Generalized Markov–Bernstein Inequalities and Stability of Dynamical Systems

Vladimir Yu. Protasov

Received January 11, 2022; revised March 25, 2022; accepted March 31, 2022

Abstract—We analyze the Markov–Bernstein type inequalities between the norms of functions and of their derivatives for complex exponential polynomials. We establish a relation between the sharp constants in these inequalities and the stability problem for linear switching systems. In particular, the maximal discretization step is estimated. We prove the monotonicity of the sharp constants with respect to the exponents, provided those exponents are real. This gives asymptotically tight uniform bounds and the general form of the extremal polynomial. The case of complex exponent is left as an open problem.

DOI: 10.1134/S0081543822050169

1. INTRODUCTION

The Markov–Bernstein type inequalities relate the norm of a polynomial to the norm of its \( \ell \)th-order derivative. The Bernstein inequality states that for a trigonometric polynomial of degree \( n \) one has \( \| p' \| \leq n \| p \| \), where \( \| p \| = \sup_{t \in \mathbb{T}} | p(t) | \) is the uniform norm on the period. The Markov inequality states that for an algebraic polynomial of degree \( n \) on the interval \( [-1, 1] \) one has \( \| p' \| \leq n^2 \| p \| \), and the equality is attained precisely when \( p \) is proportional to the Chebyshev polynomial \( T_n \). We are interested in generalizations of these inequalities to the exponential polynomials, i.e., to the functions of the form \( p(t) = \sum_{k=1}^{n} p_k e^{-h_k t} \), where \( h_1, \ldots, h_n \) are given complex numbers (in the case of multiplicity \( m \) the corresponding exponential is multiplied by the powers \( t^k, k = 0, \ldots, m - 1 \)). For each \( \ell \geq 1 \), for arbitrary \( h_k \), and for an interval (domain), we consider the inequality \( \| p^{(\ell)} \| \leq C \| p \| \) with the sharp constant \( C \). The case where \( \ell = 1 \) and the exponents are purely imaginary and form an arithmetic progression corresponds (after a proper change of variables) to the Bernstein inequality. If the exponents are real and form an arithmetic progression, then this case corresponds to the Markov inequality.

Note that in the case of real rational exponents \( h_k \), the exponential polynomial \( p(t) \) is transformed by the change of variables \( x = e^{-t/N} \) to an algebraic polynomial \( P(x) \). However, this does not solve the problem of finding the sharp constant, because the polynomial \( P(x) \) is not arbitrary: its degree can significantly exceed the number of its nonzero coefficients. Therefore, the constant obtained in this way may be much greater than the sharp constant.

We consider inequalities for derivatives on the positive half-axis \( \mathbb{R}_+ \), naturally assuming that \( \text{Im} \ h_k > 0 \) and, therefore, all the polynomials \( p(t) \) tend to zero as \( t \to +\infty \). The main problem is to find the exact value or at least to estimate the sharp constant \( C \) in the inequality \( \| p^{(\ell)} \| \leq C \| p \| \). For the derivative \( p^{(\ell)} \), we always choose the norm in \( C(\mathbb{R}_+) \), while the norm \( \| p \| \) of the polynomial can be more general. In particular, some of the results are obtained for an arbitrary monotone norm on \( \mathbb{R}_+ \) (the monotonicity means that \( \| f \| \geq \| g \| \) if \( | f(t) | \geq | g(t) | \) at every point \( t \in \mathbb{R}_+ \)).

\(^{a}\) University of L’Aquila, piazza Santa Margherita 2, 67100 L’Aquila, Italy.
\(^{b}\) Lomonosov Moscow State University, Moscow, 119991 Russia.

E-mail address: v-protassov@yandex.ru

237
We are interested both in the values of the constants $C = C(\ell, \mathbf{h})$ in the Markov–Bernstein inequality for fixed exponents $\mathbf{h}$ and in the uniform constants for all values of $\mathbf{h}$ from a given domain. This problem is motivated by applications to the study of trajectories of linear systems.

**Problem of the maximal initial velocity of a bounded trajectory.** If a function $x \in C^n([0, +\infty))$ is a solution of a linear ODE with constant coefficients $x^{(n)} = \sum_{k=0}^{n-1} a_k x^{(k)}(t)$, then how large can its initial derivative $x'(0)$ be under the condition $\|x\|_{C([0, +\infty))} \leq R$, where $R > 0$ is a fixed number? In other words, with what speed can the trajectory be launched so that it does not leave the ball of radius $R$? The same question can be formulated for the higher order derivatives $x^{(l)}(0)$. The answer is that $\|x^{(l)}(0)\| \leq R C(\ell, \mathbf{h})$, where $C(\ell, \mathbf{h})$ is the sharp constant in the Markov–Bernstein inequality for exponential polynomials in the system $\{e^{-h_k t}\}_{k=1}^n$ and $\{-h_k\}_{k=1}^n$ are the roots of the characteristic polynomial of the ODE. Similar problems are considered for other norms, for example, for $\|x\|_{L_p} \leq R$.

**Stability problem for linear switching systems.** This is probably the most important application and the main motivation of this research. We address it in Section 7. A linear switching system is a linear ODE $\dot{x}(t) = A(t)x(t)$, $x(0) = x_0$, where $x(t) \in \mathbb{R}^n$ and $A(\cdot)$ is an arbitrary measurable function taking values in some compact set of $n \times n$ matrices $\mathcal{A}$. Such systems regularly arise in engineering applications. Their systematic study began in the late 1970s (see [20, 24–27, 30] and the references therein). A system is asymptotically stable if its trajectories $x(t)$ tend to zero as $t \to +\infty$ for all functions $A(\cdot)$. One of the methods for proving stability is the discretization of the system, i.e., its approximation by a discrete-time system: $x(k+1) = (I + \tau A(k))x(k)$, $k = 0, 1, \ldots$, where $I$ is the identity matrix. It is known [27] that if the system becomes stable after the discretization with some step $\tau$, then it is stable for all smaller steps and is also stable as a continuous-time system. The problem consists in finding the longest possible step $\tau$ for which the converse is also true with some precision $\varepsilon > 0$. This means that if the system $\dot{x} = Ax$ is stable, then the discrete system $x(k+1) = (I + \tau (A(k) - \varepsilon I))x(k)$ is stable as well. Since there are efficient methods for determining the discrete-time stability [14, 16, 22, 28, 33], it follows that an efficient estimate for the discretization step $\tau$ makes those methods also applicable to continuous-time systems. This idea was developed in [2, 10, 31, 32] and other papers. In Section 7 we estimate $\tau$ in terms of the sharp constant $C(\ell, \mathbf{h})$ in the Markov–Bernstein inequality for $\ell = 2$ and $\mathbf{h} = -\text{sp}(A)$, $A \in \mathcal{A}$, where $\text{sp}(A)$ denotes the set of eigenvalues of the matrix $A$. These results are formulated in Theorems 4 and 5.

**Main results.** We obtain estimates for the step size $\tau$ in terms of the constants $C(\ell, \mathbf{h})$ for arbitrary vectors $\mathbf{h}$, while the estimates and the sharp values of these constants are found only for the real vectors. In the latter case the system of exponents is a Chebyshev system, so, to find the extremal function, we can apply the alternance. A similar idea of estimating the constants in the Markov–Bernstein inequality for real exponential polynomials have been exploited in [6–9, 23, 29, 34]. We establish a comparison theorem (Theorem 1) according to which the constant $C(\ell, \mathbf{h})$ is strictly decreasing in each exponent $h_k$. This is true not only for the uniform norm in $C(\mathbb{R}_+)$ but also for every $L_p$-norm and, more generally, for all monotone norms on $\mathbb{R}_+$. Therefore, this constant attains its greatest value on the set $h_k \leq \alpha_k$, $k = 1, \ldots, n$, at the point $\mathbf{h} = (\alpha_1, \ldots, \alpha_n)$. This result allows us to obtain uniform estimates for the constant $C(\ell, \mathbf{h})$ for various ranges of the vector $\mathbf{h}$. For example, under the conditions $h_k \leq 1$, $k = 1, \ldots, n$, the maximum values are attained for the polynomial $p(t) = e^{-t}R_{n-1}(t)$, where $R_{n-1}$ is the Chebyshev polynomial with the Laguerre weight $e^{-t}$ (Theorem 2). Applying known estimates for the derivative of those polynomials [11, 13, 23, 36] and especially the asymptotically sharp bound of V. P. Sklyarov [35], we obtain uniform estimates for the constant in the Markov–Bernstein inequality (Theorem 3). They lead to practically acceptable estimates for the discretization step $\tau$ in the stability problem for linear switching systems. Those estimates, however, can be used only for systems of real exponents, i.e., for switching systems defined by matrices with purely real spectra.
Problem of generalization to arbitrary exponents. A significant disadvantage of the obtained estimates is that they hold for real exponents. The same applies to the comparison theorem. This restriction looks especially strange in the problem of the discretization step length, where our estimates are true only when all the matrices of the system have real spectrum. This fact does not cause problems for a particular switching system, since the value $C(\ell, h)$ can be found for every $h$ as the solution of a convex problem $p^{(\ell)}(0) \to \max, \|p\| \leq 1$. However, to derive uniform estimates, one needs the comparison theorem, whose proof cannot be extended to arbitrary systems of exponents since they do not form a Chebyshev system. We leave the proof of the comparison theorem for complex exponents as an open problem (Conjecture 1). If the answer is affirmative, then all estimates obtained for real exponents can be extended to the complex case.

