ISSN 1064-2307, Journal of Computer and Systems Sciences International, 2015, Vol. 54, No. 2, pp. 165–183. © Pleiades Publishing, Ltd., 2015.
Original Russian Text © N.E. Zubov, E.A. Mikrin, M.Sh. Misrikhanov, A.V. Proletarskii, V.N. Ryabchenko, 2015, published in Izvestiya Akademii Nauk. Teoriya i Sistemy Upravleniya, 2015, No. 2, pp. 3–21.

CONTROL IN DETERMINISTIC SYSTEMS

Reduction of Large-Scale Dynamical Systems by the Krylov Subspaces Method: Analysis of Approaches

N. E. Zubov, E. A. Mikrin, M. Sh. Misrikhanov, A. V. Proletarskii, and V. N. Ryabchenko

Bauman State Technical University, Vtoraya Baumanskaya ul. 5, Moscow, 105005 Russia
Korolev Rocket and Space Corporation "Energia", ul. Lenina 4a, Korolev, Moscow oblast, 141070 Russia
Moscow Institute of Physics and Technology, Institutskii proezd 9, Dolgoprudnyi, Moscow oblast, 141700 Russia
e-mail: Nik.zubov@gmail.com

Received July 24, 2014; in final form, October 27, 2014

Abstract—Approaches and numerical algorithms for reducing the order of mathematical models of multidimensional dynamical systems that are based on the Krylov's subspaces method are described. To calculate matrices representing the reduced models in the state space, Lanczos' and Arnoldi's methods are used. Practical examples of the reduction of large-scale systems are presented.

DOI: 10.1134/S1064230715020148

INTRODUCTION

In 1931, academician A.N. Krylov published work [1] on the subject that is currently known as *Krylov subspaces*, or the *Krylov subspaces method*. The paper was devoted to the calculation of coefficients of the characteristic polynomial of a given number matrix. In this work, Krylov discussed, in particular, the efficiency of calculations and defined computational costs as the number of multiplication operations (which is not typical of a mathematical publication in the 1930s). He suggested a method that war superior to other computational methods known by that time and, since then, is widely used all over the world, especially for iterative solving systems of linear matrix equations of high orders.

The Krylov subspaces method is also widely used in the analysis of stability, controllability, and observability of linear dynamical systems in the state space, as well as in synthesis of feedback control laws (see, e.g., [2–14] and references therein).

Another application of the Krylov subspaces method is reduction of mathematical models of dynamical systems given, for example, in the form of Markov's processes. The reduction of mathematical models of dynamical systems based on the *Krylov subspaces* method is an alternative to the well-known reduction method based on the SVD decomposition. It is applied to large-scale dynamical systems with the dimension of the state space $n \gg 1$ in the case of sparse system matrices.

The key component of the reduction based on the Krylov subspaces is the so-called *moment fitting* (adjustment of Markov's parameters). The idea of the reduction is to establish correspondence between the moments of the transfer function (matrix) of the original higher-order system Σ in its expansion in a Laurent series and the moments of the lower-order model Σ_r. In so doing, the controlled and observed subspaces of the system are used.

This paper is devoted to the analysis of the known approaches to reducing mathematical models of dynamical systems and numerical algorithms constructed on the basis of the Krylov subspaces method and to solving complicated problems with the use of the algorithms presented.

1. DESCRIPTION OF DYNAMICAL SYSTEMS

We consider a linear time-invariant (LTI) system Σ of the form

$$\Sigma : \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{cases} \quad \Leftrightarrow \quad \Sigma = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \tag{1.1}$$

where $\mathbf{x} = \left(x_1 \middle| \cdots \middle| x_n\right)^{\mathrm{T}} \in \mathbb{R}^n$ is a state vector, $\mathbf{u} = \left(u_1 \middle| \cdots \middle| u_m\right)^{\mathrm{T}} \in \mathbb{R}^m$ is an input vector, $\mathbf{y} = \left(y_1 \middle| \cdots \middle| y_p\right)^{\mathrm{T}} \in \mathbb{R}^p$ is an output vector, and $\mathbf{x}_0$ is an initial state.

If $m = p = 1$, (1.1) is a SISO (single-input, single-output) system. If $m, p > 1$, then (1.1) is a MIMO (multi-input, multi-output) system. When $m = 1$, $p > 1$ or $m > 1$, $p = 1$, we have a SIMO (single-input, multi-output) or MISO (multi-input, single-output) system, respectively.

