Automatica 74 (2016) 297–307

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Large deviations for non-zero initial conditions in linear systems ☆

Boris T. Polyak [a,b], Georgi Smirnov [c]

[a] Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia
[b] ITMO University, St-Petersburg, Russia
[c] University of Minho, Braga, Portugal

ARTICLE INFO

Article history:
Received 8 April 2015
Received in revised form
8 June 2016
Accepted 5 July 2016
Available online 2 October 2016

Keywords:
Large deviation of trajectories
Peak effect
Stabilization
Lower bounds
Transient response
Linear control systems

ABSTRACT

Transient response of linear systems with non-zero initial conditions was at the center of attention for engineers and researchers at early stages of classical control theory. However this field was not intensively investigated later. For instance, the breakthrough result on unavoidable peaking effect for systems with strong damping factor was obtained by Izmailov in 1987, but it did not attract much attention. We try to continue this line of research and provide explicit worst-case lower bound. Then we exhibit large deviation effects for other pole locations and estimate lower bounds for them. The upper bounds for deviations of trajectories are much better studied; to obtain the smallest deviations by static linear feedback the techniques of linear matrix inequalities can be exploited. We demonstrate that for such closed-loop systems the upper and lower bounds have the same asymptotic behavior.

1. Introduction

Transient response of linear systems with non-zero initial conditions was at the center of attention for engineers and researchers at early stages of classical control theory. However this field was not intensively investigated later. We try to continue this line of research.

Consider a single-input continuous time linear control system

$$\dot{x} = Ax + bu, \quad u \in R, \ b \in R^n. \tag{1}$$

If the vectors $b, Ab, \ldots, A^{n-1}b$ are linearly independent, then the system is controllable and, by the pole assignment theorem, there exists a linear feedback defined by a vector K such that the equilibrium position $x = 0$ of the closed-loop system

$$\dot{x} = Ax + bK^T x = Fx \tag{2}$$

is asymptotically stable. Moreover, one can generate a linear system with any given set $\Lambda \in C$ of eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$. (We assume that if $\lambda \in \Lambda$ and $\mathrm{Im}\lambda \neq 0$, then $\mathrm{Re}\lambda - i\mathrm{Im}\lambda \in \Lambda$.) Therefore, by choosing an appropriate linear feedback it is possible to obtain a closed-loop system with an arbitrary given damping speed. However, the trajectories of the closed-loop system with fast damping significantly deviate from the equilibrium position during the initial phase of the stabilization for some non-zero initial conditions. This phenomenon is called the *peak* effect and the large deviation is referred to as an *overshoot*. Although this phenomenon always attracted attention of scientific community (see, e.g., Feldbaum, 1948), the estimate for the peak effect in general situation was proved in Izmailov (1987). More precisely, Izmailov showed that there exists a constant $\gamma = \gamma(A, b) > 0$, such that, if $\{\lambda_1, \ldots, \lambda_n\}$, are the eigenvalues of $F = A + bK^T$, then the condition $\mathrm{Re}\lambda_j \leq -\sigma < 0, j = \overline{1, n}$, implies

$$\sup_{0 \leq t \leq \frac{1}{\sigma}} \ \sup_{|x(0)|=1} |x(t)| \geq \gamma \sigma^{n-1} \tag{3}$$

for solutions of the closed-loop system (2). Here $|x|$ is an arbitrary norm of vector x (below we shall specify it). The proof given by Izmailov was significantly simplified in Bushenkov and Smirnov (1997). In Sussman and Kokotovic (1991) Izmailov's result was generalized to obtain estimates for peaking effect for outputs. From (3) we also conclude that even for moderate decay rate ($\sigma = 2 \sim 5$) the peak effect grows exponentially as a function of the system dimension n.

In what follows we always treat lower bounds in the sense of (3), i.e. as worst-case transient response with respect to initial

☆ The material in this paper was partially presented at the 19th World Congress of the International Federation of Automatic Control (IFAC 2014), August 24–29, 2014, Cape Town, South Africa. This paper was recommended for publication in revised form by Associate Editor Yoshio Ebihara under the direction of Editor Richard Middleton.

E-mail addresses: boris@ipu.ru (B.T. Polyak), smirnov@math.uminho.pt (G. Smirnov).

http://dx.doi.org/10.1016/j.automatica.2016.07.047
0005-1098/

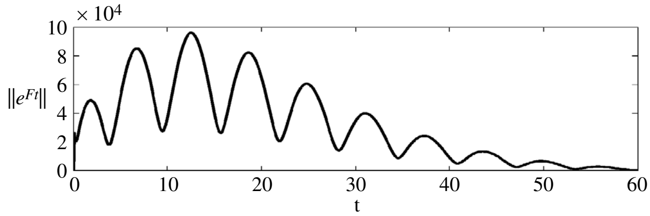

Fig. 1. Boeing 767 stabilization, $n = 55$.

conditions. Of course it does not imply large deviations for all initial conditions; for any stable system $\dot{x} = Fx$ there exist initial conditions such that $|x(t)| \le |x(0)|$ for all $t > 0$.

The effect of possible large deviations of trajectories is generally recognized beyond control community, for instance, in numerical analysis. The solution of the closed-loop system (2) is $x(t) = \exp(Ft)x(0)$, $F = A + bK^T$ thus $\max_{|x(0)|=1} |x(t)| = \|\exp(Ft)\|$, where $\|\cdot\|$ is matrix norm associated with vector norm $|\cdot|$. Hence deviations of trajectory are closely related to behavior of matrix exponent. In the famous paper Moler and Van Loan (1978) (and its continuation Moler & Van Loan, 2003) there are numerous examples of matrix exponents, having big *humps*. Fig. 1, borrowed from Moler and Van Loan (2003), exhibits transition process related to stabilization of Boeing 767 example ($\|e^{Ft}\|$ as function of t). Notice that the values of the humps are of order 10^5.

The problem of large deviations is highly significant, because non-zero initial conditions naturally arise in many applications. Typical example—observers, where true coordinates of the observed trajectory are never known. It is interesting to mention that Izmailov's result (as well as previous publications on peak effect Polotskij, 1981) are formulated for observers. Similar situation is met in switching systems (Liberzon, 2003), where the trajectory is always in non-zero position after switching. The large deviations of control systems trajectories from the equilibrium position during the transition process represent a serious obstacle to the design of cascade control systems (Sussman & Kokotovic, 1991) and to guidance stabilization Bushenkov and Smirnov (1997). Therefore, it is an important problem to estimate possible characteristics of the transition process and to describe the set of eigenvalues $\{\lambda_1, \dots, \lambda_n\}$ causing large deviations. Surprisingly this effect is not specific for large eigenvalues only. In Smirnov, Bushenkov, and Miranda (2009) it was shown that the large deviation of the solutions from the equilibrium position at the beginning of stabilization occurs for the set of eigenvalues $\{\lambda_1, \dots, \lambda_n\}$ with $|\lambda_k| \ll 1$ and $|\lambda_k| \gg 1$, $k = \overline{1, n}$.

It is necessary to note that the effect of significant growth of solution at transient times, similar to the peak effect, is also observed for two- and three-dimensional Poiseuille and Couette flows, even when all the eigenmodes decay exponentially (Reddy & Henningson, 1993). This phenomenon caused by the near-linear dependence of the eigenfunctions is of importance in the study of hydrodynamic stability and transition to turbulence and is discussed in the control theory framework in Bewley and Liu (1998). Transient effects play significant role in distributed control, where system dimension is large, see e.g. recent research on vehicular platoons (Martinec, Z, & Sebek, 2015).

In this work we analyze the situation with transient response in linear systems with non-zero initial conditions more deeply. First we treat the systems in companion form and provide examples with all eigenvalues equal real $-\sigma < 0$ where deviations of the trajectory for specific initial conditions can be estimated explicitly (Section 2). We show that the large deviation effect is present both for σ large or small. In Section 3 we focus on main results—lower bounds for worst deviations for systems in canonical form. First we calculate constant γ in (3). Next, the cases of other locations of eigenvalues are examined, for instance if some of the eigenvalues have very big or very small moduli. Extension of the results for systems in general (not canonical) form is given in Section 4. In Section 5 we address upper bounds for the peak effect. These problems are much better investigated, see Balandin and Kogan (2009), Bulgakov (1980), Hinrichsen, Plischke, and Wurth (2002), Whidborne and Amar (2011) and Whidborne and McKernan (2007). We provide a version of upper bounds, based on construction of invariant ellipsoids for the closed-loop system using semidefinite programming (SDP) approach (Boyd, El Ghaoui, Ferron, & Balakrishnan, 1994), and compare numerically various bounds for systems in companion form with desired damping. Theoretical comparison of the bounds for arbitrary linear systems is also discussed. Finally, we consider some open problems and directions for future research.

The conference version of the paper has been submitted at 19th IFAC World Congress (Polyak & Smirnov, 2014). Recently the authors with several additional coauthors have published the paper Polyak, Tremba, Khlebnikov, Shcherbakov, and Smirnov (2015) also related to large deviations problem. However there are serious differences if compared with the present work. (a) The detailed proofs of main results (Theorems 1 and 2) are obtained here for the first time, while in Polyak et al. (2015) related propositions are given with no proofs. (b) The same is true for constants in lower bounds, which are estimated below explicitly. (c) Some important results on mixed eigenvalue locations (Theorems 3 and 4) have not been addressed in Polyak et al. (2015). (d) Same is true for basic statement on asymptotic comparison of the bounds (Theorem 7). On the other hand, the paper Polyak et al. (2015) is focused on connections of the problem with Feldbaum's result (Feldbaum, 1948) and on large deviations in the case of zero initial conditions and non-zero inputs.