Notation. We denote vectors by bold letters and scalars by standard (light) letters. Thus, $h = (h_1, \ldots, h_n)$. We use the standard notation $\mathbb{R}_+$ and $\mathbb{R}_{++}$ for the sets of nonnegative reals and positive reals, respectively; similarly, $C_{++} = \{z \in \mathbb{C} \mid \text{Re } z > 0\}$ is an open right half-plane of the complex plane; the $L_p$-norm on $\mathbb{R}_+$ will be denoted by $\|\cdot\|_p$. In particular, if $f \in C(\mathbb{R}_+)$, then $\|f\|_\infty = \|f\|_{C(\mathbb{R}_+)}$.

2. STATEMENT OF THE PROBLEM

Given a vector $h = (h_1, \ldots, h_n) \in \mathbb{C}_{++}^n$, we consider the exponential system $E_h = \{e^{-h_k t}\}_{k=1}^n$. Some of the numbers $h_k$ may coincide, in which case the corresponding exponentials are multiplied by powers of $t$. If, for instance, the components $h_1, \ldots, h_r$ are equal and are different from the others, i.e., the exponent $h_1$ has multiplicity $r$, then the functions $e^{-h_1t}, \ldots, e^{-h_rt}$ are replaced by $e^{-h_1t}, te^{-h_1t}, \ldots, t^{r-1}e^{-h_1t}$, respectively. A polynomial in the system $\{e^{-h_k t}\}_{k=1}^n$, or a quasipolynomial, is an arbitrary linear combination of these exponentials with complex coefficients. The linear space of polynomials in a given system on the half-line $\mathbb{R}_+$ is denoted by $\mathcal{P}_h$. This is an $n$-dimensional subspace of the space $C_0(\mathbb{R}_+)$ of continuous functions on $\mathbb{R}_+$ that tend to zero as $t \to +\infty$. The map $h \mapsto \mathcal{P}_h$ is well defined and continuous [17, 19].

The real part of a quasipolynomial is a real linear combination of the functions $t^me^{-\alpha_k t}\cos \beta_k t$ and $t^me^{-\alpha_k t}\sin \beta_k t$, $k = 1, \ldots, n$, where $\alpha_k$ and $\beta_k$ are the real and imaginary parts of $h_k$, and the degree $m$ is less than the multiplicity of $h_k$. The linear combinations of these functions with real coefficients form the space of real quasipolynomials $\Re \mathcal{P}_h$.

Consider an arbitrary monotone norm $\|\cdot\|$ in the space $\mathcal{P}_h$. The monotonicity means that if $|f_1(t)| \geq |f_2(t)|$ for all $t \geq 0$, then $\|f_1\| \geq \|f_2\|$. For example, the $L_p$-norms, $p \in [1, +\infty]$, and the weighted $L_p$-norms are monotone. For every monotone norm and for an arbitrary convex functional $F: \mathcal{P}_h \to \mathbb{R}$, we consider the following problem:

$$F(p) \to \min, \quad p \in \mathcal{P}_h, \quad \|p\| \leq 1. \quad (2.1)$$

The value of this problem (the minimum value of the objective function) will be denoted by $\Phi(h)$. Since this problem is convex, $\Phi(h)$ can be found by standard tools of convex programming. However, we are interested not in the numerical solution but in the description of the extremal polynomial and in uniform estimates for $\Phi(h)$ over all vectors $h$ from a given set. We will deal with the sets $\mathcal{D}_n = \{z \in \mathbb{C}^n \mid |z_k| \leq 1, \text{Re } z_k > 0, k = 1, \ldots, n\}$ and $\mathcal{I}_n = \{x \in \mathbb{R}^n \mid 0 < x_k \leq 1, k = 1, \ldots, n\}$. We also use the simplified notation $\mathcal{D}_1 = \mathcal{D}$ (half-disc) and $\mathcal{I}_1 = \mathcal{I}$ (half-open interval $(0,1)$).

In this section we consider problem (2.1) in the uniform norm $\|\|_{\infty}$ on $\mathbb{R}_+$ for the functional $F(p) = -p^{(\ell)}(0)$, where $p^{(\ell)}$ is the $\ell$th derivative of $p$, $\ell \geq 1$. Thus, one finds the value of the $\ell$th derivative of the polynomial $p$ at zero under the assumption $\|p\|_{\infty} \leq 1$. This is equivalent to finding the maximal norm $\|p^{(\ell)}\|_{\infty}$ on the unit ball, which is equal to the sharp constant in the Markov–Bernstein inequality for polynomials from $\mathcal{P}_h$. 

PROCEEDINGS OF THE STEKLOV INSTITUTE OF MATHEMATICS Vol. 319 2022
Definition 1. Given a vector $h \in \mathbb{C}^n_{++}$, we set

$$M_\ell(h) = \max_{\|p\|_{\infty} \leq 1, \ p \in \mathcal{P}_h} \|p^{(\ell)}\|_{\infty} \quad \text{and} \quad M_{\ell,n} = \max_{h \in \mathcal{D}_n} M_\ell(h).$$

If we search the extremal polynomial among exponential polynomials with real exponents $h_i$, then $h$ is a real positive vector and the set $\mathcal{D}_n$ is replaced by the nonclosed unit cube $I_n = (0,1)^n$:

$$m_\ell(h) = \max_{\|p\|_{\infty} \leq 1, \ p \in \mathcal{RP}_h} \|p^{(\ell)}\|_{\infty} \quad \text{and} \quad m_{\ell,n} = \max_{h \in I_n} m_\ell(h).$$

Clearly, $m_{\ell,n} \leq M_{\ell,n}$. Our main conjecture is that those two numbers are actually equal.

Conjecture 1. For all $\ell, n \in \mathbb{N}$, we have $m_{\ell,n} = M_{\ell,n}$.

The value $M_\ell(h)$ is the sharp constant in the Markov–Bernstein type inequality for exponential polynomials: $\|p^{(\ell)}\|_{\infty} \leq M_\ell(h)\|p\|_{\infty}, \ p \in \mathcal{P}_h$. In fact, the same constant $M_\ell(h)$ is also sharp in the same inequality for the class of real quasipolynomials $\mathcal{RP}_h$. To see this, it suffices to observe that multiplying the polynomial $p$ by a proper number $e^{\alpha t}$, $\alpha \in \mathbb{R}$, and translating, we obtain the equality $\|p^{(\ell)}\|_{\infty} = \Re p^{(\ell)}(0)$. Thus, $M_\ell(h)$ is the value of the problem $\Re p^{(\ell)}(0) \to \max, \|p\|_{\infty} \leq 1$. The replacement of the polynomial $p$ by its real part, which is a real quasipolynomial $q$, does not change the value $\Re p^{(\ell)}(0) = q^{(\ell)}(0)$ and does not increase the norm. Therefore, the restriction of our problem to the class of real quasipolynomials does not change its value.

Then we will return to an arbitrary monotone norm in problem (2.1) and will focus on real exponents $h_1, \ldots, h_n$. We will prove the comparison theorem (Theorem 1) according to which the value of the problem is strictly increasing in each component $h_i$. Therefore, the value $m_\ell(h)$ is also increasing in each component $h_i$. This allows us to find the optimal value over various domains of the parameter $h$ and obtain uniform bounds over those domains. Further we apply these estimates to the problem of stability of linear switching systems.

3. THE CASE OF REAL EXPONENTS

A system $E_h = \{e^{-h_k t}\}_{k=1}^n$ is called real if $h_k$ are all real. In this case all these numbers are strictly positive and are assumed to be arranged in ascending order: $0 < h_1 \leq \ldots \leq h_n$. The space of polynomials in the system $E_h$ on $\mathbb{R}_+$ with real coefficients is called the space of real exponential polynomials (or real quasipolynomials) and is denoted by $\mathcal{RP}(h)$. The main results of this section are easily generalized to polynomials with complex coefficients (but still with real exponents $h_i$). For the sake of simplicity, we consider only the case of real coefficients. We need several basic properties of the space $\mathcal{RP}(h)$.

1. $E_h$ is a Cartesian system; i.e., it satisfies the Cartesian rule of counting zeros (see, for example, [12]). This implies in particular that if a polynomial $p \in \mathcal{P}_h$ has $n-1$ zeros on $\mathbb{R}_+$ counting multiplicities, then all its coefficients are nonzero and their signs alternate.

The Cartesian property can be proved by approximating all the numbers $h_i$ by rational numbers $\tilde{h}_i$ and by changing the variables $x = e^{-\tilde{h}_i t/N}$, where $N$ is such that $\tilde{h}_i N \in \mathbb{N}, \ i = 1, \ldots, n$. Thus, $E_h$ can be approximated with arbitrary precision by algebraic polynomials, which, as we know, form a Cartesian system.

