Correspondingly, the norms are denoted by v_1, v_2.

We need the following lemma.

Lemma 5.1 *Let $\mathcal{M}_1, \mathcal{M}_2 \in I(\mathbb{K}^{n\times n})$ satisfy*

$$\mathcal{M}_1 \subsetneq \mathcal{M}_2, \tag{12}$$

then for every $x \in \mathbb{K}^n$ there exist $S \in \mathcal{S}_{\infty,1}$ such that

$$v_1(x) = \|Sx\|$$

and such that there exists a factorization $S = \rho(\mathcal{M}_1)^{-k} A_1 \cdots A_k S_k$, $A_j \in \mathcal{M}_1, j = 1, \dots, k, S_k \in \mathcal{S}_{\infty,1}$ *such that*

$$\{A_1 \cdots A_{k-1} B S_k x \mid B \in \mathcal{M}_2\}$$

contains more than one element.

Proof: Assume $\rho(\mathcal{M}_1) = 1$ and fix $x \in \mathbb{K}^n$. Choose $S \in \mathcal{S}_{\infty,1}$ with $v_1(x) = \|Sx\|$. By Proposition 3.2 (iv) we can choose a factorization $S = A_1 S_1$ with $A_1 \in \mathcal{M}, S_1 \in \mathcal{S}_{\infty,1}$. If there is a $B \in \mathcal{M}_2$ with $BS_1 x \neq A_1 S_1 x$ we are done. Otherwise $Sx = BS_1 x$ for all $B \in \mathcal{M}_2$ and then by irreducibility of $\mathcal{M}_2$

$$S_1 x \notin \operatorname{span}\{A_1 S_1 x\}.$$

Now factorize $S_1 = A_2 S_2$, $A_2 \in \mathcal{M}_1, S_2 \in \mathcal{S}_{\infty,1}$. If there is a $B_2 \in \mathcal{M}_2$ such that $A_2 S_2 x \neq BS_2 x$, then for some $A \in \mathcal{M}_1$ we have $AA_2 S_2 x \neq ABS_2 x$, as otherwise $(A_2 - B)S_2 x$ is contained in the kernel of all $A \in \mathcal{M}_1$ in contradiction with irreducibility. On the other hand, due to our construction $\|AS_1 x\| = \|Sx\|$ and we are done. Otherwise, again due to irreducibility of $\mathcal{M}_2$, we obtain that

$$S_2 x \notin \operatorname{span}\{A_1 S_1 x, A_2 S_2 x\}. \tag{13}$$

We continue this procedure inductively. The argument breaks down in the n-th step as it is impossible that the corresponding version of (13) given by

$$S_n x \notin \operatorname{span}\{A_1 S_1 x, A_2 S_2 x, \dots, A_n S_n x\}.$$

can be satisfied, as the set on the right is linear independent if we could proceed up to this step. This shows that for some $1 \leq k \leq n$ the assertion must be satisfied. □

The main result of this section is the following proposition which states that the generalized spectral radius of a set of matrices $\mathcal{M}_2$ is strictly greater than that of a set of matrices $\mathcal{M}_1$, if $\mathcal{M}_1$ is contained in the interior of the convex hull of $\mathcal{M}_2$ where the interior is taken relative to the affine subspace generated by $\mathcal{M}_2$. Note that this result is a bit surprising because a similar statement for the maximum of the spectral radii is false, see for instance [21, Example 12].

In the following statement we denote by $\operatorname{aff}\mathcal{M}$ the affine subspace generated by $\mathcal{M}$, that is, the smallest affine subspace containing $\mathcal{M}$. The relative interior with respect to $\operatorname{aff}\mathcal{M}$ is denoted by $\operatorname{int}_{\operatorname{aff}\mathcal{M}}$.

Proposition 5.2 *Let $\mathcal{M}_1, \mathcal{M}_2 \in I(\mathbb{K}^{n\times n})$ satisfy $\mathcal{M}_1 \neq \mathcal{M}_2$ and*

$$\mathcal{M}_1 \subset \operatorname{int}_{\operatorname{aff}\mathcal{M}_2} \operatorname{conv} \mathcal{M}_2\,, \tag{14}$$

then

$$\rho(\mathcal{M}_1) < \rho(\mathcal{M}_2)\,.$$

Remark 5.3 Note that in the extremal case that $\mathcal{M}_2$ is a singleton set, our assumption (14) does not guarantee that $\mathcal{M}_1 \neq \mathcal{M}_2$, so that the extra assumption is necessary. □