Notations. By default $|x|$ for vector $x \in R^n$ with components x_k, $k = \overline{1, n}$ stands for $|x|_\infty = \max_k |x_k|$ norm. When needed, the Euclidean norm is denoted by $|x|_2$. The same notation $|\lambda|$ for a complex number λ is used for its modulus. The entries of a matrix M are denoted $m_{i,k}$ and its transpose is M^T. The kth column of the identity matrix I is e_k. The spectrum of a closed-loop matrix F is the set $\Lambda \in C$ of eigenvalues $\{\lambda_1, \dots, \lambda_n\}$, and we use the following characteristics of eigenvalues:

$$\sigma = -\max_j \mathrm{Re}\lambda_j; \qquad \omega = \max_j |\lambda_j|; \qquad \rho = \min_j |\lambda_j|.$$

That is σ is stability degree of F, ω is its spectral radius and ρ is inverse of spectral radius of F^{-1}.

2. Motivating examples

First we provide some examples on the lower bounds for deviations for linear systems in input–output form. The characteristic polynomial has all roots equal $-\sigma < 0$. We observe that for some specific initial conditions large deviations arise both for σ small and σ large.

Example 1. Linear differential equation

$$(s + \sigma)^n y(t) = 0, \quad s = \frac{d}{dt} \tag{4}$$

where $\sigma > 0$ and $n \ge 2$, with the initial conditions $y(0) = 1$, $y^{(k)}(0) = 0$, $k = \overline{1, n-1}$.

The respective solution is given by the formula

$$y(t) = \left(\sum_{j=0}^{n-1} (\sigma t)^j / j! \right) e^{-\sigma t}. \tag{5}$$

Since $y'(t) = -\frac{-\sigma^n t^{n-1}}{(n-1)!} e^{-\sigma t}$, we get $y^{(n-1)}(t) = -\frac{d^{n-2}}{dt^{n-2}} \frac{\sigma^n t^{n-1}}{(n-1)!} e^{-\sigma t} = -\frac{\sigma^n t e^{-\sigma t}}{n-1} L_{n-2}^{(1)}(\sigma t)$, where $L_k^{(\alpha)}(\theta) = \frac{\theta^{-\alpha} e^{\theta}}{k!} \frac{d^k}{d\theta^k}\left(\theta^{n+\alpha} e^{-\theta}\right)$ is a generalized Laguerre polynomial. The function $\left|\theta e^{-\theta} L_{n-2}^{(1)}(\theta)\right|, \theta \geq 0$, achieves its maximum at a point θ_n. Thus for $\gamma_n = \frac{\left|\theta_n e^{-\theta_n} L_{n-2}^{(1)}(\theta_n)\right|}{n-1}$ we have

$$\left|y^{(n-1)}\left(\frac{\theta_n}{\sigma}\right)\right| = \gamma_n \sigma^{n-1}. \tag{6}$$

From this we see that for large values of σ the absolute value of $(n-1)$th derivative of the solution $y(\cdot)$ at $t = \theta_n/\sigma$ is $O\left(\sigma^{n-1}\right)$.

The numerical values of γ_n, θ_n for small n can be found in Polyak et al. (2015).

Example 2. Eq. (4) with the initial conditions $y^{(k)}(0) = 1$, $k = \overline{0, n-1}$.

The solution is given by

$$y(t) = \left(\sum_{j=0}^{n-1} \frac{((1+\sigma)t)^j}{j!}\right) e^{-\sigma t}. \tag{7}$$

Taking the last term in the sum we get

$$y\left(\frac{n-1}{\sigma}\right) \geq \left(\frac{n-1}{\sigma e}\right)^{n-1} \frac{1}{(n-1)!} \geq \frac{\sigma^{-(n-1)}}{e\sqrt{n-1}}. \tag{8}$$

Thus for small values of $\sigma > 0$ we have $|y((n-1)/\sigma)| = O\left(\sigma^{-(n-1)}\right)$. Here large overshoot occurs for t large.

Denoting $x(t) = (y(t), y'(t), \ldots, y^{(n-1)}(t))^T$, we conclude that $\max_{t>0} \max_{|x(0)|=1} |x(t)|$ is greater than or equal to $\gamma_n \max\{\sigma^{n-1}, \sigma^{-(n-1)}\}$, thus for equal roots of the characteristic polynomial large deviations are unavoidable both for σ large and σ small. However the situation is different: for σ large we have peaking effect for the last derivative in the initial period of time, while for σ small the trajectory itself has large values and it happens for t large enough. To combine these effects we use the expression *large deviations* instead of the *peak*.

3. Lower bounds for systems in companion form

In this section we consider system (1) in companion form, i.e. with

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \qquad b = e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{9}$$

The closed-loop system with the feedback $u(x) = (K, x)$ has the system matrix $F = A + bK^T$

$$= \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix}, \quad a_{n-i+1} = -K_i \tag{10}$$

and its characteristic polynomial is

$$\Delta(s) = s^n + a_1 s^{n-1} + \cdots + a_n. \tag{11}$$

We can arbitrarily choose the coefficients a_i; our goal is to link the location of the matrix F spectrum $\Lambda = \{\lambda_1, \ldots, \lambda_n\}$ (roots of $\Delta(s)$) with the transient behavior of solutions $x(t)$. Of course only stable systems are of interest, $\mathrm{Re}\lambda_i \leq -\sigma < 0$.

3.1. Large eigenvalues

One of the first papers on dependence of transient response on spectrum location is due to Feldbaum (1948). He investigated the behavior of the first component of the solution $x_1(t)$; it decreases monotonically and exhibits no peaking. However it is not hard to see that the last component $x_n(t)$ behaves differently and peaking effect is strong for eigenvalues with large negative real parts (Polyak et al., 2015). The results in Feldbaum (1948) are true for all real eigenvalues (with one possible complex pair); below we get rid of this assumption. The proof of the theorem (as well as of all other results) can be found in the Appendix; its statement without the proof has been published in Polyak et al. (2015).

Theorem 1. *If* $\{\lambda_1, \ldots, \lambda_n\}$, *with* $\mathrm{Re}\lambda_j < 0$, $j = \overline{1, n}$, *are the eigenvalues of* F, $\rho = \min\{|\lambda_j| : j = \overline{1, n}\}$, $\omega = \max\{|\lambda_j| : j = \overline{1, n}\}$, *and* $x(0) = e_1$, *then*

$$\left|x\left(\frac{\log 2}{n\omega}\right)\right| \geq \frac{2\log 2 - 1}{n}\rho^{n-1} > \frac{0.3862}{n}\rho^{n-1}. \tag{12}$$

We conclude that for large eigenvalues and large n unavoidable deviations can exceed any practically acceptable values. For instance, $\sigma = 4, n = 10$ imply deviations of order 10^5. It is interesting to note that Theorem 1 covers wider class of spectrum locations if compared with Izmailov's theorem. It deals with $\min_j |\lambda_j|$, not with $\min_j |\mathrm{Re}\lambda_j|$. Moreover, it always gives a better estimate than that of Theorem 2 in Polyak et al. (2015), based on Feldbaum's result. Theorem 1 (as well Theorems 2–4) provides also particular initial conditions and time instances when peak effect is met; but it is just a lower bound and possibly larger deviations may be achieved.

3.2. Small eigenvalues

Now we address the case with $\omega = \max\{|\lambda_j| : j = \overline{1, n}\}$ small, where $\{\lambda_1, \ldots, \lambda_n\}$, are the eigenvalues of F.

Theorem 2. *If* $\mathrm{Re}\lambda_j < 0$, $j = \overline{1, n}$ *and* $x(0) = e_n$ *then*

$$\left|x\left(\frac{\theta_n}{\omega}\right)\right| \geq \gamma_n \frac{1}{\omega^{n-1}}, \tag{13}$$

where $\theta_n = \frac{n-1}{n}\log 2$ *and* $\gamma_n = \frac{(\log 2)^{n-1}}{en!}$.

We conclude that small eigenvalues (with $\omega < 0.2/n$, for example) cause large deviations for t large.

3.3. Mixed eigenvalues

We consider the case when both small and large eigenvalues are present. Theorem 3 exhibits that dominating (see (15)) large eigenvalues cause peak effect in spite of existence of small eigenvalues, while Theorem 4 demonstrates large deviations when small eigenvalues dominate (18). Notice that the corresponding initial conditions are different.

Assume that the eigenvalues $\lambda_1, \ldots, \lambda_n$ satisfy the inequalities

$$\omega > |\lambda_1| \geq \cdots |\lambda_\nu| \geq \eta > \xi > |\lambda_{\nu+1}| \geq \cdots \geq |\lambda_n| \tag{14}$$

and $\mathrm{Re}\lambda_j < 0, j = 1, \ldots, n$.