2. Every Cartesian system is also a Chebyshev system (or Haar system); i.e., every nontrivial polynomial from $\mathcal{P}_h$ has at most $n-1$ zeros (see [12, 18, 19]). Thus, $E_h$ is a Chebyshev system on $\mathbb{R}_+$. For every $n$, for arbitrary $t_i \in \mathbb{R}_+$, and for arbitrary numbers $c_i \in \mathbb{R}, \ i = 1, \ldots, n$, there exists a polynomial $p \in \mathcal{P}_h$ for which $p(t_i) = c_i$.

3. For each $\ell \in \mathbb{N}$, the $\ell$th derivative of the polynomial $p \in \mathcal{P}_h$ has at most $n - \ell$ zeros on $\mathbb{R}_+$ counting multiplicities.
Indeed, if we denote all zeros of \( p \) by \( a_1 < \ldots < a_k < +\infty \), assuming now that they are all different, and adding an extra zero \( a_{k+1} = +\infty \) (since \( p(+\infty) = 0 \)), we conclude that each interval \((a_i, a_{i+1})\), \( i = 1, \ldots, k \), contains at least one zero of the derivative \( p' \). Therefore, \( p' \) has at least \( k \) zeros on \( \mathbb{R}_+ \). Applying induction, we extend this property to all derivatives \( p^{(\ell)} \). The case of multiple roots follows by a limit passage.

4. COMPARISON THEOREM

For a system of real exponents \( \mathbf{h} \), it is possible not only to compute the value \( m_\ell(\mathbf{h}) \) but also to analyze its behaviour as a function of the arguments \( h_1, \ldots, h_n \). In other words, one can find the asymptotics of the \( \ell \)th derivative of the polynomial in the unit ball of the space \( \mathcal{P}_h \) depending on the vector \( \mathbf{h} \). Moreover, this can be done not only for the unit ball in the uniform norm \( \| \cdot \|_\infty \) but also for every monotone norm in \( \mathbb{R}_+ \), in particular, for the \( L_p \)-norm. As we mentioned in the previous section, it is sufficient to solve this problem in the space of quasipolynomials with real coefficients \( \mathcal{RP}_h \).

Consider problem (2.1) for the functional \( F(p) = -p^{(\ell)}(0) \) and for a fixed monotone norm \( \| \cdot \| \) in \( \mathcal{RP}_h \). Thus, it suffices to find the maximum of the \( \ell \)th derivative at zero under the assumption \( \| \cdot \| \leq 1 \). Denote by \( \Phi_\ell(\mathbf{h}) \) the value of this problem. The polynomial \( p \) on which this maximum is attained has \( n - 1 \) roots (counting multiplicities) on the ray \( \mathbb{R}_+ \). Otherwise, there would exist a polynomial \( q \in \mathcal{P}_h \) that has the same roots as \( p \) and is such that \( p(t)q(t) < 0 \) for all \( t \in \mathbb{R}_+ \) different from the roots of \( p \) and for which \( p^{(\ell)}(0)q^{(\ell)}(0) > 0 \). Then for sufficiently small \( \lambda \), we have \( \|p + q\| < 1 \) and \( F(p + q) > F(p) \), which contradicts the optimality of \( p \).

**Theorem 1** (comparison theorem). Let \( \| \cdot \| \) be an arbitrary monotone norm in \( C_0(\mathbb{R}_+) \) and \( \Phi_\ell(\mathbf{h}) \) be the maximum value of \( p^{(\ell)}(0) \) for all possible \( p \in \mathcal{RP}_h \) such that \( \|p\| \leq 1 \). If the vectors \( \mathbf{h}', \mathbf{h} \in \mathbb{R}_{++} \) are such that \( \mathbf{h}' \geq \mathbf{h} \) and \( \mathbf{h}' \neq \mathbf{h} \), then for all \( \ell \in \mathbb{N} \) we have \( \Phi_\ell(\mathbf{h}') > \Phi_\ell(\mathbf{h}) \).

We see that the value of problem (2.1) increases in every variable \( h_i \). Applying this theorem for the \( L_\infty \)-norm, we obtain

**Corollary 1.** If \( \mathbf{h}' \geq \mathbf{h} \) and \( \mathbf{h}' \neq \mathbf{h} \), then for each \( \ell \in \mathbb{N} \) we have \( m_\ell(\mathbf{h}') > m_\ell(\mathbf{h}) \).

For \( \ell = 1 \), this corollary is analogous to the comparison theorem for hyperbolic sines [8], but the method of the proof is different and is based on the following key fact: a small perturbation of a Chebyshev system \( E_{\mathbf{h}} \) generates a larger Chebyshev system (i.e., a system of a greater number of functions). This provides an additional degree of freedom for the choice of the parameter and makes it possible to reduce the norm of the polynomial and to simultaneously increase the value of the objective function \( F \).

We will use the following fact that can be easily derived from the convexity of a norm.

**Lemma 1.** If \( p \) and \( q \) are elements of a normed space, \( \|p\| = 1 \), and \( \|p + q\| < 1 \), then there exists a positive constant \( c \) such that \( \|p + \lambda q\| < 1 - \lambda c \) for all \( \lambda \in (0, 1] \).

**Proof of Theorem 1.** Without loss of generality we assume that \( p^{(\ell)}(0) > 0 \). It suffices to consider the case when the vectors \( \mathbf{h} \) and \( \mathbf{h}' \) differ in only one coordinate, say, the \( r \)th: \( \mathbf{h}' = \mathbf{h} + \tau e_r \), \( \tau > 0 \). After this the proof follows by changing the coordinates one by one. Moreover, it suffices to consider only small variations of \( \tau \), because locally increasing functions are globally increasing. Further, we can assume that the exponent \( h_r \) is simple. Indeed, if it is multiple, then the inequality \( F(p_{h'}) \geq F(p_h) \) follows by continuity, and the strict inequality \( F(p_{h'}) > F(p_h) \) follows by replacing the shift \( h_r \to h_r' \) by two consecutive shifts \( h_r \to h_r'' \to h_r' \) with simple exponent \( h_r'' \). The first shift does not reduce \( F \), and the second one increases it. Thus, the exponent \( h_r \) is simple. Assume that so are all the other \( h_k \) (in the general case the proof is similar).

The maximum of \( F \) is attained at some polynomial \( p(t) = \sum_{k=1}^n p_k e^{-h_k t} \) that has \( n - 1 \) roots on \( \mathbb{R}_+ \). Suppose all of them are simple: \( 0 < \mu_1 < \ldots < \mu_{n-1} < +\infty \). The general case is treated in
the same way. Consider a vector \( \mathbf{y} \in \mathbb{R}^n \) and a number \( \delta > 0 \), which will be specified later. For arbitrary \( \lambda > 0 \), take the \((\lambda \mathbf{y}, \lambda \delta)\)-variation of \( p \):

\[
p_\lambda(t) = (p_r + \lambda y_r)e^{-(h_r + \lambda \delta)t} + \sum_{k \neq r}(p_k + \lambda y_k)e^{-h_k t}.
\]

Thus, \( p_\lambda \in \mathcal{P}_{h'} \) with \( h' = h + \lambda \delta e_r \). We aim to find \( \mathbf{y} \) and \( \delta \) such that \( p_\lambda \) has a small norm and a large value of \( F \) for sufficiently small \( \lambda > 0 \). Then \( e^{-(h_r + \lambda \delta)t} = e^{-h_t t}(1 - \lambda \delta t) + O(\lambda^2) \), and therefore

\[
p_\lambda(t) - p(t) = \lambda \left[ -y_r \delta t e^{-h_r t} + \sum_{k=1}^{n} y_k e^{-h_k t} \right] + O(\lambda^2), \quad \lambda \to 0. \tag{4.1}
\]

Since \( t e^{-h_r t} \to 0 \) as \( t \to +\infty \), the last term in (4.1) is \( O(\lambda^2) \) uniformly for all \( t \in [0, +\infty) \). Hence, the polynomial \( p_\lambda - p \) can be approximated with precision \( O(\lambda^2) \) by a polynomial consisting of \( n + 1 \) exponentials (the right-hand side of (4.1)). Now we choose the coefficients of those polynomials so that \( \|p_\lambda\| < 1 \) and \( F(p_\lambda) > F(p) \). Consider the polynomial \( q(t) = \sum_{k=1}^{n} q_k e^{-h_k t} \) consisting of \( n + 1 \) exponentials (\( h_r \) has multiplicity 2) that vanishes at \( n \) points \(-\alpha = \mu_0, \mu_1, \ldots, \mu_{n-1}\), where \( \alpha < 0 \) with \( q(0) > 0 \). Clearly, \( q(t)p(t) < 0 \) at all points \( t \in \mathbb{R}_+ \) different from \( \mu_k \), and so \( |q(t)| < 1 \); hence \( \|p + q\| < 1 \). Furthermore, the derivative \( q^{(\ell)}(t) \) changes its sign \( n \) times on the interval \([-\alpha, +\infty)\), while \( p^{(\ell)}(t) \) changes its sign \( n - 1 \) times and the signs of these polynomials are different as \( t \to +\infty \). Therefore, they have the same sign at the point \( t = -\alpha \). Thus, \( q^{(\ell)}(-\alpha) > 0 \). Choosing a small \( \alpha \), we can assume that \( q^{(\ell)}(0) > 0 \), and so \( F(q) > 0 \).