Proof: Assume the assertion is false, so that we may assume $\rho(\mathcal{M}_1) = \rho(\mathcal{M}_2) = 1$. Fix a strictly convex norm $\|\cdot\|$ on $\mathbb{K}^n$. We will show that under our assumption for $x \neq 0$ we have

$$v_1(x) := \max_{S\in\mathcal{S}_{\infty,1}} \|Sx\| < \max_{S\in\mathcal{S}_{\infty,2}} \|Sx\| =: v_2(x)\,. \tag{15}$$

This implies for some $c > 1$ that $v_2(x) > cv_1(x), x \neq 0$ by a compactness argument. By definition of $\mathcal{S}_{\infty,2}$ it follows in particular that for $x_0, v_1(x_0) = 1$ there exists an $S_1 \in \mathcal{S}_2$ with

$$\|S_1 x_0\| > \frac{c}{2}$$

and arguing inductively there are $S_1, \dots, S_k \in \mathcal{S}_2$ with

$$\|S_k \cdots S_1 x_0\| > \left(\frac{c}{2}\right)^k\,.$$

Hence there exists an unbounded trajectory for the discrete inclusion generated by $\mathcal{M}_2$. This implies that $\rho(\mathcal{M}_2) > 1$ as $\mathcal{M}_2$ is irreducible, a contradiction.

Thus it remains to show that (15) holds under the assumption $\rho(\mathcal{M}_1) = \rho(\mathcal{M}_2) = 1$. Note that this assumption implies in particular, that $\mathcal{S}_{\infty,1} \subset \mathcal{S}_{\infty,2}$. Also due to (14) it holds that whenever we have a set of the form

$$D := \{ABx \mid B \in \mathcal{M}_2\}\,,$$

then

$$\max\{\|ABx\| \mid B \in \mathcal{M}_2\} > \max\{\|ABx\| \mid B \in \mathcal{M}_1\}\,,$$

unless D is a singleton set. The reason for this is the linearity of the map $B \mapsto ABx$ and the strict convexity of our norm.

Fix $0 \neq x \in \mathbb{K}^n$ and let $S \in \mathcal{S}_{\infty,1}$ be such that

$$\|Sx\| = v_1(x)\,.$$

By Proposition 3.2 (iv) and Lemma 5.1 we can factorize $S = A_1 \cdots A_k S_k$ with $A_j \in \mathcal{M}_1, j = 1, \dots, k, S_k \in \mathcal{S}_{\infty,1}$ such that the set

$$\{A_1 \cdots A_{k-1} B S_k x \mid B \in \mathcal{M}_2\} \tag{16}$$

consists of more than one element. Then it follows

$$v_2(x) \geq \max\{\|A_1 \cdots A_{k-1} B S_k x\| \mid B \in \mathcal{M}_2\}$$

$$> \max\{\|A_1 \cdots A_{k-1} B S_k x\| \mid B \in \mathcal{M}_1\} = v_1(x)\,.$$

This completes the proof. □

Remark 5.4 It is worth pointing out, that the proof of the above result would be much simplified if we knew, that there exists a strictly convex Barabanov norm v_1 for $\mathcal{M}_1$. In this case (assuming $\rho(\mathcal{M}_1) = 1$) we would conclude immediately from (14) and strict convexity that for each $x \neq 0$ there is some $A \in \mathcal{M}_2$ such that $v_1(Ax) > v_1(x)$, which implies $\rho(\mathcal{M}_1) < \rho(\mathcal{M}_2)$. To show that such an approach is not possible, let us demonstrate that for some irreducible sets of matrices no Barabanov norm is strictly convex.

In fact, we return to the set $\mathcal{M}$ introduced in Example 3.3. As we have already calculated $\mathcal{S}_\infty$, we see immediately, that for any norm w the corresponding Barabanov norm is given by

$$v\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \max\left\{w\left(\begin{bmatrix} x_1 \\ 0 \end{bmatrix}\right), w\left(\begin{bmatrix} 0 \\ x_2 \end{bmatrix}\right)\right\}.$$

This norm is not strictly convex. □

Before we note a consequence for strictly increasing function with values in $\mathcal{K}(\mathbb{K}^{n\times n})$ we need the following remark. If a bounded set $\mathcal{M} \subset \mathbb{K}^{n\times n}$ is reducible, then after a suitable change of coordinates all matrices $A \in \mathcal{M}$ are of the form