Theorem 3. *Let* $\nu \geq 2$ *and*

$$\eta > \left(\frac{10(1+\xi)^n}{\omega n}\right)^{\frac{1}{\nu}}. \tag{15}$$

Then for $x(0) = e_{n-\nu+1}$ *the following inequality holds:*

$$\left|x\left(\frac{\log 2}{\omega n}\right)\right| \geq \gamma_n \eta^{\nu-1}, \quad \gamma_n = \frac{1 - 1.1\log 2}{n}\log 2. \tag{16}$$

Assume now that the eigenvalues $\lambda_1, \ldots, \lambda_n$ satisfy more restrictive inequalities

$$|\mathrm{Re}\lambda_1| \geq \cdots |\mathrm{Re}\lambda_\nu| \geq \eta > \xi > |\lambda_{\nu+1}| \geq \cdots \geq |\lambda_n|, \tag{17}$$

$\omega > |\lambda_j|$ and $\mathrm{Re}\lambda_j < 0$, $j = 1, \ldots, n$.

Theorem 4. *Suppose that*

$$\begin{aligned} &\nu \leq \min\{n-3, (n-2)/2\}, \qquad \eta \geq 3, \\ &\quad \xi \leq \min\{c/\omega, 1\}, \end{aligned} \tag{18}$$

where $c = \left(\frac{3}{4e^2}\frac{(n-\nu-1)!}{1-e/(n-\nu)}\right)^{1/(\nu-1)}$, *and* $x(0) = e_n$. *Then the following inequality is satisfied*

$$\left|x\left(\frac{1}{\xi}\right)\right| \geq \frac{\gamma_n}{\xi^{n-2\nu-1}}. \tag{19}$$

Here γ_n can be estimated explicitly (see the proof). From this theorem we see that if small eigenvalues dominate (that is their number is large enough and their values are small (18)), then condition (17) implies the effect of large deviations for large t.

4. Lower bounds for systems in general form

Let $l \in R^n$ satisfy $(l, A^{n-1}b) = 1$ and $(l, A^{j-1}b) = 0$, $j = \overline{1, n-1}$ (such l exists due to the controllability assumption). In the coordinates $z_j = (l, A^{j-1}x)$, $j = \overline{1, n}$ system (1) is converted into companion form and for closed-loop system $\dot{z} = Gz$ the eigenvalues of matrices G and F coincide. We have $z = Mx$, where matrix M has the rows $l^T A^{j-1}$, $j = \overline{1, n}$. Since $|x|_2 \geq |z|_2/\|M\| \geq |z|_\infty/\kappa$, where κ is spectral norm of M, that is the maximal singular value of the matrix M, all the estimates, previously obtained for systems in canonical form, can be reformulated in terms of Euclidean norm in the original space after division by κ. Notice that matrix M depends only on the pair (A, b) and does not depend on the choice of the eigenvalues (i.e. on the feedback K), see Wonham (1979). For example, combining Theorem 1 with this result, we get

Theorem 5. *For any solution of* (2) *the constraint* $\mathrm{Re}\lambda_j \leq -\sigma < 0$, $j = \overline{1, n}$ *for eigenvalues of the closed-loop system implies the estimate*

$$\sup_{0 \leq t \leq \frac{1}{\sigma}} \sup_{|x(0)|_2 = 1} |x(t)|_2 \geq \gamma \sigma^{n-1} \tag{20}$$

with $\gamma = (2\log 2 - 1)/(n\kappa)$.

This is Izmailov's theorem with specified constant in lower bound. Of course this estimate can be more conservative than ones given by Theorem 1 for systems in companion form. It is of interest to obtain less conservative lower bounds for some particular classes of matrices, for instance, for matrices in Jordan form or for tridiagonal matrices.

Illustrative examples

To illustrate the previous theorems we address a series of examples of control systems in companion form. We choose initial conditions provided by the theorems, they are not necessarily "the worst-case" conditions. The results are presented as infinity-norm of $x(t)$, thus non-smooth plots are due to relative domination of different states at different time periods.

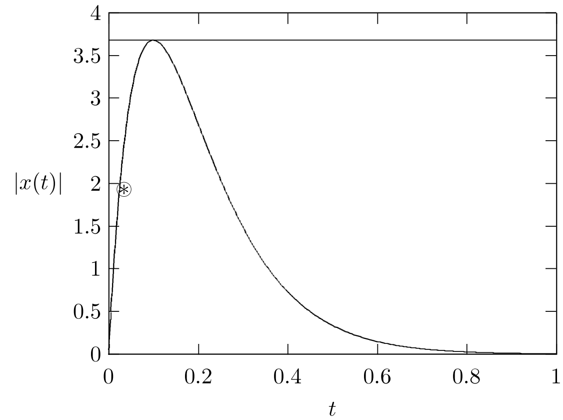

Fig. 2. Large eigenvalues, $n = 2$, $\Lambda = \{-10, -10\}$.

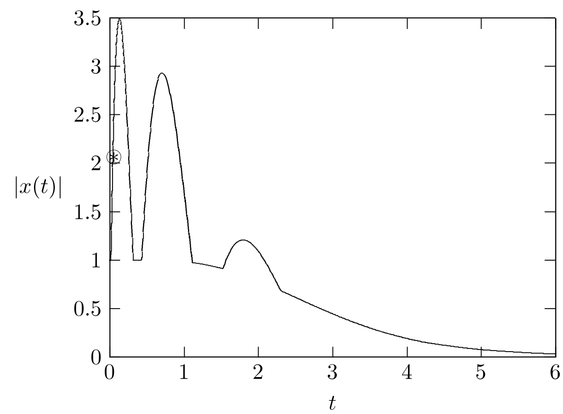

Fig. 3. $n = 6$, $\Lambda = \{-2, -2, -2, -2, -2, -2\}$.

First consider a two dimensional system. The eigenvalues of the closed loop system matrix are $\lambda_1 = \lambda_2 = -10$. The norm $|x(t)|$ of the solution starting at $(1, 0)$ is shown in Fig. 2. The true deviation is compared with lower bounds provided by (6) and (12). We see that lower bound (6), shown by horizontal line, is tight (it is specially oriented for equal eigenvalues, as in this example), while estimate provided by Theorem 1 (shown here and in the other examples with an asterisk in a circle) is conservative. However, it approximately predicts the time instant with peaking effect and its value.

When n increases, the peak effect is strongly exhibited for moderate stability degree. In Fig. 3, one can see that for $n = 6$ large deviations are met for $\sigma = 2$.

To illustrate Theorem 2 consider the closed loop system with $\Lambda = \{-0.1, -0.1\}$; norm of the solution starting at $(0, 1)$ is shown in Fig. 4.

The case of mixed eigenvalues is presented in Fig. 5. Here $n = 5$, $\Lambda = \{-1, -1000, -1000, -1000, -1000\}$ and Theorem 3 can be applied with $\nu = 4$, $\eta = 1000$, $\xi = 1$, $n - \nu + 1 = 2$. The norm of trajectory with $x(0) = e_2$ demonstrates peak effect for t extremely small.

5. Upper bounds

In contrast with lower bounds, upper bounds for transient process in linear systems with non-zero initial conditions are much better studied, see Balandin and Kogan (2009), Bulgakov (1980), Hinrichsen et al. (2002), Nechepurenko (2002), Van Dorsselaer,

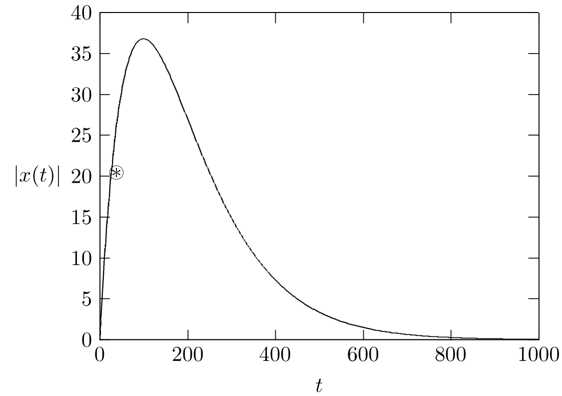

Fig. 4. Small eigenvalues, $n = 2$, $\Lambda = \{-0.1, -0.1\}$.

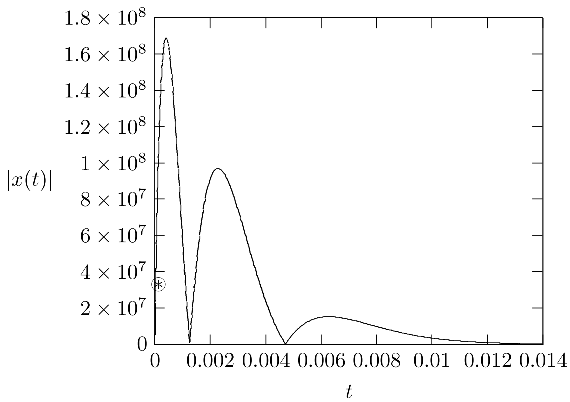

Fig. 5. Mixed eigenvalues, $n = 5$, $\Lambda = \{-1, -1000, -1000, -1000, -1000\}$.

Kraaijevanger, and Spijker (1993), Whidborne and Amar (2011) and Whidborne and McKernan (2007). Our main goal is to compare lower and upper bounds (Theorem 7).

If we have stable closed-loop linear system $\dot{x} = Fx$, then we can try to find the quadratic Lyapunov function $V(x) = (P^{-1}x, x)$ such that the invariant ellipsoid $E = \{x : V(x) \le 1\}$ contains the unit ball $\{x : |x|_2 \le 1\}$ and has minimal ratio of the semi-major and the semi-minor axes (that is condition number of P is minimal). This is equivalent to solving Semi-Definite Programming (SDP Boyd et al., 1994) problem

$$\|P\| \to \min, \qquad PF^T + FP \preceq 0, \quad I \preceq P. \tag{21}$$

Here $P = P^T \in R^{n\times n}$ is matrix variable, $M \preceq 0$ denotes semidefinite matrix M, and $\|P\|$ stands for spectral norm of P. Then we can guarantee the estimate

$$\max_{t\ge 0} \max_{|x(0)|_2 \le 1} |x(t)|_2 \le \|\hat{P}\|, \tag{22}$$

where $\hat{P}$ is the solution of the above SDP. This is a well known approach to get upper bounds for deviations in closed-loop systems.