Now we choose the coefficients \( y_1, \ldots, y_n \) and \( \delta \) so that the polynomial in brackets on the right-hand side of (4.1) coincides with \( q \). In this case \( y_k = q_k \) for all \( k \) and \(-y_r \delta = q_{r,1} \). Thus, \( \delta = -q_{r,1}/q_r \), and this number is positive. Indeed, the polynomial \( q \) consists of \( n + 1 \) exponentials and has \( n \) zeros; hence, all its coefficients are nonzero and have alternating signs. Since \( q_{r,1} \) and \( q_r \) are two consecutive coefficients, we have \( q_{r,1}/q_r < 0 \). Substituting into (4.1), we obtain

\[
p_\lambda(t) = p(t) + \lambda q(t) + O(\lambda^2), \quad \lambda \to 0.
\]

On the other hand, \( \|p + q\| < 1 \), and consequently \( \|p + \lambda q\| < 1 - c\lambda \), \( \lambda \in (0, 1] \). Therefore, \( \|p_\lambda\| < 1 - c\lambda \) for all positive \( \lambda \) small enough. Moreover, since \( F(p + \lambda q) = F(p) + \lambda F(q) \) with \( F(q) > 0 \), it follows that \( F(p_\lambda) > F(p) + c_1 \lambda \) for all positive \( \lambda \) small enough, where \( c_1 > 0 \) is a constant. This contradicts the optimality of \( p \). \[ \square \]

5. COROLLARIES AND SPECIAL CASES

Theorem 1 states that the maximum value of the \( \ell \)th derivative of a real quasipolynomial on the unit ball (in a monotone norm) is strictly increasing in each component \( h_i \). This, in particular, gives the largest value of the \( \ell \)th derivative over all polynomials when the vector \( h \) runs through a rectangular parallelepiped in \( \mathbb{R}^n \).

**Corollary 2.** If the parameters \( h_i \) lie in a rectangular parallelepiped \( \{h \in \mathbb{R}_+^n \mid h_i \in (0, \alpha_i]\} \), then the maximum value of \( \Phi_\ell(h) \) on the unit ball of an arbitrary monotone norm is attained at a unique point, the vertex \( h = (\alpha_1, \ldots, \alpha_n) \).

Let us now apply the comparison theorem to some particular norms. We will say that a monotone norm on \( \mathbb{R}_+ \) is shift-monotone if \( \|f(\cdot + a)\|_{\mathbb{R}_+} \leq \|f(\cdot)\| \) for every positive \( a \). All the \( L_p \)-norms possess this property. For an arbitrary shift-monotone norm \( \|\cdot\| \), the maximum value of \( \|p^{(\ell)}\|_\infty \) under the constraint \( \|p\| \leq 1 \) is equal to \( \|p^{(\ell)}(0)\| \); i.e., it is attained at zero. To prove this, assume the contrary and then shift the optimal polynomial so that its maximum point is at zero. Thus, the following holds.
Corollary 3. Suppose that a monotone norm $\|\cdot\|$ is also shift-monotone. Then for every $\ell \in \mathbb{N}$ the maximum value of $\|p^{(\ell)}\|_\infty$ on the unit ball $p \in \mathcal{RP}$, $\|p\| \leq 1$, is an increasing function of $h$.

As is well known, the $L_\infty$-norm of a polynomial in an arbitrary Chebyshev system takes the largest value $p^{(\ell)}(0)$ on the ball $\|p\|_\infty \leq 1$ at a unique polynomial, which possesses an alternance and is called the Chebyshev polynomial for that system. In our case this is the Chebyshev polynomial in the system $E_h$ that has an alternance on $[0, +\infty)$ of $n$ points including the point 0. We denote this polynomial by $T_h$. For every $h \in \mathbb{R}^n_+$, the polynomial $T_h$ can be found numerically by the Remez algorithm [34], which makes it possible to compute $\Phi_\ell(h) = m_\ell(h)$. We are interested in the uniform sharp constant $m_{\ell,n}$.

6. MARKOV–BERNSTEIN INEQUALITIES FOR EXPONENTIAL POLYNOMIALS

For an arbitrary vector $h \in \mathbb{C}_{++}$ and for every $\ell \geq 0$, the value $\Phi_\ell(h)$ of problem (2.1) for the functional $F(p) = -p^{(\ell)}(0)$ is also the sharp constant in the Markov–Bernstein inequality for exponential polynomials in the system $E_h$:

$$\|p^{(\ell)}\|_\infty \leq \Phi_\ell(h)\|p\|.$$  \hspace{1cm} (6.1)

Let us note that every shift-monotone norm $\|\cdot\|$ on $\mathbb{R}_+$ has its own constant $\Phi_\ell(h)$; therefore, it would be more correct to denote it by $\Phi_\ell(h, \|\cdot\|)$. However, we will keep the short notation. If $h$ is a real vector in the unit cube $I_n = \{h \in \mathbb{R}^n \mid 0 < h_i \leq 1, i = 1, \ldots, n\}$, then the maximum value of the constant $\Phi_\ell(h)$, according to Corollary 2, is attained at $h = e = (1, \ldots, 1)$. Therefore, the maximum is always attained at a polynomial from $\mathcal{P}_e$, i.e., at a function of the form $p(t) = e^{-t}q(t)$, where $q(t)$ is an arbitrary algebraic polynomial of degree $n - 1$. In particular, for $\|\cdot\| = \|\cdot\|_\infty$, the extremal polynomial is $p(t) = e^{-t}R_{n-1}(t)$, where $R_{n-1}$ is a Chebyshev polynomial of degree $n - 1$ with the Laguerre weight $e^{-t}$, for which the function $e^{-t}R_{n-1}(t)$ possesses $n$ points of alternance on the positive half-line [11, 36]. Let us collect all these facts in the following theorem.

**Theorem 2.** For an arbitrary shift-invariant norm $\|\cdot\|$ on $\mathbb{R}_+$ and arbitrary $\ell \geq 1$ and $h \in I_n$, we have

$$\|p^{(\ell)}\|_\infty \leq \Phi_\ell(e)\|p\|, \quad p \in \mathcal{RP}_h.$$  \hspace{1cm} (6.2)

This inequality becomes an equality at a unique (up to multiplication by a constant) polynomial from $\mathcal{RP}_e$. For $\|\cdot\| = \|\cdot\|_\infty$, one has $\Phi_\ell(e) = m_{\ell,n}$, and a unique extremal polynomial is the exponential Chebyshev polynomial $T_e = e^{-t}R_{n-1}(t)$, where $R_{n-1}$ is an algebraic Chebyshev polynomial of degree $n - 1$ with the Laguerre weight $e^{-t}$. Moreover, $m_{\ell,n} = |T_e^{(\ell)}(0)|$.

Thus, $m_{\ell,n}$ is equal to the $\ell$th derivative of the polynomial $e^{-t}R_{n-1}(t)$ at zero. For each $n$, this polynomial can be explicitly found; hence, $m_{\ell,n}$ is efficiently computable. However, a natural question arises as to how fast it grows in $n$ and in $\ell$. This problem reduces to the estimation of the $\ell$th derivative of the polynomial $R_{n-1}$ at zero. The first upper bound was presented in 1964 by Szegő [36], who proved that $|R_{n-1}^{(\ell)}(0)| \leq Cn$. Then this result was successively improved in [11, 13, 23]. Asymptotically sharp estimates for all $\ell$ have been obtained by Sklyarov [35].

**Theorem A (V. P. Sklyarov, 2009).** For an arbitrary natural $n \geq 2$, let $R_{n-1}$ denote the algebraic Chebyshev polynomial of degree $n - 1$ with the Laguerre weight $e^{-t}$. Then for each $\ell \in \mathbb{N}$ we have

$$\frac{8^\ell (n-1)! \ell!}{(n-1-\ell)! (2\ell)!} \left(1 - \frac{\ell}{2(n-1)}\right) \leq |R_{n-1}^{(\ell)}(0)| \leq \frac{8^\ell (n-1)! \ell!}{(n-1-\ell)! (2\ell)!}.$$  \hspace{1cm} (6.3)

Thus, the ratio of the upper and lower bounds (6.3) tends to one as $n \to \infty$ when $\ell$ is fixed. To estimate $m_{\ell,n}$, one needs to evaluate the derivatives of the polynomial $T_e$. They can easily be
Table 1. The upper bound on \( m_{2,n} \) for \( n = 2, \ldots, 10 \)

| \( n \) | Upper bound for \( m_{2,n} \) | Bound (6.8) | \( n \) | Upper bound for \( m_{2,n} \) | Bound (6.8) |
|---|---|---|---|---|---|
| 2 | 8.182 | 9 | 7 | 198.420 | 209 |
| 3 | 25.157 | 27 \( \frac{3}{2} \) | 8 | 268.283 | 281 |
| 4 | 52.587 | 57 | 9 | 348.788 | 363 \( \frac{2}{3} \) |
| 5 | 90.585 | 97 | 10 | 439.938 | 457 |
| 6 | 139.191 | 147 \( \frac{3}{2} \) |

obtained with the derivatives of the polynomial \( R_{n-1} \):

\[
\left[ e^{-t} R_{n-1}(t) \right]^{(\ell)} = e^{-t} \sum_{j=0}^{\ell} (-1)^{\ell-j} \binom{k}{j} R_{n-1}^{(j)}(t).
\]

Taking into account that the signs of \( R_{n-1}^{(j)}(0) \) alternate in \( j \), we get

\[
m_{\ell,n} = 1 + \sum_{j=1}^{\ell} \binom{\ell}{j} |R_{n-1}^{(j)}(0)|. \tag{6.4}
\]

Now we invoke (6.3) and after elementary simplifications obtain

\[
\sum_{j=0}^{\ell} \left(1 - \frac{j}{2(n-1)}\right) 8^j (n-1) \binom{j}{2j} \leq m_{\ell,n} \leq \sum_{j=0}^{\ell} 8^j (n-1) \binom{j}{2j} \tag{6.5}
\]

(all the terms with \( j \geq n \) are zeros; \( \binom{n}{n} = 1 \) for each \( n \geq 0 \)). We have not succeeded in simplifying these expressions. Combining this with Theorem 2, we obtain the following.