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & \dots & A_{1d} \\ 0 & A_{22} & A_{23} & \dots & A_{2d} \\ 0 & 0 & A_{33} & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ & & & & \\ 0 & & \dots & 0 & A_{dd} \end{bmatrix}, \tag{17}$$

where each of the sets $\mathcal{M}_{ii} := \{A_{ii}; A \in \mathcal{M}\}, i = 1 \dots d$ is irreducible. By Lemma 2 (c) in [2] it holds that

$$\rho(\mathcal{M}) = \max_{i=1,\dots,d} \rho(\mathcal{M}_{ii}) \,. \tag{18}$$

Corollary 5.5 *Let $f : \mathbb{R}_+ \to \mathcal{K}(\mathbb{K}^{n\times n})$ be a function such that $f(\theta_1) \subset f(\theta_2)$ satisfy* (14) *for all $\theta_1 < \theta_2 \in \mathbb{R}_+$. Then*

(i) there exists a $\theta_0 \in \mathbb{R}_+$ such that $\rho \circ f$ is constant on $[0, \theta_0)$ and strictly increasing on $[\theta_0, \infty)$,

(ii) if $f(\theta) \in I(\mathbb{K}^{n\times n})$ for some $\theta > 0$, then $\theta_0 < \theta$,

(iii) if f is continuous then $\rho \circ f$ is continuous,

(iv) if f is locally Lipschitz continuous then $\rho \circ f$ is locally Lipschitz continuous on $[0, \infty) \setminus F$, where F contains at most $n - 1$ points.

Proof: (i) The interval $(0, \infty)$ can be partitioned into at most n intervals on which the invariant subspaces of $f(\theta)$ are constant. That is, there are numbers $0 = a_0 < a_1 < a_{k-1} < a_k = \infty$, $k \le n$ such that on $(a_j, a_{j+1}), j = 0, \dots, k-1$ all matrices $A \in \cup_{\theta \in (a_j, a_{j+1})} f(\theta)$ are of a fixed block-diagonal structure of the form (17), where for each $\theta \in (a_j, a_{j+1})$ and each $i = 1, \dots, d$ the set $\mathcal{M}_i(\theta) := \{A_{ii} | A \in f(\theta)\}$ is irreducible. The assumptions do not guarantee that the family $\mathcal{U}_i := \{\mathcal{M}_i(\theta) \mid \theta \in (a_j, a_{j+1})\}$ is strictly increasing. Nevertheless, we know that

$$\operatorname{conv} \mathcal{M}_i(\theta_1) \subset \operatorname{int}_{\operatorname{aff} \mathcal{M}_i(\theta_2)} \operatorname{conv} \mathcal{M}_i(\theta_2) \,,$$

for $\theta_1 < \theta_2 \in (a_j, a_{j+1})$. This implies that the only possibility for $\mathcal{U}_i$ not to be increasing at $\theta_0 \in (a_j, a_{j+1})$ is that $\mathcal{U}_i$ is a singleton set.

Hence, for the map $\rho_i : \theta \mapsto \rho(\mathcal{M}_i(\theta))$, $\theta \in (a_j, a_{j+1})$ there are three possibilities

(i) ρ_i is constant on (a_j, a_{j+1}), if $\mathcal{M}_i(\theta) \equiv \{A_0\}$ on (a_j, a_{j+1}),

(ii) ρ_i is strictly increasing, if $\mathcal{U}_i$ is strictly increasing,

(iii) there is a constant $\theta_0 \in (a_j, a_{j+1})$ such that ρ_i is constant on (a_j, θ_0) and strictly increasing on (θ_0, a_{j+1}).

Due to (18) the same is true for $\rho \circ f$ on (a_j, a_{j+1}). Now it follows that if there are $\theta_1 < \theta_2 \in \mathbb{R}_+$ with $\rho \circ f(\theta_1) < \rho \circ f(\theta_2)$ then $\rho \circ f$ is strictly increasing on $[\theta_2, \infty)$, because in θ_2 the maximum of the joint spectral radii

ρ_i is attained in a block, which necessarily contains two distinct matrices, because otherwise the same value would be attained in θ_1. In this i-th block ρ_i is thus strictly increasing and merging of blocks does not change this. As the assumptions guarantee that $\rho \circ f$ is increasing the only possibility for this function to be constant is on an interval of the form $[0, \theta_0)$. This shows the first assertion.

(ii) is an immediate consequence of Proposition 5.2, while (iii) follows from Lemma 3.5.

(iv) If f is Lipschitz continuous then by Corollary 4.2 ρ_i is locally Lipschitz continuous on the intervals (a_j, a_{j+1}) and thus also the maximum of these functions. Thus F contains at most the points $a_1, \dots, a_{k-1}$, and of these there are at most $n-1$.