For open-loop systems we can similarly design a linear feedback in order to guarantee the minimal possible deviations with desired damping. Consider SDP

$$\|P\| \to \min, \tag{23}$$

$$P(A + bK^T)^T + (A + bK^T)P \preceq -2\sigma P, \tag{24}$$

$$I \preceq P. \tag{25}$$

Lyapunov inequality (24) guarantees that the decay rate of the closed-loop system exceeds σ and that the ellipsoid $E = \{x \in R^n : (x, P^{-1}x) \le 1\}$ is its invariant set, while conditions (23) and (25) imply that E contains the unit ball $\{x : |x| \le 1\}$ and has minimal ratio of the semi-major and the semi-minor axes. Inequality (24) is nonlinear with respect to the unknown variables P and K. This difficulty can be easily overcome by introducing new variable $Y = PK$. In terms of variables P and Y (24) reads

$$AP + PA^T + bY^T + Yb^T \preceq -2\sigma P. \tag{26}$$

This is a typical SDP problem and it can be solved numerically (Grant & Boyd, n.d.). Its solution $K = \hat{P}^{-1}\hat{Y}$ can be a good candidate for a feedback with desired stability degree and small deviation for all non-zero initial conditions. Similar approaches can be found in the references mentioned above.

It is of interest to compare lower and upper bounds for systems in companion form with desired decay rate σ. The lower bounds are given by Theorem 1; having in mind that $\rho \ge \sigma$, estimate (12) implies

$$\max_{0\le t} \max_{|x(0)|=1} |x(t)| \ge \gamma_n \sigma^{n-1} = \underline{\kappa}_n, \quad \gamma_n = \frac{0.3863}{n}. \tag{27}$$

To obtain upper bounds we fix the desired σ and solve SDP (23), (25), (26) with its solution $\hat{P}$. Then we guarantee the estimate

$$\max_{0\le t} \max_{|x(0)|_2=1} |x(t)|_2 \le \|\hat{P}\| = \overline{\kappa}_n. \tag{28}$$

Minor incomparability of the bounds is that we used ∞-norm for lower bounds and 2-norm for upper bounds; however $|\cdot|_\infty \le |\cdot|_2$ and this leads to more conservative results.

Finally we calculate the true maximal deviation in 2-norm as

$$\kappa_n = \max_{t>0} \max_{|x(0)|_2=1} |x(t)|_2 = \max_t \|e^{Ft}\|_2 \tag{29}$$

where F is closed-loop system matrix with the controller $K = \hat{P}^{-1}\hat{Y}$.

In numerical calculations we fixed $\sigma = 2$, the results are collected in Table 1:

Table 1
Numerical values of κ_n.

n	2	3	4	5
$\underline{\kappa}_n$	0.3863	0.5151	0.7726	1.2362
κ_n	1.60	5.75	17.18	66.46
$\overline{\kappa}_n$	17.94	337.00	$6.8 \cdot 10^3$	$1.47 \cdot 10^5$

We see that both lower and upper bounds are highly conservative even for small dimensions n.

Another problem of interest is to investigate the closed-loop system obtained by solving upper bound (23), (25), (26) with fixed $\sigma = 2$. For system in canonical form we calculate the matrix $F = A + bK^T$, $K = \hat{P}^{-1}\hat{Y}$, its eigenvalues $\Lambda(F)$ and stability degree $\sigma(F) = -\max_k \mathrm{Re}\lambda_k$; they are shown in Table 2 for several n.

Table 2
Eigenvalues and stability degree for closed-loop systems.

n	2	3	4	5
$\sigma(F)$	2.31	2.39	2.17	2.26
$\Lambda(F)$	$-2.31 \pm 2.21i$	-18.27 $-2.40 \pm 1.26i$	$-2.96 \pm 5.13i$ $-2.17 \pm 0.91i$	$-4.38 \pm 5.79i$ $-2.39 \pm 1.55i$ -2.26

We conclude that true stability degree $\sigma(F)$ is close to desired decay rate $\sigma = 2$. Surprisingly the spectrum $\Lambda(F)$ exhibits no regularity; eigenvalues are both real and complex and some of them (e.g. for $n = 3$) have large modulus.

We have seen that lower and upper bounds differ strongly. However for large eigenvalues their asymptotic behavior (for σ large) is similar. To provide this result we reformulate Theorem 5 (relating to systems in general form and large σ) in terms of quadratic Lyapunov functions $V(x) = (P^{-1}x, x)$. We denote $\chi(P)$ the condition number of matrix P, that is $\chi(P) = \|P\|\|P^{-1}\| = M/m$, M, m being the largest and the smallest eigenvalues of P. Obviously large $\chi(P)$ correspond to ellipsoids $E = \{x \in R^n : (x, P^{-1}x) \leq 1\}$ with large ratio of its biggest and smallest half-axis.

Theorem 6. *For any feedback $u = K^T x$ such that the closed-loop matrix $F = A + bK^T$ has eigenvalues λ_i with $\mathrm{Re}\lambda_i \leq -\sigma < 0$ and for any $V(x) = (P^{-1}x, x)$ which is Lyapunov function for closed-loop system $\dot{x} = Fx$ the following estimate holds:*

$$\chi(P) \geq \underline{\gamma}\sigma^{2(n-1)}. \tag{30}$$

Here $\underline{\gamma} > 0$ depends on A, b but not on K.

Indeed for $|x(0)|_2 = 1$ due to Theorem 6 and the Lyapunov function properties, we have

$$\begin{aligned} 1/m &= (1/m)|x(0)|_2^2 \geq V(x(0)) \\ &\geq \max_{t\geq 0} V(x(t)) \geq (1/M) \max_{t>0} |x(t)|_2^2 \\ &\geq (1/M) \max_{t>0} |x(t)|^2 \geq (1/M)\gamma_n^2 \sigma^{2(n-1)}. \end{aligned} \tag{31}$$

Thus condition number of P can be used for comparison of upper and lower bounds for deviations of the trajectories in 2-norm. Now we are in a position to present the main result of this section.

Theorem 7. *For controllable system* (1) *and any $\sigma > 0$ there exist a feedback $u = K^T x$ and a quadratic form $V(x) = (P^{-1}x, x)$ such that the closed-loop matrix $F = A + bK^T$ has eigenvalues λ_i with $\mathrm{Re}\lambda_i \leq -\sigma$ and $V(x)$ is its Lyapunov function while*

$$\chi(P) \leq \overline{\gamma}(\sigma + 3/2)^{2(n-1)}, \tag{32}$$

where $\overline{\gamma}$ depends on A, b only.

Of course dependence of $\underline{\gamma}$ and $\overline{\gamma}$ on n (for instance for systems in companion form) can be different; moreover terms $\sigma^{2(n-1)}$ and $(\sigma + 3/2)^{2(n-1)}$ in lower and upper bounds (30), (32) are also different. However both bounds are of order $O(\sigma^{2(n-1)})$ for σ large.

6. Conclusions and future research

We provided some results on transient response in linear systems with non-zero initial conditions. First we investigated systems in companion form and obtained lower bounds for maximal deviations in such systems. The bounds depend on pole location of closed-loop systems; large deviations are unavoidable for various locations. The results are extended to systems in general form. Comparison confirms asymptotic equivalence of lower and upper bounds.

However these are just the first steps in the challenging area. We can mention few open problems relating to transient response research:

(1) The results on lower bounds for systems in companion form (Section 3) are formulated for specific eigenvalue locations. But what can be said for arbitrary eigenvalues? How to estimate $\alpha_n = \inf_\Lambda \sup_{t>0} \sup_{|x(0)|=1} |x(t)|$? Here $x(t)$ is the solution of equation with matrix (10) having spectrum Λ. Our initial conjecture was that $\Lambda = \{-\sigma, \ldots, -\sigma\}$ with real $\sigma > 0$ depending on n is the answer. However this conjecture is false; counterexamples can be constructed for small n. Probably the question has sense only if some restrictions on Λ are added, for example $|\lambda_i| \leq \omega,\ i = 1, \ldots, n$. To the best of our knowledge, the problem remains open.

(2) Similar question: what are the best constants in all theorems of Section 3?

(3) Lower bounds for systems in general form (Section 4) can be rather conservative. Are there some classes of systems (beyond canonical) which allow more accurate estimates?

(4) The case of discrete-time systems is challenging. It is well known that for Shur stable matrix A its powers A^k can increase at initial iterations. But what is an analog for lower bound results of Section 3?

(5) Lower bounds for deviations of transient response from steady-state outputs for systems with zero initial conditions and harmonic inputs deserve similar analysis. Some results in this direction are presented in Polyak et al. (2015).

Acknowledgments

The authors are grateful to P. Shcherbakov, M. Khlebnikov, A. Tremba, M. Kogan, A. Polinova, I. Caiado, J. Whidborne and anonymous referees for numerous helpful comments and references.