**Theorem 3.** For all \( \ell, n \in \mathbb{N}, n \geq 2 \), and \( h \in \mathcal{I}_n \), we have

\[
\max_{p \in \mathcal{P}(h)} \|p^{(\ell)}\|_{\infty} \leq m_{\ell,n} \|p\|_{\infty}, \tag{6.6}
\]

where the equality is attained at the polynomial \( T_{e} = e^{-t} R_{n-1} \), and

\[
\sum_{j=0}^{\ell} \left(1 - \frac{j}{2(n-1)}\right) 8^j (n-1) \binom{j}{2j} \leq m_{\ell,n} \leq \sum_{j=0}^{\ell} 8^j (n-1) \binom{j}{2j} \tag{6.7}
\]

(all the terms with \( j \geq n \) are zeros).

**Corollary 4.** Under the assumptions of Theorem 2, if the interval \((0,1]\) for \( h_i \) is replaced by \((0,\alpha]\), then both sides of (6.7) are multiplied by \( \alpha^\ell \). This inequality is asymptotically sharp with the extremal polynomials \( e^{-t/\alpha} R_{n}(t/\alpha) \).

Thus, for all \( \ell \) and \( n \), we have an upper bound for the sharp constant \( m_{\ell,n} \) in the Markov–Bernstein inequality for real exponential polynomials. In the next section we deal with applications to dynamical systems, where we need only the case \( \ell = 2 \), for which the right inequality in (6.7) takes the form

\[
m_{2,n} \leq \frac{16n^2 - 24n + 11}{3}. \tag{6.8}
\]

This estimate is asymptotically sharp as \( n \to \infty \). For small \( n \), it can be computed precisely. Table 1 presents the values of \( m_{2,n} \) for \( n \leq 10 \). We see that the general estimate of Theorem 3 in the third column is fairly close to the sharp estimate (second column).
Remark 1. Since $M_{\ell,n} \geq m_{\ell,n}$, the upper bound for $m_{\ell,n}$ need not be an upper bound for $M_{\ell,n}$. In general, Theorem 3 is not valid for complex exponential polynomials. In particular, we do not know whether estimates (6.6) also give estimates for $M_{\ell,n}$. In Conjecture 1 we suppose that the answer is affirmative and $M_{\ell,n} = m_{\ell,n}$.

7. STABILITY OF LINEAR SWITCHING SYSTEMS

A linear switching system is a linear ODE $\dot{x}(t) = A(t)x(t)$ for a vector function $x : \mathbb{R}_+ \to \mathbb{R}^n$ with the initial condition $x(0) = x_0$ in which $A(t)$ is a matrix function taking values in a given compact set $A$, called a control set. A control function, or a switching law, is an arbitrary measurable function $A : \mathbb{R}_+ \to A$. Linear switching systems arise naturally in problems of robotics, electronic engineering, mechanics, planning, etc. [20]. One of the main issues is to find or estimate the fastest possible growth of trajectories of the system and analyze its stability. The Lyapunov exponent $\sigma(A)$ of the system is the infimum of the numbers $\alpha$ for which $\|x(t)\| \leq C e^{\alpha t} \|x_0\|$, $t \in [0, +\infty)$. We will identify the linear switching system with the corresponding matrix family (control set) $A$. The Lyapunov exponent does not change after replacing the control set by its convex hull. Therefore, without lost of generality we assume that $A$ is convex. Moreover, it can also be assumed that the system is irreducible, i.e., its matrices do not share common invariant nontrivial subspaces.

A system is said to be asymptotically stable (or just stable for short) if all its trajectories tend to zero. If $\sigma < 0$, then the system is obviously stable. The converse is less obvious: the stability implies that $\sigma < 0$ (see [27]). For one-element control sets $A = \{A\}$, the stability is equivalent to the fact that $A$ is a Hurwitz matrix, i.e., all its eigenvalues have negative real parts. If $A$ contains more than one matrix, the stability problem becomes much harder. It is well known [24–26, 30] that the stability is equivalent to the existence of a Lyapunov norm in $\mathbb{R}^n$, i.e., a norm for which there exists a $\delta > 0$ such that $\|x(t)\| \leq e^{-\delta t} \|x_0\|$, $t \in \mathbb{R}_+$, for all trajectories $x(t)$. N. E. Barabanov [1] showed that for an arbitrary convex control set $A$, there exists an invariant Lyapunov norm such that $\|x(t)\| \leq e^{\sigma t} \|x_0\|$ for every trajectory $x(t)$, and for every point $x_0$ there exists a trajectory $\bar{x}(t)$ such that $\bar{x}(0) = x_0$ and $\|\bar{x}(t)\| = e^{\sigma t} \|x_0\|$.

The problem of computing the Lyapunov exponent of the system is equivalent to solving the stability problem. Indeed, for an arbitrary $s$, the inequality $\sigma(A) < s$ is equivalent to $\sigma(A - sI) < 0$, i.e., is equivalent to the stability of the system $A - sI$, where $I$ is the identity matrix and $A - sI = \{A - sI \mid A \in A\}$. Consequently, if we are able to solve the stability problem for every matrix family, then we can compare the Lyapunov exponent with an arbitrary number, which allows us to compute the Lyapunov exponent merely by the bisection method. Unfortunately, the general stability problem is quite difficult; the existing methods either work in low dimensions (as a rule, at most 4 or 5) or give too rough estimates. For example, the method of common quadratic Lyapunov function (CQLF) gives only necessary stability conditions, which are far from being necessary [20, 21]. Other methods, for example, those of piecewise linear or piecewise quadratic Lyapunov functions [3, 4], the extremal polytope method [15], etc., are sharper but, as a rule, are implementable only in low dimensions. The discretization method considered below is well known and looks quite promising.

7.1. Discrete and continuous systems. The discretization method reduces the stability problem for a linear switching system to the stability of an appropriate discrete-time system: $x(k+1) = B(k)x(k)$, $B(k) \in \mathcal{B}$, $k \geq 0$, where $\mathcal{B}$ is a compact matrix set (see [2, 10, 31, 32]). The following key statement was established in [27].

Fact 1. If for some $\tau_0 > 0$ the discrete system generated by the family of matrices $\mathcal{B} = I + \tau_0 A$ is stable, then it is stable for all $\tau < \tau_0$ and, moreover, so is the continuous-time system $A$.

Since $e^{\tau A} = I + \tau A + O(\tau^2)$, the matrix $B = I + \tau A$ gives a linear approximation of $e^{\tau A}$. In fact, we approximate every trajectory of the system $A$ by the Euler method with step size $\tau$ and
analyze the growth of the corresponding piecewise linear trajectories. If the step size tends to zero, then, obviously, the stability of the piecewise linear approximation implies the stability of the initial system. Fact 1, however, states more (which is not so obvious): if the piecewise linear approximation is stable for some step size $\tau_0$, not necessarily small, then it is stable for every smaller step size and the system $\mathcal{A}$ is stable.

Thus, the Euler method with step $\tau$ not only approximates the trajectories of the system but also gives a sufficient condition for its stability. The main question is to what extent that condition is necessary. The main problem is formulated as follows.

**Problem 1.** Given a continuous-time system $\mathcal{A}$ and an arbitrary $\varepsilon > 0$, find $\tau_0 > 0$ such that the inequality $\sigma(\mathcal{A}) < -\varepsilon$ implies the stability of the discrete-time system $\mathcal{B} = I + \tau \mathcal{A}$ for all $\tau \leq \tau_0$.