□

6 Extremal norms

We now investigate conditions for the existence of extremal norms. For this we need a notion of "defectiveness" of the generalized spectral radius in the case that $\mathcal{M}$ is reducible, which in some sense generalizes the notion of a defective eigenvalue with a modulus equal to the spectral radius. We intend to generalize the well known result that for a matrix A there exists an operator norm v with

$$v(A) = r(A)\,,$$

if and only if all eigenvalues $\lambda \in \sigma(A)$ with $|\lambda| = r(A)$ are nondefective. Unfortunately we are not able to recover the "only if" part of this statement.

For a set $\mathcal{M}$ of matrices of the form (17) let $J := \{1 \le i \le d \mid \rho(\mathcal{M}_{ii}) = \rho(\mathcal{M})\}$ denote the set of indices for which the generalized spectral radius is attained.

Definition 6.1 *A compact set of matrices $\mathcal{M} \subset \mathbb{K}^{n\times n}$ is said to have* nondefective generalized spectral radius *if there is a basis of $\mathbb{K}^n$ such that every matrix $A \in \mathcal{M}$ is of the from* (17) *and for all $i \in J$, $i < j \le \max J$ and all $A \in \mathcal{M}$ it holds that*

$$A_{ij} = 0\,.$$

Note that instead of requiring "zero rows" to the right of $A_{ii}, i \in J$ we could also have required "zero columns", that is for $i \in J$, $i < j \le \max J$, $A \in \mathcal{M}$ we have $A_{ji} = 0$. These two notions are equivalent, as one form is always similar to the other.

In particular, the above definition is satisfied if $\mathcal{M}$ is irreducible. Our proof is based on the following lemma.

Lemma 6.2 *Let $\mathbb{K}^n = \mathbb{K}^m \bigoplus \mathbb{K}^p$ and let $\mathcal{M} \in \mathcal{K}(\mathbb{K}^{n\times n})$ satisfy that every $A \in \mathcal{M}$ is of the form*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

with $A_{11} \in \mathbb{K}^{m\times m}, A_{12} \in \mathbb{K}^{m\times p}, A_{22} \in \mathbb{K}^{p\times p}$. Denote

$$\mathcal{M}_1 := \{A_{11} \mid A \in \mathcal{M}\} \subset \mathbb{K}^{m\times m}, \quad \mathcal{M}_2 := \{A_{22} \mid A \in \mathcal{M}\} \subset \mathbb{K}^{p\times p}.$$

(i) If $\rho(\mathcal{M}_1) < \rho(\mathcal{M}_2)$ and there is an extremal norm v_2 on $\mathbb{K}^p$ corresponding to $\mathcal{M}_2$ then there exists an extremal norm w on $\mathbb{K}^n$ corresponding to $\mathcal{M}$.

(ii) If $\rho(\mathcal{M}_1) > \rho(\mathcal{M}_2)$ and there is an extremal norm v_1 on $\mathbb{K}^m$ corresponding to $\mathcal{M}_1$ then there exists an extremal norm w on $\mathbb{K}^n$ corresponding to $\mathcal{M}$.

Proof: (i) Fix a norm v_1 on $\mathbb{K}^m$ and $\varepsilon > 0$ such that

$$v_1(A_{11}) \le \rho(\mathcal{M}_1) + \varepsilon < \rho(\mathcal{M}_2), \text{ for all } A \in \mathcal{M}.$$

This is possible by Lemma II(b) in [2]. Let v_{12} denote the operator norm from $(\mathbb{K}^p, v_2)$ to $(\mathbb{K}^m, v_1)$ and denote

$$c := \max\{v_{12}(A_{12}) \mid A \in \mathcal{M}\}. \tag{19}$$

Fix $a > 0$ so that $\rho(\mathcal{M}_1) + \varepsilon + c/a < \rho(\mathcal{M}_2)$ and define the norm

$$w(x_1, x_2) := \max\{v_1(x_1), av_2(x_2)\}.$$

We claim that w is an extremal norm for $\mathcal{M}$ on $\mathbb{K}^n$. To this let $A \in \mathcal{M}, x_1 \in \mathbb{K}^m, x_2 \in \mathbb{K}^p$ be arbitrary, then

$$w(Ax) = \max\{v_1(A_{11}x_1 + A_{12}x_2), av_2(A_{22}x_2)\}$$

$$\le \max\{(\rho(\mathcal{M}_1) + \varepsilon)v_1(x_1) + cv_2(x_2), a\rho(\mathcal{M}_2)v_2(x_2)\}.$$

If $v_1(x_1) \le av_2(x_2)$ this implies

$$w(Ax) \le \max\{(\rho(\mathcal{M}_1) + \varepsilon + c/a), \rho(\mathcal{M}_2)\}av_2(x_2) = \rho(\mathcal{M}_2)w(x).$$