The work was supported by the Portuguese Foundation for Science and Technologies (FCT), the Portuguese Operational Programme for Competitiveness Factors (COMPETE), the Portuguese Strategic Reference Framework (QREN), the European Regional Development Fund (FEDER) through Project PTDC/MAT/111809/2009 VARIANT. The work of the first author was also supported by Government of Russian Federation, Grant 074-U01, Ministry of Education and Science of Russian Federation, Project 14.Z50.31.0031.

Appendix. Proofs

A.1. Auxiliary results

Assume that all the eigenvalues $\lambda_1, \ldots, \lambda_n$ of system (2) are different. Consider the Cauchy problem

$$\dot{\Phi}(t) = F\Phi(t), \qquad \Phi(0) = I, \tag{33}$$

where I is the identity matrix. The components of $\Phi(t)$ are denoted by $\phi_{i,j}(t)$. Recall some results from Smirnov et al. (2009) used below. Denote by $\mathcal{L}(\Phi)(s)$ the Laplace transform of the matrix $\Phi(t)$, the solution of problem (33). Applying the Laplace transform to differential equation (33), we obtain $\mathcal{L}(\dot{\Phi})(s) = F\mathcal{L}(\Phi)(s)$. From this we get

$$\mathcal{L}(\Phi)(s) = (sI - F)^{-1} = \Delta(s)^{-1}M(s), \tag{34}$$

where $M(s)$ is a matrix with polynomial elements and $\Delta(s) = \det(sI - F) = \prod_{j=1}^n (s - \lambda_j)$. Using the inversion theorem, we have

$$\Phi(t) = \frac{1}{2\pi i}\int_\Gamma \exp(st)\,\frac{M(s)}{\Delta(s)}ds, \tag{35}$$

where Γ is a closed curve containing inside all points λ_k, $k = \overline{1, n}$. From this we have

$$\phi_{n,l}(t) = -\frac{1}{2\pi i}\int_\Gamma \exp(st)\,\frac{\sum_{\beta=0}^{l-1} a_{n-\beta}s^{n+\beta-l-1}}{(s-\lambda_1)\ldots(s-\lambda_n)}ds \tag{36}$$

and

$$\phi_{1,n}(t) = \frac{1}{2\pi i}\int_\Gamma \exp(st)\,\frac{ds}{(s-\lambda_1)\ldots(s-\lambda_n)}. \tag{37}$$

Expanding the exponential function in series, using the equality

$$(1-h_1)^{-1}\dots(1-h_n)^{-1} = \sum_{q=0}^{\infty}\sum_{q_1+\cdots+q_n=q} h_1^{q_1}\dots h_n^{q_n}, \quad |h_j|<1,\ j=\overline{1,n}, \tag{38}$$

with $b_j = \lambda_j/\lambda$, and calculating the residue at infinity, from (36) and (37) we obtain

$$\phi_{n,l}(t) = -\sum_{\beta=0}^{l-1}\sum_{1\le i_1<\cdots<i_{n-\beta}\le n}\lambda_{i_1}\dots\lambda_{i_{n-\beta}} \times \sum_{\alpha\ge l-\beta}\frac{t^\alpha}{\alpha!}\sum_{q_1+\cdots+q_n=\alpha+\beta-l}\lambda_1^{q_1}\dots\lambda_n^{q_n} \tag{39}$$

and

$$\phi_{1,n}(t) = \sum_{\alpha=n-1}^{\infty}\frac{t^\alpha}{\alpha!}\sum_{q_1+\cdots+q_n=\alpha-n+1}\lambda_1^{q_1}\dots\lambda_n^{q_n} \tag{40}$$

respectively. The details can be found in Smirnov et al. (2009).

A.2. Proof of Theorem 1

Assume that all the eigenvalues $\lambda_1,\dots,\lambda_n$ of system (2) are different. From (39) we have

$$\phi_{n,1}(t) = -\lambda_1\dots\lambda_n\sum_{\alpha=1}^{\infty}\frac{t^\alpha}{\alpha!}\sum_{q_1+\cdots+q_n=\alpha-1}\lambda_1^{q_1}\dots\lambda_n^{q_n}. \tag{41}$$

Let $\omega = \max\{|\lambda_j| : j=\overline{1,n}\}$ and $t_* = \theta/(\omega n)$, θ being a parameter. Then we have

$$|\phi_{n,1}(t_*)| = \frac{|\lambda_1\dots\lambda_n|}{\omega n}\delta(\theta), \tag{42}$$

where

$$\delta(\theta) = \theta\left|1+\sum_{\alpha=2}^{\infty}\frac{t_*^{\alpha-1}}{\alpha!}\sum_{q_1+\cdots+q_n=\alpha-1}\lambda_1^{q_1}\dots\lambda_n^{q_n}\right|. \tag{43}$$

The number of terms in the last sum equals

$$\binom{\alpha+n-2}{n-1} = \left(1+\frac{n-1}{\alpha-1}\right)\left(1+\frac{n-1}{\alpha-2}\right)\dots\left(1+\frac{n-1}{1}\right)\le n^{\alpha-1}. \tag{44}$$

Since $|\lambda_j|\le\omega$, we get

$$\begin{aligned}\delta(\theta) &\ge \theta\left(1-\sum_{\alpha=2}^{\infty}\frac{\theta^{\alpha-1}}{\alpha!(\omega n)^{\alpha-1}}\sum_{q_1+\cdots+q_n=\alpha-1}\omega^{\alpha-1}\right)\\ &= \theta\left(1-\sum_{\alpha=2}^{\infty}\frac{\theta^{\alpha-1}}{\alpha!(\omega n)^{\alpha-1}}\omega^{\alpha-1}\binom{\alpha+n-2}{n-1}\right)\\ &\ge \theta\left(1-\sum_{\alpha=2}^{\infty}\frac{\theta^{\alpha-1}}{\alpha!}\right) = 2\theta+1-e^\theta.\end{aligned} \tag{45}$$

Since the function $\psi(\theta) = 2\theta+1-e^\theta$, $\theta\ge 0$, attains its maximum at $\theta = \log 2$, we obtain the result for different $\lambda_1,\dots,\lambda_n$. But the matrix of system (2) continuously depends on $\lambda_1,\dots,\lambda_n$ and $\phi_{n,1}(t_*)$ continuously depends on the matrix, thus the result is true for arbitrary $\lambda_1,\dots,\lambda_n$. □

A.3. Proof of Theorem 2

Assume that all the eigenvalues $\lambda_1,\dots,\lambda_n$ of system (2) are different. From (40) we obtain

$$\begin{aligned}\phi_{1,n}(t) &= \sum_{\alpha=n-1}^{\infty}\frac{t^\alpha}{\alpha!}\sum_{q_1+\cdots+q_n=\alpha-n+1}\lambda_1^{q_1}\dots\lambda_n^{q_n}\\ &= \frac{t^{n-1}}{(n-1)!}+\sum_{\alpha=n}^{\infty}\frac{t^\alpha}{\alpha!}\sum_{q_1+\cdots+q_n=\alpha-n+1}\lambda_1^{q_1}\dots\lambda_n^{q_n}.\end{aligned} \tag{46}$$

Let $\omega = \max\{|\lambda_j| : j=\overline{1,n}\}$, $t_* = \theta/\omega$, where parameter θ will be specified later. Then, as in the previous proof, we have

$$\begin{aligned}|\phi_{1,n}(t_*)| &\ge \frac{\theta^{n-1}}{\omega^{n-1}(n-1)!}-\sum_{\alpha=n}^{\infty}\frac{\theta^\alpha\omega^{\alpha-n+1}}{\alpha!\omega^\alpha}\binom{\alpha}{n-1}\\ &= \frac{\theta^{n-1}}{\omega^{n-1}(n-1)!}\left(1-\sum_{\alpha=n}^{\infty}\frac{\theta^{\alpha-n+1}}{(\alpha-n+1)!}\right)\\ &= \frac{\theta^{n-1}}{\omega^{n-1}(n-1)!}\left(2-e^\theta\right).\end{aligned} \tag{47}$$

Since $2-e^\theta\ge 1-\theta/\log 2$, whenever $0\le\theta\le\log 2$, maximizing the function $\psi(\theta) = \theta^{n-1}(1-\theta/\log 2)$ we obtain

$$\theta^* = \arg\max\psi(\theta) = \frac{n-1}{n}\log 2, \quad \psi(\theta^*)\ge\frac{(\log 2)^{n-1}}{en}. \tag{48}$$

This leads to (13). Thus, the theorem is proved for different $\lambda_1,\dots,\lambda_n$. We conclude as in the previous proof. □

A.4. Proof of Theorem 3

We need the following auxiliary estimate.