In other words, how small should the discretization step $\tau$ be to guarantee that the stability of the obtained discrete system approximates the stability of the continuous-time system with precision $\varepsilon$? The stability of the discrete system $\mathcal{B}$ can be expressed in terms of the joint spectral radius:

$$\rho(\mathcal{B}) = \lim_{k \to \infty} \max_{B(j) \in \mathcal{B}, j=1,...,k} \|B(k)\cdots B(1)\|^{1/k}.$$ 

A discrete-time system $\mathcal{B}$ is stable precisely when $\rho(\mathcal{B}) < 1$ (see [24–26]). Thus, if for some $\tau_0$ we have $\rho(I + \tau_0 \mathcal{A}) < 1$, then $\rho(I + \tau \mathcal{A}) < 1$ for all $\tau < \tau_0$ and, moreover, $\sigma(\mathcal{A}) < 0$. The inequality $\rho(\mathcal{B}) < 1$ is equivalent to the existence of a norm in $\mathbb{R}^n$ in which $\|B\| < 1$, $B \in \mathcal{B}$, i.e., all operators $B \in \mathcal{B}$ are contractions in the corresponding norm. Considering the unit ball in that norm gives an equivalent condition: there exists a symmetrized convex body $G \in \mathbb{R}^n$ such that $B(G) \subset \text{int} G$ for all $B \in \mathcal{B}$. For a one-matrix family $\mathcal{B}$, the joint spectral radius becomes the ordinary spectral radius, i.e., the largest modulus of eigenvalues of the matrix. If $\mathcal{B}$ contains at least two matrices, then the computation of its joint spectral radius becomes an NP-hard problem [5]. Nevertheless, several efficient methods to compute $\rho(\mathcal{B})$ for generic matrix families have been presented recently. They work well in dimensions $n \leq 25$, and for nonnegative families the dimension can be increased to several thousands [16, 22, 28]. Hence, the stability problem for discrete-time systems can be efficiently solved by computing the joint spectral radius. The solution of Problem 1 makes it possible to extend this method to continuous-time systems. To this end, the discretization step $\tau$ should not be too small. Otherwise, all matrices of the family $I + \tau \mathcal{A}$ will be close to the identity matrix, in which case all known methods for computing the joint spectral radius work slowly [33].

**Definition 2.** Given a compact family of matrices $\mathcal{A}$ and an arbitrary $\varepsilon > 0$, let

$$S_\mathcal{A}(\varepsilon) = \begin{cases} +\infty, & \text{if } \sigma(\mathcal{A}) \geq -\varepsilon, \\ \sup\{\tau > 0 \mid \rho(I + \tau \mathcal{A}) < 1\}, & \text{if } \sigma(\mathcal{A}) < -\varepsilon. \end{cases}$$

(7.1)

We use the simplified notation $S_\mathcal{A}(\varepsilon) = S$. Thus, if $\sigma(\mathcal{A}) < -\varepsilon$, then $\rho(I + \tau \mathcal{A}) < 1$ for all $\tau < S$. The number $S$ has the following meaning: if we do not know the value of $\sigma(\mathcal{A})$ but have a lower bound for $S$, then we choose an arbitrary value $\tau < S$ and compute the joint spectral radius $\rho(I + \tau \mathcal{A})$. If it is greater than or equal to 1, then $\sigma \geq -\varepsilon$; otherwise $\sigma < 0$. We also set

$$S_r(\varepsilon) = \inf\left\{ S_\mathcal{A}(\varepsilon) \mid \max_{A \in \mathcal{A}} \rho(A) \leq r \right\}.$$ 

Thus, $S_r(\varepsilon)$ is a lower bound for $S_\mathcal{A}(\varepsilon)$ which is uniform over all families $\mathcal{A}$ of $n$ matrices with the spectral radii not exceeding $r$. In Theorems 4 and 5 we express $S_\mathcal{A}(\varepsilon)$ and $S_r(\varepsilon)$ in terms of $M_2(\mathbf{h})$ and $M_{2,n}$ for $\mathbf{h} = -\text{sp}(A)$, $A \in \mathcal{A}$.

Now our plan is to estimate the quantity $S_\mathcal{A}(\varepsilon)$ from below by applying the Lyapunov norm of the family $\mathcal{A}$ (Propositions 1 and 2). After this the problem will be reduced to the solution of the extremal problem (7.2), which is equivalent to finding the sharp constant in the Markov–Bernstein type inequality for exponential polynomials (Proposition 3). Then deriving a uniform bound $S_r(\varepsilon)$
will be equivalent to minimizing the value $S_{\mathcal{A}}(\varepsilon)$ over all families of matrices $\mathcal{A}$ with spectral radii at most $r$. This will be done in Theorem 5.

7.2. Reduction to an extremal problem. We are going to estimate the discretization step $S_{\mathcal{A}}(\varepsilon)$ from below by solving a special extremal problem. Let \( h = (h_1, \ldots, h_n) \in \mathbb{C}^n \) and $\mathcal{RP}_h$ be the corresponding space of real quasipolynomials. Given an $\varepsilon > 0$, we denote by $s(h, \varepsilon)$ the value of the following problem:

$$
\frac{1 - p(0)}{p'(0) - \varepsilon p(0)} \rightarrow \min, \quad p \in \mathcal{RP}_h, \quad \|p\|_{\mathbb{R}^+} \leq 1, \quad p'(0) > \varepsilon p(0) .
$$

(7.2)

The geometric meaning of this quantity becomes clear from Proposition 1. To formulate it, we need to introduce some further notation. For a Hurwitz matrix $B$ and a vector $x \in \mathbb{R}^n$, we set $G_B(x) = \text{co}\{\pm e^{tB}x \mid t \in [0, +\infty)\}$. Thus, $G_B(x)$ is the convex hull of the curve $\{e^{tB}x \mid t \in [0, +\infty)\} \cup \{-e^{tB}x \mid t \in [0, +\infty)\}$, which connects the points $x$ and $-x$.

**Proposition 1.** Let $B$ be a Hurwitz matrix, $x \in \mathbb{R}^n \setminus \{0\}$ an arbitrary point, and $\varepsilon > 0$. Then the greatest number $\tau$ for which $x + \tau (B - \varepsilon I)x \in G_B(x)$ is equal to $s(h, \varepsilon)$, where $h = -\text{sp}(B)$ (Fig. 1).

**Proof.** Set $G_B(x) = G$ for short. Assume that $B$ has $n$ different eigenvalues; the general case then follows by the limit passage. By the Carathéodory theorem, an arbitrary point of $G$ is a convex combination of at most $n + 1$ extreme points of $G$, i.e., points of the form $\pm e^{tB}x$, $t \geq 0$. Hence, there are $n + 1$ nonnegative numbers $\{t_k\}_{k=1}^{n+1}$ and $n + 1$ numbers $\{q_k\}_{k=1}^{n+1}$ such that

$$
x + \tau (B - \varepsilon I)x = \sum_{k=1}^{n+1} q_k e^{t_k B} x, \quad \sum_{k=1}^{n+1} |q_k| = 1.
$$

(7.3)

In the basis of eigenvectors of $B$, we obtain the complex matrix $\tilde{B} = \text{diag}(-h_1, \ldots, -h_n)$ and a point $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_n) \in \mathbb{C}^n$. Moreover, $\text{Re} h_k > 0$, $k = 1, \ldots, n$. We assume that all coordinates $\tilde{x}_k$ are nonzero; the general case will then follow by the limit passage. Equality (7.3) has the following form in the new basis:

$$
(I - \tau \text{diag}(h_1 + \varepsilon, \ldots, h_n + \varepsilon)) \tilde{x} = \sum_{k=1}^{n+1} q_k e^{-t_k \tilde{B}} \tilde{x}.
$$

For each component $\tilde{x}_j$, we obtain

$$(1 - \tau (h_j + \varepsilon))\tilde{x}_j = \sum_{k=1}^{n+1} q_k e^{-t_k h_j} \tilde{x}_j;$$

then we divide by $\tilde{x}_j$:

$$
1 - \tau (h_j + \varepsilon) = \sum_{k=1}^{n+1} q_k e^{-t_k h_j}, \quad j = 1, \ldots, n.
$$

(7.4)
Thus, for each $A$ without switches, i.e., for the stationary control. Thus, for an arbitrary matrix family $A$, and denote by $\hat{G}$ the set of all suitable $\{t_k\}_{k=1}^{n+1}$ and $\{q_k\}_{k=1}^{n+1}$. This is a convex body in $\mathbb{C}^n$ over the field of real numbers, i.e., a convex body in $\mathbb{R}^{2n}$. By the convex separation theorem, for an arbitrary $\epsilon > 0$, the point $e + \delta(\hat{B} - \epsilon I)e$ does not belong to $\hat{G}$ precisely when it can be strictly separated from it with a linear functional, i.e., when there exists a nonzero vector $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$ for which

$$\Re\langle e + \delta(\hat{B} - \epsilon I)e, z \rangle > \max_{t \geq 0} \Re\left(\sum_{j=1}^{n} z_j e^{-\epsilon t_j}\right). \quad (7.5)$$

Denote by $q(t) = \sum_{j=1}^{n} z_j e^{-\epsilon t_j} \in \mathcal{P}_h$ a complex polynomial and by $p(t) = \Re q(t)$ the corresponding real quasipolynomial. Then the right-hand side of (7.5) is equal to $\|p\|_{\mathbb{R}_+}$. Then we modify the left-hand side as follows:

$$(e + \delta(\hat{B} - \epsilon I)e, z) = \sum_{j=1}^{n} z_j + \delta \sum_{j=1}^{n} (-h_j - \epsilon)z_j = q(0) + \delta q'(0) - \delta \epsilon q(0).$$