Conversely, if $v_1(x_1) > av_2(x_2)$ then

$$w(Ax) \le \max\{(\rho(\mathcal{M}_1) + \varepsilon + c/a), \rho(\mathcal{M}_2)\}v_1(x_1) = \rho(\mathcal{M}_2)w(x).$$

(ii) Again fix a norm v_2 and $\varepsilon > 0$ such that

$$v_2(A_{22}) \le \rho(\mathcal{M}_2) + \varepsilon < \rho(\mathcal{M}_1), \text{ for all } A \in \mathcal{M}.$$

Let v_{12} denote the same operator norm as in (i) and let c be defined as in (19). Fix $a > 0$ such that $\rho(\mathcal{M}_2) + \varepsilon + c/a < \rho(\mathcal{M}_1)$. For the norm defined by

$$w(x_1, x_2) = v_1(x_1) + av_2(x_2), x_1 \in \mathbb{K}^m, x_2 \in \mathbb{K}^p$$

a calculation similar to the one in (i) shows that it is extremal for $\mathcal{M}$. $\square$

Now we are in a position to prove our main result on extremal norms.

Theorem 6.3 *Let $\mathcal{M} \subset \mathbb{K}^{n\times n}$ be compact with nondefective generalized spectral radius. Then there exists an extremal norm for $\mathcal{M}$ on $\mathbb{K}^n$.*

Proof: Assume that we have chosen a basis such that all matrices $A \in \mathcal{M}$ are in the form (17), with $A_{ii} \in \mathbb{K}^{n_i\times n_i}, i = 1, \dots, d$. If $d = 1$ the result is immediate from Theorem 1.1 so assume $d > 1$. Let $J = \{i_1 < \dots < i_k\} \subset \{1, \dots, d\}$ be the set of indices satisfying $\rho(\mathcal{M}_{ii}) = \rho(\mathcal{M})$. We will work inductively backwards on the set J. In the first step consider the matrices

$$\mathcal{M}_k := \left\{ \begin{bmatrix} A_{i_{k-1}+1,i_{k-1}+1} & * & \dots & \dots & * \\ 0 & \ddots & * & \dots & * \\ 0 & 0 & \ddots & * & \vdots \\ \vdots & & \ddots & A_{i_k-1,i_k-1} & * \\ 0 & & \dots & 0 & A_{i_k,i_k} \end{bmatrix} \middle| \; A \in \mathcal{M} \right\}.$$

Note that $\rho(\mathcal{M}_k) = \rho(\mathcal{M})$ and all blocks except for the one in the right lower corner have a generalized spectral radius strictly smaller than $\rho(\mathcal{M})$. Thus Lemma 6.2 (i) applies and there is an extremal norm w_k on

$$\bigoplus_{i=i_{k-1}+1}^{i_k} \mathbb{K}^{n_i}$$

corresponding to $\mathcal{M}_k$. Now on $\bigoplus_{i=i_{k-1}}^{i_k} \mathbb{K}^{n_i}$ all matrices are of the form

$$\begin{bmatrix} A_{i_{k-1},i_{k-1}} & 0 \\ 0 & A_k \end{bmatrix}, \quad A_k \in \mathcal{M}_k.$$

Thus again applying Theorem 1.1 it is clear that there is an extremal norm on $\bigoplus_{i=i_{k-1}}^{i_k} \mathbb{K}^{n_i}$.

Now we may apply the same argument for the blocks corresponding to $\bigoplus_{i=i_{k-2}+1}^{i_k} \mathbb{K}^{n_i}$ to successively obtain extremal norms by repeatedly applying Lemma 6.2 (i). As a result we obtain an extremal norm on $\bigoplus_{i=i_1}^{i_k} \mathbb{K}^{n_i}$. Now the result follows after a further application of Lemma 6.2 (i) and (ii) to the remaining blocks with indices smaller than i_1, respectively larger than i_k. $\square$

Remark 6.4 Note that we cannot assume to be able to order the blocks in an order such that the generalized spectral radii are increasing or decreasing in (17) as this would imply properties of the invariant subspaces of $\mathcal{M}$. For instance for the set

$$\mathcal{M} := \left\{ \begin{bmatrix} \frac{1}{2} & a \\ 0 & 1 \end{bmatrix} \;\middle|\; a \in [0,1] \right\}$$

the only nontrivial invariant subspace is span $[1,0]'$ which is associated to the eigenvalue $1/2$. Hence no similarity transformation will transform $\mathcal{M}$ into a set of matrices of the form

$$\begin{bmatrix} 1 & * \\ 0 & \frac{1}{2} \end{bmatrix}.$$

This somewhat explains the awkward proof of Theorem 6.3. $\square$

A further interesting feature of extremal norms is that they allow to make the inequality in (3) more precise.