Lemma 1. *Assume that* $\mathrm{Re}\lambda_j<0$, $j=1,\dots,n$, *and* (14) *is satisfied. Then the following inequality holds*

$$\left|\sum_{1\le i_1<\cdots<i_\nu\le n}\lambda_{i_1}\dots\lambda_{i_\nu}\right|\ge|\lambda_1\dots\lambda_\nu|\ge\eta^\nu. \tag{49}$$

Proof. Since the coefficients of the characteristic polynomial $\Delta(\lambda)$ are real, the roots have the following structure: $\lambda_j\in R$, $j\in J=\{j_1,\dots,j_\mu\}$, and $\lambda_j=\alpha_j\pm i\beta_j$, $\alpha_j,\ \beta_j\in R$, $j\notin J$. (The set J can be empty or can coincide with the set of all indices.) We obviously have

$$\begin{aligned}\Delta(\lambda) &= \prod_{j=1}^{n}(\lambda-\lambda_j)\\ &= \lambda^n+\cdots+(-1)^\nu\sum_{1\le i_1<\cdots<i_\nu\le n}\lambda_{i_1}\dots\lambda_{i_\nu}\lambda^{n-\nu}+\cdots\\ &\quad+(-1)^n\lambda_1\dots\lambda_n\\ &= \prod_{r=1}^{\mu}(\lambda-\lambda_{j_r})\prod_{r=\mu+1}^{(n-\mu)/2}((\lambda-\alpha_{j_r})^2+\beta_{j_r}^2)\\ &= \prod_{r=1}^{\mu}(\lambda+|\lambda_{j_r}|)\prod_{r=\mu+1}^{(n-\mu)/2}((\lambda+|\alpha_{j_r}|)^2+\beta_{j_r}^2)\\ &= \prod_{r=1}^{\mu}(\lambda+|\lambda_{j_r}|)\prod_{r=\mu+1}^{(n-\mu)/2}(\lambda^2+|\lambda_{j_r}|^2+2\lambda|\alpha_{j_r}|).\end{aligned} \tag{50}$$

From this we see that the coefficient of $\lambda^{n-\nu}$ of the polynomial $\Delta(\lambda)$ is a sum of positive terms and therefore has the form

$$(-1)^{\nu} \sum_{1 \le i_1 < \cdots < i_\nu \le n} \lambda_{i_1} \ldots \lambda_{i_\nu} = |\lambda_1 \ldots \lambda_\nu| + \delta, \tag{51}$$

where $\delta > 0$. □

Proof of Theorem 3. Assume that all the eigenvalues $\lambda_1, \ldots, \lambda_n$ are different. Set $l = n - \nu + 1$. Then from (39) we have

$$\begin{aligned}
\phi_{n,n-\nu+1}(t) = &- \sum_{1 \le i_1 < \cdots < i_\nu \le n} \lambda_{i_1} \ldots \lambda_{i_\nu} t \\
&- \sum_{1 \le i_1 < \cdots < i_\nu \le n} \lambda_{i_1} \ldots \lambda_{i_\nu} \\
&\times \sum_{\alpha \ge 2} \frac{t^\alpha}{\alpha!} \sum_{q_1 + \cdots + q_n = \alpha - 1} \lambda_1^{q_1} \ldots \lambda_n^{q_n} \\
&- \sum_{\beta=0}^{n-\nu-1} \sum_{1 \le i_1 < \cdots < i_{n-\beta} \le n} \lambda_{i_1} \ldots \lambda_{i_{n-\beta}} \\
&\times \sum_{\alpha \ge n-\nu-\beta+1} \frac{t^\alpha}{\alpha!} \sum_{q_1 + \cdots + q_n = \alpha + \beta - n + \nu - 1} \lambda_1^{q_1} \ldots \lambda_n^{q_n}.
\end{aligned} \tag{52}$$

Since $\exp(a) - (1 + a) \le \exp(a) a^2/2$, $a \ge 0$, we get

$$\begin{aligned}
&\left| \sum_{\alpha \ge 2} \frac{t^\alpha}{\alpha!} \sum_{q_1 + \cdots + q_n = \alpha - 1} \lambda_1^{q_1} \ldots \lambda_n^{q_n} \right| \\
&\le \sum_{\alpha \ge 2} \frac{t^\alpha}{\alpha!} \sum_{q_1 + \cdots + q_n = \alpha - 1} \omega^{\alpha - 1} \\
&= \sum_{\alpha \ge 2} \frac{t^\alpha}{\alpha!} \omega^{\alpha - 1} \binom{\alpha + n - 2}{n - 1} \\
&\le \sum_{\alpha \ge 2} \frac{1}{\omega n} \frac{(\omega n t)^\alpha}{\alpha!} \le \frac{\omega n t^2}{2} \exp(\omega n t)
\end{aligned} \tag{53}$$

(see the proof of Theorem 1). Similarly, we obtain

$$\begin{aligned}
&\left| \sum_{\alpha \ge n-\nu-\beta+1} \frac{t^\alpha}{\alpha!} \sum_{q_1 + \cdots + q_n = \alpha + \beta - n + \nu - 1} \lambda_1^{q_1} \ldots \lambda_n^{q_n} \right| \\
&\le \sum_{\alpha \ge n-\nu-\beta+1} \frac{t^\alpha}{\alpha!} (\omega n)^{\alpha + \beta - n + \nu - 1} \\
&= \sum_{\alpha \ge n-\nu-\beta+1} \frac{(\omega n t)^\alpha}{\alpha!} (\omega n)^{\beta - n + \nu - 1} \\
&\le \frac{t^{n-\nu-\beta+1}}{(n - \nu - \beta + 1)!} \exp(\omega n t) \le \frac{t^2}{2} \exp(\omega n t),
\end{aligned} \tag{54}$$

whenever $\beta \le n - \nu - 1$ and $t < 1$. Using Lemma 1 and assumption (14) we get

$$\begin{aligned}
&\frac{\sum_{\beta=0}^{n-\nu-1} \left| \sum_{1 \le i_1 < \cdots < i_{n-\beta} \le n} \lambda_{i_1} \ldots \lambda_{i_{n-\beta}} \right|}{\left| \sum_{1 \le i_1 < \cdots < i_\nu \le n} \lambda_{i_1} \ldots \lambda_{i_\nu} \right|} \\
&= \frac{\sum_{m=\nu+1}^{n} \left| \sum_{1 \le i_1 < \cdots < i_m \le n} \lambda_{i_1} \ldots \lambda_{i_m} \right|}{\left| \sum_{1 \le i_1 < \cdots < i_\nu \le n} \lambda_{i_1} \ldots \lambda_{i_\nu} \right|}
\end{aligned}$$

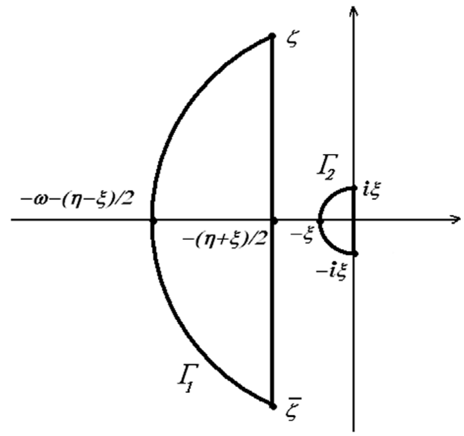

Fig. 6. Curves Γ_1 and Γ_2.

$$\begin{aligned}
&\le \sum_{m=\nu+1}^{n} \sum_{1 \le i_1 < \cdots < i_m \le n} \frac{\xi^m}{\eta^\nu} = \sum_{m=\nu+1}^{n} \binom{n}{m} \frac{\xi^m}{\eta^\nu} \\
&\le \sum_{m=0}^{n} \binom{n}{m} \frac{\xi^m}{\eta^\nu} = \frac{(1 + \xi)^n}{\eta^\nu}.
\end{aligned} \tag{55}$$

Set $t_* = \theta/(\omega n)$; θ will be chosen later. Then we have

$$\begin{aligned}
|\phi_{n,n-\nu+1}(t_*)| \ge &\left| \sum_{1 \le i_1 < \cdots < i_\nu \le n} \lambda_{i_1} \ldots \lambda_{i_\nu} \right| t_* \\
&\times \left(1 - \frac{t_* \exp(\omega n t_*)}{2} \left(\omega n + \frac{(1 + \xi)^n}{\eta^\nu} \right) \right) \\
&\ge \frac{\eta^\nu}{\omega n} \theta \left(1 - \alpha \frac{\theta e^\theta}{2} \right), \quad \alpha = 1 + \frac{(1 + \xi)^n}{\omega n \eta^\nu}.
\end{aligned} \tag{56}$$

From (15) we get $\alpha \le 1.1$. Taking $\theta = \log 2$ we have (16).

Arguing as at the end of the proof of Theorem 2 we obtain the result. □

A.5. Proof of Theorem 4

Let ζ and $\bar{\zeta}$ be the points of intersection of the circumference $|s| = \omega + (\eta - \xi)/2$ with the line $\mathrm{Re}s = -(\eta + \xi)/2$. Consider the closed curves $\Gamma_1 = \{s \mid |s| = \omega + (\eta - \xi)/2, \mathrm{Re}s < -(\eta + \xi)/2\} \cup [\bar{\zeta}, \zeta]$ and $\Gamma_2 = \{s \mid |s| = \xi, \mathrm{Re}s < 0\} \cup [-i\xi, i\xi]$ (see Fig. 6).