Thus, the left-hand side of (7.5) is equal to $p(0) + \delta p'(0) - \delta \epsilon p(0)$, and we arrive at the inequality

$$p(0) + \delta p'(0) - \delta \epsilon p(0) > \|p\|_{\mathbb{R}_+};$$

hence,

$$\delta(p'(0) - \epsilon p(0)) > \|p\|_{\mathbb{R}_+} - p(0).$$

The right-hand side is obviously nonnegative; consequently, $p'(0) - \epsilon p(0) > 0$. Normalizing $p$ so that $\|p\|_{\mathbb{R}_+} = 1$, we obtain $\delta > (1 - p(0))/(p'(0) - \epsilon p(0))$. Thus, the point $e + \delta(\hat{B} - \epsilon I)e$ does not belong to the convex body $G_B(e)$ if and only if there exists a quasipolynomial $p \in \mathcal{RP}_h$, $\|p\|_{\mathbb{R}_+} \leq 1$, satisfying that inequality. $\square$

**Remark 2.** In fact, we have proved slightly more: if $\tau < s(h, \epsilon)$ and $x \neq 0$, then the point $x + \tau(B - \epsilon I)x$ lies inside $G_B(x)$.

A lower bound for $S_A(\epsilon)$ is provided by the maximum possible value of $s(h, \epsilon)$ over all vectors $h = -\text{sp}(A) - \epsilon e$, $A \in \mathcal{A}$. We prove this in the following proposition by applying the invariant norm of the matrix family $\mathcal{A}$.

**Proposition 2.** For an arbitrary matrix family $A$, we have

$$S_A(\epsilon) \geq \min_{A \in \mathcal{A}} s(h, \epsilon), \quad (7.6)$$

where $h = -\text{sp}(A) - \epsilon e$.

**Proof.** We need to show that if $\sigma(A) < -\epsilon$, then $\rho(I + \tau A) < 1$ for all $\tau$ which are smaller than the right-hand side of (7.6). Since $\sigma(A) < -\epsilon$, it follows that there exists a Lyapunov norm in $\mathbb{R}^n$ for which $\|x(t)\| < e^{-\epsilon t} \|x_0\|, t > 0$, for every trajectory $x(\cdot)$. This is true, in particular, for trajectories without switches, i.e., for the stationary control. Thus, for an arbitrary matrix $A \in \mathcal{A}$, we have $\|e^{tA}x_0\| < e^{-\epsilon t} \|x_0\|$, and therefore $\|e^{(A + \epsilon I)t}x_0\| < \|x_0\|$ for all $t > 0$. Hence, for every point $y$ from the symmetrized convex hull of the set $\{e^{(A + \epsilon I)t}x_0 \mid t \in \mathbb{R}_+\}$, we have $\|y\| \leq \|x_0\|$. Applying Proposition 1 to the matrix $B = A + \epsilon I$ and taking into account Remark 2, we conclude that the point $(I + \tau A)x_0$ belongs to the interior of the set $G_{A+\epsilon I}(x_0)$, and therefore, $\|(I + \tau A)x_0\| < \|x_0\|$. Thus, for each $A \in \mathcal{A}$, the operator norm of $I + \tau A$ is strictly smaller than 1. Consequently, $\rho(I + \tau A) < 1$. $\square$
**Remark 3.** The value \( s(h, \varepsilon) \) of problem (7.2) is inversely proportional to its parameters: for every \( \lambda > 0 \), we have
\[
s(\lambda h, \lambda \varepsilon) = \lambda^{-1} s(h, \varepsilon). \tag{7.7}
\]
To show this, it suffices to change the variable \( u = \lambda t \) and observe that \( p'(0) = \lambda p_u'(0) \). Therefore, the computation or estimation of the value \( s(h, \varepsilon) \) is reduced to the case \( |h_n| \leq 1 \).

It remains to compute \( s(h, \varepsilon) \), i.e., to solve problem (7.2). This can be done numerically by applying convex optimization tools, since the set of real quasipolynomials \( q \) satisfying all the constraints of the problem is convex and the objective function is quasiconvex (all its level sets are convex). Therefore, for a particular vector \( h \in \mathbb{C}^n \), the problem can be efficiently solved, for instance, by the gradient relaxation methods. We, however, need a general lower bound for the value \( s(h, \varepsilon) \).

**Proposition 3.** For any \( \varepsilon > 0 \) and \( h \) with \( |h_n| \leq 1 \), we have
\[
s(h, \varepsilon) > \frac{2\varepsilon}{M_2(h) + 2\varepsilon^2},
\]
where \( s(h, \varepsilon) \) is the value of problem (7.2).

**Proof.** Let us first forget for a while that \( p \) is a quasipolynomial, keeping only the assumption that \( \|p''\| \leq M_2(h)\|p\| \). For the sake of simplicity we also let \( M_2(h) = m \). Thus, we need to compute the minimum of \((1 - p(0))/(p'(0) - \varepsilon p(0))\) under the assumptions \( p \in C^2(\mathbb{R}_+) \), \( \|p\| \leq 1 \), and \( \|p''\| < m \). Set \( x = 1 - p(0) \); then the objective function in problem (7.2) takes the form \( f(x) = x/(p'(0) - \varepsilon(1 - x)) \). The numerator and the denominator are both positive; hence the minimum of this fraction is attained when \( p'(0) \) is maximal. Denote \( p'(0) \) by \( \alpha \). Since \( p''(t) \geq -m \), we have
\[
p(t) \geq p(0) + p'(0)t - \frac{ml^2}{2} = 1 - x + \alpha t - \frac{ml^2}{2}
\]
for every \( t \in \mathbb{R}_+ \). On the other hand, \( p(t) \leq 1 \), because \( \|p\|_{\mathbb{R}_+} \leq 1 \). Therefore, the maximum of the quadratic function \( 1 - x + \alpha t - ml^2/2 \) over \( t \in \mathbb{R}_+ \) does not exceed 1. Computing the value at the vertex of the parabola, we obtain \( \alpha^2/(2m) + 1 - x \leq 1 \), and so \( \alpha \leq \sqrt{2mx} \). Hence,
\[
f(x) \geq \frac{x}{\sqrt{2mx} - \varepsilon(1 - x)}.
\]
The minimum of the right-hand side of this inequality under the positivity constraints for the numerator and denominator is attained at the point \( x_{\text{min}} = 2\varepsilon^2/m \); therefore,
\[
f(x) \geq f(x_{\text{min}}) = \frac{2\varepsilon}{m + 2\varepsilon^2}.
\]
Thus, the minimum of the objective function \((1 - p(0))/(p'(0) - \varepsilon p(0))\) is equal to \( 2\varepsilon/(m + 2\varepsilon^2) \) and is attained at the polynomial \( p \) defined above. Since the quadratic function does not belong to \( \mathcal{RP}_h \) and the unit ball in the space \( \mathcal{RP}_h \) is compact, it follows that the minimum of the objective function is greater than \( 2\varepsilon/(m + 2\varepsilon^2) \). \( \square \)

Combining Propositions 2 and 3, we obtain an estimate for the length of the discretization interval \( S_A(\varepsilon) \).

**Theorem 4.** For any numbers \( \varepsilon > 0 \) and \( r > \varepsilon \) and every system of operators \( A \) such that \( \max_{A \in A} \rho(A) \leq r \), we have
\[
S_A(\varepsilon) \geq \max_{A \in A} \frac{2\varepsilon}{M_2(h) + 2\varepsilon^2}, \tag{7.8}
\]
where \( h = -\text{sp}(A) - \varepsilon e \).
Estimate (7.8) is computed individually for every system of matrices $A$. A uniform bound $S_r(\varepsilon)$ over all families of matrices is given by the following theorem.

**Theorem 5.** For any numbers $\varepsilon > 0$ and $r > \varepsilon$ we have

$$S_r(\varepsilon) \geq \frac{2\varepsilon}{r^2 M_{2,n}}.$$  

**Proof.** Applying Propositions 1 and 2, we conclude that the value $S_A(\varepsilon)$ is not less than the minimum value of $s(h, \varepsilon)$ over all vectors $h = -\text{sp}(A) - \varepsilon e$, $A \in A$. Proposition 3 yields

$$S_A(\varepsilon) \geq \min_{h = -\text{sp}(A) - \varepsilon e, A \in A} \frac{2\varepsilon}{M_2(h) + 2\varepsilon^2}.$$  

On the other hand,

$$s(h, \varepsilon) \geq \frac{2\varepsilon}{(\rho(A) - \varepsilon)M_{2,n} + 2\varepsilon^2/(\rho(A) - \varepsilon)}.$$  

If $\rho(A) \leq r$ for all $A \in A$, then

$$S_r(\varepsilon) \geq \frac{2\varepsilon}{(r - \varepsilon)^2 M_{2,n} + 2\varepsilon^2}.$$  

Since $M_{2,n} > 2$ and $r \geq \varepsilon$, we see that the denominator of the fraction is at least $r^2 M_{2,n}$, which completes the proof.  