Lemma 6.5 *Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Assume that $\mathcal{M} \subset \mathbb{K}^{n \times n}$ is bounded.*

(i) If there exists an extremal norm v for $\mathcal{M}$, then there exists a constant $M > 0$ such that for all $t \geq 1$

$$|\log \hat{\rho}_t(\mathcal{M}) - \log \rho(\mathcal{M})| < M t^{-1}.$$

(ii) Otherwise there exists an $M > 0$ such that for all $t \geq 1$

$$|\log \hat{\rho}_t(\mathcal{M}) - \log \rho(\mathcal{M})| < M \frac{1 + \log t}{t}.$$

Proof: Let v be the extremal norm for $\mathcal{M}$. As all norms on finite dimensional vector spaces are equivalent it follows with (3) that

$$0 \leq \frac{1}{t} \log \sup_{S_t \in \mathcal{S}_t} \|S_t\| - \log \rho(\mathcal{M}) \leq \frac{1}{t} \log \sup_{S_t \in \mathcal{S}_t} cv(S_t) - \rho(\mathcal{M}) = \frac{1}{t} \log c. \quad (20)$$

This proves the assertion.
(ii) This follows from Lemma 2.3 in [29]. $\square$

Remark 6.6 Note that we cannot expect a similar statement for the lower bound $\overline{\rho}_t$. If we return to our Example 3.3 then we see that in this case $\overline{\rho}_{2k}(\mathcal{M}) = \rho(\mathcal{M}) = 1$ and $\overline{\rho}_{2k+1} = 0$ for all $k \in \mathbb{N}$. $\Box$

We also note the following consequence of local uniform convergence of $\hat{\rho}(\mathcal{M})$ to $\rho(\mathcal{M})$.

Corollary 6.7 *Let $P \subset I(\mathbb{K}^{n\times n})$ be compact then there is a constant $M > 0$ such that for all $\mathcal{M} \in P$ and all $t \geq 1$ it holds that*

$$|\hat{\rho}_t(\mathcal{M}) - \rho(\mathcal{M})| < \rho(\mathcal{M})(M^{1/t} - 1)\,,$$

i.e. $\hat{\rho}_t$ converges locally uniformly to $\rho(\mathcal{M})$ on $I(\mathbb{K}^{n\times n})$.

Proof: Just note that the constant c in the proof of Lemma 6.5 (i) can be chosen independently of $\mathcal{M} \in P$ by Theorem 4.1. $\Box$

7 Conclusion

We have studied extremal norms for linear discrete and differential inclusions. For the special case of irreducible inclusions we give a constructive procedure for a special extremal norm. This approach yields Lipschitz continuity of the generalized spectral radius and a monotonicity property as a byproduct. A more general sufficient criterion guaranteeing the existence of an extremal norm has also been presented. Furthermore, we have pointed out that the convergence of $\hat{\rho}_t$ to the generalized spectral radius is linear if an extremal norm exists, in particular in the irreducible case.

References

[1] N. E. Barabanov. Lyapunov indicator of discrete inclusions. I–III. *Autom. Remote Control*, 49:152–157, 283–287, 558–565, 1988.

[2] M. A. Berger and Y. Wang. Bounded semigroups of matrices. *Lin. Alg. Appl.*, 166:21–27, 1992.

[3] V. D. Blondel and J. N. Tsitsiklis. The boundedness of all products of a pair of matrices is undecidable. *Systems & Control Letters*, 41(2):135–140, 2000.

[4] T. Bousch and J. Mairesse. Asymptotic height optimization for topical IFS, Tetris heaps and the finiteness conjecture. submitted, 2000.

[5] F. Colonius and W. Kliemann. The Lyapunov spectrum of families of time varying matrices. *Trans. Am. Math. Soc.*, 348:4389–4408, 1996.

[6] F. Colonius and W. Kliemann. *The Dynamics of Control.* Birkhäuser, Boston, 2000.

[7] I. Daubechies and J. C. Lagarias. Sets of matrices all infinite products of which converge. *Lin. Alg. Appl.*, 161:227–263, 1992.

[8] L. Elsner. The generalized spectral-radius theorem: An analytic-geometric proof. *Linear Algebra and its Applications*, 220:151–159, 1995.

[9] L. Elsner and W.-J. Beyn. Infinite products and paracontracting matrices. *Electr. J. Linear Algebra*, 2:1–8, 1997.

[10] L. Elsner and S. Friedland. Norm conditions for convergence of infinite products. *Lin. Alg. Appl.*, 250:133–142, 1997.

[11] G. Gripenberg. Computing the joint spectral radius. *Lin. Alg. Appl.*, 234:43–60, 1996.