From (40) we have $\phi_{1,n}(t) = I_1 + I_2$, where

$$I_k = \frac{1}{2\pi i} \int_{\Gamma_k} \exp(st) \frac{ds}{(s - \lambda_1) \ldots (s - \lambda_n)}, \quad k = 1, 2. \tag{57}$$

Obviously, we have

$$|I_1| \le \frac{\pi(\omega + (\eta - \xi)/2)}{2\pi} \frac{e^{-(\eta+\xi)t/2}}{((\eta - \xi)/2)^n}. \tag{58}$$

Taking $t = 1/\xi$ and having in mind that due to (18) $\eta - \xi \ge 2$ and $\xi < 1$, we obtain

$$|I_1| \le c_1 \omega, \quad c_1 = 0.75 e^{-2}. \tag{59}$$

Now let us estimate I_2. Expanding the exponential function and using (38), we get $I_2 =$

$$\frac{1}{2\pi i} \int_{\Gamma_2} \frac{\sum_{q \ge 0, \alpha \ge 0, q_{\nu+1} + \cdots + q_n = s} \frac{t^\alpha}{\alpha!} s^{\alpha - n + \nu - s} \lambda_{\nu+1}^{q_{\nu+1}} \ldots \lambda_n^{q_n}}{\prod_{j=1}^{\nu} (s - \lambda_j)} ds. \tag{60}$$

Calculating the residue at $s = 0$, we get

$$I_2 = \frac{(-1)^\nu}{\prod_{j=1}^{\nu} \lambda_j} \sum_{\substack{\alpha \geq n-\nu-1 \\ q_{\nu+1}+\cdots+q_n \\ =\alpha-n+\nu+1}} \frac{t^\alpha}{\alpha!} \lambda_{\nu+1}^{q_{\nu+1}} \ldots \lambda_n^{q_n}. \tag{61}$$

Set $t = 1/\xi$. Then we have

$$I_2 = \frac{(-1)^\nu}{\prod_{j=1}^{\nu} \lambda_j} \frac{(1/\xi)^{n-\nu-1}}{(n-\nu-1)!} \times \left(1 + (n-\nu-1)! \sum_{\substack{\alpha \geq n-\nu \\ q_{\nu+1}+\cdots+q_n \\ =\alpha-n+\nu+1}} \frac{\left(\frac{1}{\xi}\right)^{\alpha-n+\nu+1} \lambda_{\nu+1}^{q_{\nu+1}} \ldots \lambda_n^{q_n}}{\alpha!} \right). \tag{62}$$

Observe that

$$\left| \sum_{\substack{\alpha \geq n-\nu \\ q_{\nu+1}+\cdots+q_n \\ =\alpha-n+\nu+1}} \frac{\left(\frac{1}{\xi}\right)^{\alpha-n+\nu+1} \lambda_{\nu+1}^{q_{\nu+1}} \ldots \lambda_n^{q_n}}{\alpha!} \right| \leq \sum_{\alpha \geq n-\nu} \left(\frac{1}{\xi}\right)^{\alpha-n+\nu+1} \frac{\xi^{\alpha-n+\nu+1}}{\alpha!} \binom{\alpha}{n-\nu} = \frac{e}{(n-\nu)!}. \tag{63}$$

Hence we get

$$|I_2| \geq \frac{c_2}{\xi^{n-\nu-1} \omega^\nu (n-\nu-1)!}, \tag{64}$$

where $c_2 \geq 1 - e/(n-\nu) \geq 1 - e/3$. Finally, we have

$$|\phi_{1,n}(1/\xi)| \geq |I_2| - |I_1|. \tag{65}$$

Combining the last inequality with (59) and (18) we obtain the result. $\square$

A.6. Proof of Theorem 7

We will need a few auxiliary lemmas. Below $s \in R$ does not belong to spectrum of A; it will be specified later (see (90)).

Lemma 2. *The vectors $b_k = (A - sI)^{-k} b$, $k = \overline{0, n-1}$, form a basis in R^n.*

Proof. Indeed, then the vectors b_k are well-defined. It suffices to show that they are linearly independent. Assume that $0 = \sum_{k=0}^{n-1} \beta_k b_k = \sum_{k=0}^{n-1} \beta_k (A - sI)^{-k} b$. Multiplying this equality by $(A - sI)^{n-1}$ we get

$$\begin{aligned} 0 &= \sum_{k=0}^{n-1} \beta_k (A - sI)^{n-k-1} b \\ &= \sum_{k=0}^{n-1} \beta_k \sum_{j=0}^{n-k-1} \binom{n-k-1}{j} (-s)^{n-k-1-j} A^j b \\ &= \sum_{j=0}^{n-1} \sum_{k=0}^{n-j-1} \beta_k \binom{n-k-1}{j} (-s)^{n-k-1-j} A^j b. \end{aligned} \tag{66}$$

Since the vectors $A^j b$, $j = \overline{0, n-1}$ are linearly independent, we have

$$\sum_{k=0}^{n-j-1} \beta_k \binom{n-k-1}{j} (-s)^{n-k-1-j} = 0, \quad j = \overline{0, n-1}. \tag{67}$$

Taking successively $j = n-1, n-2, \ldots$, we obtain $\beta_0 = 0, \beta_1 = 0$, etc. $\square$

Any vector $x \in R^n$ can be represented as $x = \sum_{k=0}^{n-1} \beta_k b_k$ and as

$$x = \sum_{m=0}^{n-1} \gamma_m A^m b. \tag{68}$$

The vectors $\beta = (\beta_0, \ldots, \beta_{n-1})$ and $\gamma = (\gamma_0, \ldots, \gamma_{n-1})$ satisfy the equality

$$\beta(s, \gamma) = M(s)\gamma, \tag{69}$$

where $M(s)$ is an $(n \times n)$-matrix.

Lemma 3. *The following representation takes place:*

$$\beta_k(s, \gamma) = \sum_{r=0}^{k} \beta_k^r(\gamma)(-s)^r, \quad k = \overline{0, n-1}. \tag{70}$$

Proof. Fix $\gamma \neq 0$ and set $\beta_k(s) = \beta_k(s, \gamma)$, $k = \overline{0, n-1}$. Multiplying the equality

$$\sum_{k=0}^{n-1} \beta_k (A - sI)^{-k} b = \sum_{m=0}^{n-1} \gamma_m A^m b \tag{71}$$

by $(A - sI)^{n-1}$ we get

$$\sum_{k=0}^{n-1} \beta_k (A - sI)^{n-k-1} b = \sum_{m=0}^{n-1} \gamma_m (A - sI)^{n-1} A^m b \tag{72}$$

or, equivalently,

$$\sum_{k=0}^{n-1} \sum_{m=0}^{n-k-1} \beta_k (-s)^{n-k-m-1} \binom{n-k-1}{m} A^m b = \sum_{m=0}^{n-1} \gamma_m (A - sI)^{n-1} A^m b. \tag{73}$$

Switching the order of summation we have

$$\sum_{m=0}^{n-1} \sum_{k=0}^{n-m-1} \beta_k \binom{n-k-1}{m} (-s)^{n-k-m-1} A^m b = \sum_{m=0}^{n-1} \gamma_m (A - sI)^{n-1} A^m b = \sum_{m=0}^{n-1} P_m(s) A^m b, \tag{74}$$

where the degree of the polynomials $P_m(s)$, $m = \overline{0, n-1}$ is less than or equal to $n-1$. From this we successively determine β_0, β_1, etc. The functions $\beta_k = \beta_k(s)$, $k = \overline{0, n-1}$, have a form of polynomials

$$\beta_k(s) = \sum_{r=0}^{r_k} \beta_k^r (-s)^r. \tag{75}$$

From

$$\sum_{m=0}^{n-1} \sum_{k=0}^{n-m-1} \sum_{r=0}^{r_k} \beta_k^r \binom{n-k-1}{m} (-s)^{n+r-k-m-1} A^m b = \sum_{m=0}^{n-1} P_m(s) A^m b, \tag{76}$$

we see that the degree of polynomials

$$\sum_{k=0}^{n-m-1} \sum_{r=0}^{r_k} \beta_k^r \binom{n-k-1}{m} (-s)^{n+r-k-m-1} \tag{77}$$

is also less than or equal to $n-1$. Set $m=n-q$, $q=\overline{1,n}$. We see that the degree of the polynomials

$$\sum_{k=0}^{q-1}\sum_{r=0}^{r_k}\beta_k^r\binom{n-k-1}{n-q}(-s)^{s+r-k-1} \tag{78}$$

does not exceed $n-1$. Considering $q=\overline{1,n}$, we get $\beta_0^r=0$ whenever $r>n-q$. Then considering $q=\overline{2,n}$, we get $\beta_1^r=0$ for $r>n-q+1$, and so on. Finally, we obtain $\beta_k^r=0$ whenever $r>k$, $k=\overline{0,n-1}$. Thus we have (70). □

As a special case of this lemma we get

Lemma 4. *The coordinates α_k, $k=\overline{0,n-1}$, of the vector Ab with respect to the basis $b_k=(A-sI)^{-k}b$, $k=\overline{0,n-1}$, are $O(|s|^{n-1})$ as $|s|\to\infty$.*

Recall that by matrix norm we intend the norm induced by the vector norm.