Theorems 4 and 5 show that the lower bound for the discretization step is linear in $\varepsilon$, which is much better than one might expect. As a rule, such estimates are exponential in $\varepsilon$ and can hardly be applied in dimensions higher than 3. The estimate provided by Theorem 5 depends on the dimension only in the coefficient of the linear function. This makes it applicable even in relatively high dimensions. However, there is one difficulty on this way: we cannot compute $M_{2,n}$. Only for families $A$ that consist of matrices with real eigenvalues the constant $M_{2,n}$ in Theorem 5 becomes efficiently computable because it can be replaced by $m_{2,n}$ (Theorem 2).

**Corollary 5.** If all matrices of the family $A$ have real spectra, then the value $M_{2,n}$ in Theorem 5 is replaced by $m_{2,n}$ and the value $M_2(h)$ in Theorem 4 is replaced by $m_2(h)$.

The value $m_2(h)$ can be found with arbitrary precision by applying the Remez algorithm. It is attained at the corresponding Chebyshev polynomial $T_h$. As follows from Theorem 2, the value $m_{2,n}$ is attained at the polynomial $e^{-1}R_{n-1}(t)$ and can be efficiently computed as well. It can be estimated by inequality (6.7) and, for small dimensions, can be evaluated by computing the Chebyshev polynomial $R_{n-1}$ with the Laguerre weight. For small dimensions, the results are presented in Table 1. We see that $m_{2,n}$ grows in $n$ not very fast; therefore, the estimate from Theorem 5 is applicable even for relatively high dimensions (few dozens).

Certainly, the case of real spectrum is rather exceptional. In the general case one needs to compute or at least to estimate the value of $M_{2,n}$ from above. We are not aware of any method to compute it with arbitrary precision. All known estimates are rough and can hardly be applicable. If Conjecture 1 is true, then $M_{2,n} = m_{2,n}$, and the problem will be solved. Otherwise, we have an open problem of finding satisfactory upper bounds for $M_{2,n}$.

**ACKNOWLEDGMENTS**

The author expresses his sincere thanks to the anonymous referee for the attentive reading and many valuable remarks.
FUNDING

This work is supported by the Theoretical Physics and Mathematics Advancement Foundation "BASIS."

REFERENCES

1. N. E. Barabanov, “Absolute characteristic exponent of a class of linear nonstationary systems of differential equations,” Sib. Math. J. 29 (4), 521–530 (1988) [transl. from Sib. Mat. Zh. 29 (4), 12–22 (1988)].

2. F. Blanchini, D. Casagrande, and S. Miani, “Modal and transition dwell time computation in switching systems: A set-theoretic approach,” Automatica 46 (9), 1477–1482 (2010).

3. F. Blanchini and S. Miani, “Piecewise-linear functions in robust control,” in Robust Control via Variable Structure and Lyapunov Techniques: Based on IEEE Workshop, Benevento, 1994 (Springer, Berlin, 1996), Lect. Notes Control Inf. Sci. 217, pp. 213–243.

4. F. Blanchini and S. Miani, “A new class of universal Lyapunov functions for the control of uncertain linear systems,” IEEE Trans. Autom. Control 44 (3), 641–647 (1999).

5. V. D. Blondel, S. Gaubert, and J. N. Tsitsiklis, “Approximating the spectral radius of sets of matrices in the max-algebra is NP-hard.” IEEE Trans. Autom. Control 45 (9), 1762–1765 (2000).

6. P. Borwein and T. Erdélyi, Polynomials and Polynomial Inequalities (Springer, New York, 1995).

7. P. Borwein and T. Erdélyi, “Upper bounds for the derivative of exponential sums,” Proc. Am. Math. Soc. 123 (5), 1481–1486 (1995).

8. P. Borwein and T. Erdélyi, “A sharp Bernstein-type inequality for exponential sums,” J. Reine Angew. Math. 476, 127–141 (1996).

9. P. Borwein and T. Erdélyi, “Newman’s inequality for Müntz polynomials on positive intervals,” J. Approx. Theory 85 (2), 132–139 (1996).

10. C. Briat and A. Seuret, “Affine characterizations of minimal and mode-dependent dwell-times for uncertain linear switched systems,” IEEE Trans. Autom. Control 58 (5), 1304–1310 (2013).

11. H. Carley, X. Li, and R. N. Mohapatra, “A sharp inequality of Markov type for polynomials associated with Laguerre weight,” J. Approx. Theory 113 (2), 221–228 (2001).

12. V. K. Dzyadyk and I. A. Shevchuk, Theory of Uniform Approximation of Functions by Polynomials (de Gruyter, Berlin, 2008).

13. G. Freud, “On two polynomial inequalities. I,” Acta Math. Acad. Sci. Hung. 22 (1–2), 109–116 (1971).

14. G. Gripenberg, “Computing the joint spectral radius,” Linear Algebra Appl. 234, 43–60 (1996).

15. N. Guglielmi, L. Laglia, and V. Protasov, “Polytope Lyapunov functions for stable and for stabilizable LSS,” Found. Comput. Math. 17 (2), 567–623 (2017).

16. N. Guglielmi and V. Protasov, “Exact computation of joint spectral characteristics of linear operators,” Found. Comput. Math. 13 (1), 37–97 (2013).

17. S. Karlin, “Representation theorems for positive functions,” J. Math. Mech. 12 (4), 599–617 (1963).

18. S. Karlin and W. J. Studden, Tchebycheff Systems: With Applications in Analysis and Statistics (Interscience, New York, 1966), Pure Appl. Math. 15.

19. M. G. Krein and A. A. Nudel’man, The Markov Moment Problem and Extremal Problems (Am. Math. Soc., Providence, RI, 1977), Transl. Math. Monogr. 50.

20. D. Liberzon, Switching in Systems and Control (Birkhäuser, Boston, 2003), Syst. Control Found. Appl.

21. D. Liberzon and A. S. Morse, “Basic problems in stability and design of switched systems,” IEEE Control Syst. Mag. 19 (5), 59–70 (1999).

22. T. Meijstrik, “Algorithm 1011: Improved invariant polytope algorithm and applications,” ACM Trans. Math. Softw. 46 (3), 29 (2020).

23. L. Milev and N. Naidenov, “Exact Markov inequalities for the Hermite and Laguerre weights,” J. Approx. Theory 138 (1), 87–96 (2006).

24. A. P. Molchanov and E. S. Pyatnitskii, “Lyapunov functions that specify necessary and sufficient conditions of absolute stability of nonlinear nonstationary control systems. I,” Autom. Remote Control 47 (3), 344–354 (1986) [transl. from Avtom. Telemekh., No. 3, 63–73 (1986)].

25. A. P. Molchanov and E. S. Pyatnitskii, “Lyapunov functions that specify necessary and sufficient conditions of absolute stability of nonlinear nonstationary control systems. II,” Autom. Remote Control 47 (4), 443–451 (1986) [transl. from Avtom. Telemekh., No. 4, 5–15 (1986)].

26. A. P. Molchanov and E. S. Pyatnitskii, “Lyapunov functions that specify necessary and sufficient conditions of absolute stability of nonlinear nonstationary control systems. III,” Autom. Remote Control 47 (5), 629–630 (1986) [transl. from Avtom. Telemekh., No. 5, 38–49 (1986)].
27. A. P. Molchanov and E. S. Pyatnitskiy, “Criteria of asymptotic stability of differential and difference inclusions encountered in control theory,” Syst. Control Lett. 13 (1), 59–64 (1989).
28. C. Möller and U. Reif, “A tree-based approach to joint spectral radius determination,” Linear Algebra Appl. 463, 154–170 (2014).
29. D. J. Newman, “Derivative bounds for Müntz polynomials,” J. Approx. Theory 18, 360–362 (1976).
30. V. I. Opoitsev, Equilibrium and Stability in Models of Collective Behavior (Nauka, Moscow, 1977) [in Russian].
31. V. Yu. Protasov and R. M. Jungers, “Is switching systems stability harder for continuous time systems?,” in 52nd IEEE Conf. on Decision and Control, 2013 (IEEE, 2013), pp. 704–709.
32. V. Yu. Protasov and R. M. Jungers, “Analysing the stability of linear systems via exponential Chebyshev polynomials,” IEEE Trans. Autom. Control 61 (3), 795–798 (2016).
33. V. Yu. Protasov, R. M. Jungers, and V. D. Blondel, “Joint spectral characteristics of matrices: A conic programming approach,” SIAM J. Matrix Anal. Appl. 31 (4), 2146–2162 (2010).
34. E. Remes, “Sur le calcul effectif des polynomes d’approximation de Tchebichef,” C. R. Acad. Sci. 199, 337–340 (1934).
35. V. P. Sklyarov, “The sharp constant in Markov’s inequality for the Laguerre weight,” Sb. Math. 200 (6), 887–897 (2009) [transl. from Mat. Sb. 200 (6), 109–118 (2009)].
36. G. Szegő, “On some problems of approximations,” Publ. Math. Inst. Hung. Acad. Sci., Ser. A 9, 3–9 (1964).

This article was submitted by the author simultaneously in Russian and English.