[12] N. Guglielmi and M. Zennaro. On the asymptotic properties of a family of matrices. Quaderni matematici, II Serie 447, Università degli Studi di Trieste, January 1999.

[13] N. Guglielmi and M. Zennaro. On the zero-stability of variable stepsize multistep methods: The spectral radius approach. Quaderni matematici, II Serie 452, Università degli Studi di Trieste, May 1999.

[14] L. Gurvits. Stability of discrete linear inclusions. *Lin. Alg. Appl.*, 231:47–85, 1995.

[15] C. Heil and G. Strang. *Linear Algebra for Signal Processing*, volume 69 of *IMA Vol. Math. Appl.*, chapter Continuity of the joint spectral radius: Application to wavelets, pages 51–61. Springer-Verlag, New York, 1995.

[16] V. S. Kozyakin. Algebraic unsolvability of problem of absolute stability of desynchronized systems. *Autom. Rem. Control*, 51(6):754–759, 1990.

[17] J. C. Lagarias and Y. Wang. The finiteness conjecture for the generalized spectral radius of a set of matrices. *Lin. Alg. Appl.*, 214:17–42, 1995.

[18] J. Levitzki. Über nilpotente Unterringe. *Mathematische Annalen*, 105:620–627, 1931.

[19] M. Maesumi. An efficient lower bound for the generalized spectral radius of a set of matrices. *Lin. Alg. Appl.*, 240:1–7, 1996.

[20] M. Omladič and H. Radjavi. Irreducible semigroups with multiplicative spectral radius. *Lin. Alg. Appl.*, 251:59–72, 1997.

[21] A. D. B. Paice and F. R. Wirth. Analysis of the local robustness of stability for maps. In *Proceedings of the 4th European Control Conference 1997*, pages TU–E–B–4, Brussels, Belgium, July 1997.

[22] H. Radjavi. On irreducibility of semigroups of compact operators. *Indiana Univ. Math. J.*, 39:499–515, 1990.

[23] G.-C. Rota and G. Strang. A note on the joint spectral radius. *Indag. Math.*, 22:379–381, 1960.

[24] M. H. Shih, J. W. Wu, and C. T. Pang. Asymptotic stability and generalized gelfand spectral radius formula. *Linear Algebra and its Applications*, 252:61–70, 1997.

[25] J. N. Tsitsiklis and V. Blondel. The Lyapunov exponent and joint spectral radius of pairs of matrices are hard - when not impossible - to compute and to approximate. *Math. Contr., Sign., and Syst.*, pages 31–40, 1997.

[26] A. Vladimirov, L. Elsner, and W.-J. Beyn. Stability and paracontractivity of discrete linear inclusions. *Lin. Alg. Appl.*, 312:125–134, 2000.

[27] F. Wirth. Dynamics of time-varying discrete-time linear systems: Spectral theory and the projected system. *SIAM J. Control Optim.*, 36(2):447–487, 1998.

[28] F. Wirth. On stability of infinite-dimensional discrete inclusions (summary). *J. Math. Syst. Estim. Control*, 8(4):507–510, 1998. Full electronic manuscript, 21 pp. 429 KB, ftp://trick.ntp.springer.de/pub/jmsec/, Retrieval code 72701.

[29] F. Wirth. On the calculation of real time-varying stability radii. *Int. J. Robust & Nonlinear Control*, 8:1043–1058, 1998.

Berichte aus der Technomathematik ISSN 1435-7968

http://www.math.uni-bremen.de/zetem/berichte.html

— Vertrieb durch den Autor —

Reports **Stand: 12. Oktober 2000**

98–01. Peter Benner, Heike Faßbender:
An Implicitly Restarted Symplectic Lanczos Method for the Symplectic Eigenvalue Problem, Juli 1998.

98–02. Heike Faßbender:
Sliding Window Schemes for Discrete Least-Squares Approximation by Trigonometric Polynomials, Juli 1998.

98–03. Peter Benner, Maribel Castillo, Enrique S. Quintana-Ortí:
Parallel Partial Stabilizing Algorithms for Large Linear Control Systems, Juli 1998.

98–04. Peter Benner:
Computational Methods for Linear–Quadratic Optimization, August 1998.

98–05. Peter Benner, Ralph Byers, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Solving Algebraic Riccati Equations on Parallel Computers Using Newton's Method with Exact Line Search, August 1998.

98–06. Lars Grüne, Fabian Wirth:
On the rate of convergence of infinite horizon discounted optimal value functions, November 1998.

98–07. Peter Benner, Volker Mehrmann, Hongguo Xu:
A Note on the Numerical Solution of Complex Hamiltonian and Skew-Hamiltonian Eigenvalue Problems, November 1998.