Lemma 5. *The norm of the matrix $M^{-1}(s)$ is $O(1)$ as $|s|\to\infty$.*

Proof. Indeed, from Lemma 3 we see that $\beta_k(s,\gamma)s^{-k-m}=o(1)$, $|s|\to\infty$, $m=\overline{1,n-1}$. Dividing (73) by $(-s)^{n-1}$ we obtain

$$\beta_0 b+c(s)=\sum_{m=0}^{n-1}\gamma_m A^m b+\sum_{m=0}^{n-1}\gamma_m d_m(s), \tag{79}$$

where $|c(s)|=o(1)$ and $|d_m(s)|=o(1)$, $m=\overline{0,n-1}$, as $|s|\to\infty$. From this we get

$$\gamma_m=\begin{cases}\beta_0+o(1), & m=0,\\ o(1), & m=\overline{1,n-1}.\end{cases} \quad \square \tag{80}$$

Proof of Theorem 7. By Lemma 2 the vectors $b_k=(A-sI)^{-k}b$, $k=\overline{0,n-1}$, form a basis in R^n. Let $x=\sum_{k=0}^{n-1}\beta_k b_k$ and $Ab=\sum_{k=0}^{n-1}\alpha_k b_k$. Choose linear feedback $u(x)=u(\beta)=s\beta_0-(\beta_1+\beta_0\alpha_0)$. Since $\dot x=\sum_{k=0}^{n-1}\dot\beta_k b_k$ and

$$\begin{aligned}Ax+u(x)b &= A\sum_{k=0}^{n-1}\beta_k b_k+(s\beta_0-(\beta_1+\beta_0\alpha_0))b\\ &=(A-sI)\sum_{k=1}^{n-1}\beta_k b_k+s\sum_{k=1}^{n-1}\beta_k b_k+\beta_0 Ab\\ &\quad+(s\beta_0-(\beta_1+\beta_0\alpha_0))b\\ &=\sum_{k=1}^{n-2}(s\beta_k+\beta_{k+1})b_k+s\beta_{n-1}b_{n-1}\\ &\quad+\beta_0\sum_{k=1}^{n-1}\alpha_k b_k+s\beta_0 b,\end{aligned} \tag{81}$$

the system $\dot x=Ax+u(x)b$ in the coordinates β takes the form

$$\dot\beta_0=s\beta_0, \tag{82}$$

$$\dot\beta_1=s\beta_1+\beta_2+\beta_0\alpha_1, \tag{83}$$

$$\dots \tag{84}$$

$$\dot\beta_{n-2}=s\beta_{n-2}+\beta_{n-1}+\beta_0\alpha_{n-2}, \tag{85}$$

$$\dot\beta_{n-1}=s\beta_{n-1}+\beta_0\alpha_{n-1}. \tag{86}$$

Put

$$\delta^2=(n-1)\max_{k=\overline{1,n-1}}\alpha_k^2 \tag{87}$$

and

$$W(\beta)=\sum_{k=1}^{n-1}\beta_k^2+\delta^2\beta_0^2. \tag{88}$$

Then we see that $dW(\beta)/dt$ is equal to

$$\begin{aligned}&=2sW(\beta)+2\left(\sum_{k=1}^{n-2}(\beta_k\beta_{k+1}+\beta_k\beta_0\alpha_k)+\beta_{n-1}\beta_0\alpha_{n-1}\right)\\ &\le 2sW(\beta)+\sum_{k=1}^{n-2}(\beta_k^2+\beta_{k+1}^2)+\sum_{k=1}^{n-1}\left(\frac{\alpha_k^2\beta_k^2(n-1)}{\delta^2}+\frac{\delta^2\beta_0^2}{n-1}\right)\\ &\le(2s+3)W(\beta).\end{aligned} \tag{89}$$

Now fix

$$s=-\sigma-3/2, \tag{90}$$

then the last inequality becomes

$$\dot W(\beta)\le-2\sigma W(\beta). \tag{91}$$

We introduced Lyapunov function $W(\beta)$ in variables β, now convert it into $V(x)$. Recall that $x=H\gamma$ (see (68)), where $H=[b\,Ab\dots A^{n-1}b]$ is the controllability matrix. Due to controllability assumption H^{-1} exists and both H, H^{-1} do not depend on σ. On the other hand $\beta=M(s)\gamma$ (see (69)), and $\|M(s)\|=O(|s|^{n-1})\le c_1(\sigma+3/2)^{n-1}$ (see (70)), while $\|M(s)^{-1}\|=O(1)$ (Lemma 5). We have $W(\beta)=(Q\beta,\beta)$, where $Q=\mathrm{diag}(\delta^2,1,\dots,1)$ (88), with $\delta^2=O(\alpha^2)$ (see (87)). From Lemma 4 we conclude that the condition number satisfies the inequality $\chi(Q)\le c_2(\sigma+3/2)^{2(n-1)}$. Collecting all these relations we obtain

$$W(\beta)=W(MH^{-1}x)=V(x)=(P^{-1}x,x), \tag{92}$$

with

$$P^{-1}=(H^{-1})^T M^T Q M H^{-1}, \tag{93}$$

and

$$\chi(P)\le c_3(\sigma+3/2)^{2(n-1)}. \tag{94}$$

Final remark is that inequality

$$\dot V(x)\le-2\sigma V(x) \tag{95}$$

for quadratic Lyapunov function $V(x)$ and linear system $\dot x=Fx$ implies $\mathrm{Re}\lambda_i\le-\sigma$ for all eigenvalues λ_i of the matrix F.

Thus we have constructed linear feedback and quadratic Lyapunov function with the desired properties. □

References

Balandin, L., & Kogan, M. (2009). Lyapunov function method for control law synthesis under one integral and several phase constraints. *Differential Equations*, *45*, 670–679.

Bewley, T., & Liu, S. (1998). Optimal and robust control and estimation of linear paths to transition. *Journal of Fluid Mechanics*, *365*, 305–349.

Boyd, S., El Ghaoui, L., Ferron, E., & Balakrishnan, V. (1994). *Linear matrix inequalities in systems and control theory*. Philadelphia: SIAM.

Bulgakov, A. (1980). An efficiently calculable parameter for the stability property of a system of linear differential equations with constant coefficients. *Siberian Mathematical Journal*, *21*, 339–347.

Bushenkov, V., & Smirnov, G. (1997). *Stabilization problems with constraints: analysis and computational aspects*. Amsterdam: Gordon and Breach.

Feldbaum, A. (1948). On the distribution of roots of characteristic equations of control systems. *Avtomatika i Telemekhanika*, 253–279.

Grant, M., & Boyd, S. (n.d.). Cvx: Matlab software for disciplined convex programming (web page and software): http://stanford.edu/boyd/cvx.

Herman, I., Martinec, D., Z, H., & Sebek, M. (2015). Nonzero bound on fiedler eigenvalue causes exponential growth of h-infinity norm of vehicular platoon. *IEEE Transactions on Automatic Control*, *60*, 2248–2253.

Hinrichsen, D., Plischke, E., & Wurth, F. (2002). State feedback stabilization with guaranteed transient bounds. In *Proc. 15th int. symp. math. theory networks & syst.* p. Paper no. 2132.

Izmailov, R. (1987). The peak effect in stationary linear systems with scalar inputs and outputs. *Automation and Remote Control*, *48*, 1018–1024.

Liberzon, D. (2003). *Switching in systems and control*. Boston: Birkhäuser.

Moler, C., & Van Loan, C. (1978). Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review*, *20*, 801–836.

Moler, C., & Van Loan, C. (2003). Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review*, *45*, 3–49.

Nechepurenko, M. (2002). Bounds for the matrix exponential based on the lyapunov equation and limits of the hausdorff set. *Computational Mathematics and Mathematical Physics*, *42*, 125–134.

Polotskij, V. (1981). Estimation of the state of single-output linear systems by means of observers. *Automation and Remote Control*, *41*, 1640–1648.

Polyak, B., & Smirnov, G. (2014). Large deviations in continuous-time linear single-input control systems. In *Proceedings 19-th world ifac congress, Cape Town* (pp. 5586–5591).

Polyak, B., Tremba, A., Khlebnikov, M., Shcherbakov, P., & Smirnov, G. (2015). Large deviations in linear control systems with nonzero initial conditions. *Automation and Remote Control*, *75*, 957–976.

Reddy, S., & Henningson, D. (1993). Energy growth in viscous channel flows. *Journal of Fluid Mechanics*, *252*, 209–238.

Smirnov, G., Bushenkov, V., & Miranda, F. (2009). Advances on the transient growth quantification in linear control systems. *International Journal of Applied Mathematics & Statistics*, *14*, 82–92.

Sussman, H., & Kokotovic, P. (1991). The peaking phenomenon and the global stabilization of nonlinear systems. *IEEE Transactions on Automatic Control*, *36*, 424–439.

Van Dorsselaer, J., Kraaijevanger, J., & Spijker, M. (1993). Linear stability analysis in the numerical solution of initial value problems. *Acta Numerica*, *2*, 199–237.

Whidborne, J., & Amar, N. (2011). Computing the maximum transient energy growth. *BIT Numerical Mathematics*, *51*, 447–557.

Whidborne, J., & McKernan, J. (2007). On minimizing maximum transient energy growth. *IEEE Transactions on Automatic Control*, *52*, 1762–1767.

Wonham, W. (1979). *Linear multivariable control: a geometric approach*. New York: Springer-Verlag.

Boris T. Polyak received Ph.D. degree in mathematics from Moscow State University in 1963 and Doctor of Science degree in engineering from Institute for Control Science, Moscow, in 1977. He has been Head of Ya.Z. Tsypkin Laboratory, Institute for Control Science of Russian Academy of Sciences, Moscow, Russia (where he is currently Chief Researcher) and Professor of Moscow University of Physics and Engineering. He is the author of more than 200 papers in peer-review journals and 4 monographs, including Introduction to Optimization, Russian and English editions, and Robust Stability and Control, coauthored with P.S. Scherbakov. He is on the Editorial Boards of numerous journals, IFAC Fellow, recipient of EURO-2012 Gold Medal. His research interests include mathematical programming, nonsmooth optimization, stochastic estimation and optimization, linear and nonlinear analysis and design, robust stability, chaos control, Monte-Carlo simulation.

Georgi Smirnov was born in 1961 in Moscow, USSR. His Ph.D. is from the Moscow State University (1988). He worked at the Computing Centre of The Soviet Academy of Sciences. Now he is professor at the University of Minho, Portugal. Georgi Smirnov's research interests lie in applied mathematics, optimization, differential equations, control theory, as well as in space flight dynamics and control.