98–08. Eberhard Bänsch, Burkhard Höhn:
Numerical simulation of a silicon floating zone with a free capillary surface, Dezember 1998.

99–01. Heike Faßbender:
The Parameterized SR Algorithm for Symplectic (Butterfly) Matrices, Februar 1999.

99–02. Heike Faßbender:
Error Analysis of the symplectic Lanczos Method for the symplectic Eigenvalue Problem, März 1999.

99–03. Eberhard Bänsch, Alfred Schmidt:
Simulation of dendritic crystal growth with thermal convection, März 1999.

99–04. Eberhard Bänsch:
Finite element discretization of the Navier-Stokes equations with a free capillary surface, März 1999.

99–05. Peter Benner:
Mathematik in der Berufspraxis, Juli 1999.

99–06. Andrew D.B. Paice, Fabian R. Wirth:
Robustness of nonlinear systems and their domains of attraction, August 1999.

99–07. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Balanced Truncation Model Reduction of Large-Scale Dense Systems on Parallel Computers, September 1999.

99–08. Ronald Stöver:
Collocation methods for solving linear differential-algebraic boundary value problems, September 1999.

99–09. Huseyin Akcay:
Modelling with Orthonormal Basis Functions, September 1999.

99–10. Heike Faßbender, D. Steven Mackey, Niloufer Mackey:
Hamilton and Jacobi come full circle: Jacobi algorithms for structured Hamiltonian eigenproblems, Oktober 1999.

99–11. Peter Benner, Vincente Hernández, Antonio Pastor:
On the Kleinman Iteration for Nonstabilizable System, Oktober 1999.

99–12. Peter Benner, Heike Faßbender:
A Hybrid Method for the Numerical Solution of Discrete-Time Algebraic Riccati Equations, November 1999.

99–13. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Numerical Solution of Schur Stable Linear Matrix Equations on Multicomputers, November 1999.

99–14. Eberhard Bänsch, Karol Mikula:
Adaptivity in 3D Image Processing, Dezember 1999.

00–01. Peter Benner, Volker Mehrmann, Hongguo Xu:
Perturbation Analysis for the Eigenvalue Problem of a Formal Product of Matrices, Januar 2000.

00–02. Ziping Huang:
Finite Element Method for Mixed Problems with Penalty, Januar 2000.

00–03. Gianfrancesco Martinico:
Recursive mesh refinement in 3D, Februar 2000.

00–04. Eberhard Bänsch, Christoph Egbers, Oliver Meincke, Nicoleta Scurtu:
Taylor-Couette System with Asymmetric Boundary Conditions, Februar 2000.

00–05. Peter Benner:
Symplectic Balancing of Hamiltonian Matrices, Februar 2000.

00–06. Fabio Camilli, Lars Grüne, Fabian Wirth:
A regularization of Zubov's equation for robust domains of attraction, März 2000.

00–07. Michael Wolff, Eberhard Bänsch, Michael Böhm, Dominic Davis:
Modellierung der Abkühlung von Stahlbrammen, März 2000.

00–08. Stephan Dahlke, Peter Maaß, Gerd Teschke:
Interpolating Scaling Functions with Duals, April 2000.

00–09. Jochen Behrens, Fabian Wirth:
A globalization procedure for locally stabilizing controllers, Mai 2000.

00–10. Peter Maaß, Gerd Teschke, Werner Willmann, Günter Wollmann:
Detection and Classification of Material Attributes – A Practical Application of Wavelet Analysis, Mai 2000.

00–11. Stefan Boschert, Alfred Schmidt, Kunibert G. Siebert, Eberhard Bänsch, Klaus-Werner Benz, Gerhard Dziuk, Thomas Kaiser:
Simulation of Industrial Crystal Growth by the Vertical Bridgman Method, Mai 2000.

00–12. Volker Lehmann, Gerd Teschke:
Wavelet Based Methods for Improved Wind Profiler Signal Processing, Mai 2000.

00–13. Stephan Dahlke, Peter Maass:
A Note on Interpolating Scaling Functions, August 2000.

00–14. Ronny Ramlau, Rolf Clackdoyle, Frédéric Noo, Girish Bal:
Accurate Attenuation Correction in SPECT Imaging using Optimization of Bilinear Functions and Assuming an Unknown Spatially-Varying Attenuation Distribution, September 2000.

00–15. Peter Kunkel, Ronald Stöver:
Symmetric collocation methods for linear differential-algebraic boundary value problems, September 2000.

00–16. Fabian Wirth:
The generalized spectral radius and extremal norms, Oktober 2000.