For zero initial conditions, we define for system (1.1) the convolution operator $\mathcal{S}$ from $\mathbf{u}(t)$ to $\mathbf{y}(t)$ as [12]

$$\mathcal{S} : \mathbf{u}(t) \mapsto \mathbf{y}(t) = \mathbf{G} * \mathbf{u} = \int_0^\infty \mathbf{G}(t-\tau)\mathbf{u}(\tau)d\tau \tag{1.2}$$

with the kernel having the form of a weight (pulse) function (matrix) $\mathbf{G}(t)$. The Laplace transformation of the function $\mathbf{G}(t)$ yields the transfer function (matrix)

$$\mathbf{F}(s) \triangleq \mathcal{L}(\mathbf{G})(s) \triangleq \int_0^\infty \mathbf{G}(t)e^{-st}dt = \mathbf{C}\left(s\mathbf{I}_n - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D}, \tag{1.3}$$

where $\mathbf{I}_n$ is the identity matrix of size $n \times n$.

In what follows, we assume that system (1.1) is completely controllable [15], i.e., the controllability matrix

$$\mathcal{C} = \left(\mathbf{B} \middle| \mathbf{AB} \middle| \mathbf{A}^2\mathbf{B} \middle| \cdots \middle| \mathbf{A}^{n-1}\mathbf{B}\right) \tag{1.4}$$

has full rank,

$$\operatorname{rank}\mathcal{C} = n, \tag{1.5}$$

and completely observable [15], i.e., the observability matrix

$$\mathcal{O} = (\mathbf{C}^{\mathrm{T}} \middle| \mathbf{A}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}} \middle| (\mathbf{A}^{\mathrm{T}})^2\mathbf{C}^{\mathrm{T}} \middle| \cdots \middle| (\mathbf{A}^{\mathrm{T}})^{n-1}\mathbf{C}^{\mathrm{T}}) \tag{1.6}$$

is also of full rank,

$$\operatorname{rank}\mathcal{O} = n. \tag{1.7}$$

Let system (1.1) be Lyapunov stable. Then, there exist Gramians $\mathcal{P}$ and $\mathcal{Q}$ given by [5]

$$\mathcal{P} \triangleq \int_0^\infty e^{\mathbf{A}\tau}\mathbf{B}\mathbf{B}^{\mathrm{T}}e^{\mathbf{A}^{\mathrm{T}}\tau}d\tau, \quad \mathcal{Q} \triangleq \int_0^\infty e^{\mathbf{A}^{\mathrm{T}}\tau}\mathbf{C}^{\mathrm{T}}\mathbf{C}e^{\mathbf{A}\tau}d\tau \tag{1.8}$$

such that $\mathcal{P} = \mathcal{P}^{\mathrm{T}} \geq 0$ and $\mathcal{Q} = \mathcal{Q}^{\mathrm{T}} \geq 0$. Moreover, matrices (1.8) are solutions to the corresponding algebraic Lyapunov equations

$$\mathbf{A}\mathcal{P} + \mathcal{P}\mathbf{A}^{\mathrm{T}} + \mathbf{B}\mathbf{B}^{\mathrm{T}} = 0, \quad \mathbf{A}^{\mathrm{T}}\mathcal{Q} + \mathcal{Q}\mathbf{A} + \mathbf{C}^{\mathrm{T}}\mathbf{C} = 0. \tag{1.9}$$

Let us return to the convolution operator $\mathcal{S}$ given by (1.2) and, confining its domain and range, introduce the Hankel operator [12] as

$$\mathcal{H} : \mathbf{u}_-(t) \mapsto \mathbf{y}_+(t) = \mathcal{H}(\mathbf{u}_-) = \int_{-\infty}^0 \mathbf{G}(t-\tau)\mathbf{u}_-(\tau)\,d\tau, \quad t \geq 0. \tag{1.10}$$

Here, speaking in non-strict terms, $\mathbf{u}_-(t)$ are "past" inputs and $\mathbf{y}_+(t)$ are "future" outputs. In other words, the Hankel operator maps "past" inputs into "future" outputs. Unlike the convolution operator, the Hankel operator has a finite rank [12], which does not exceed n, and, hence, a finite set of singular values $\sigma_i(\Sigma)$, which are defined as follows.

Let system (1.1) be completely controllable, completely observable, and asymptotically stable. Then, the singular values $\sigma_i(\Sigma)$ are positive square roots of the eigenvalues λ_i of the product $\mathcal{P}\mathcal{Q}$ of Gramians from (1.8) or (1.9), i.e.,

$$\sigma_i(\Sigma) = \sqrt{\lambda_i(\mathcal{P}\mathcal{Q})}, \quad i = \overline{1,n}. \tag{1.11}$$

The greatest singular value $\sigma_{\max}(\Sigma)$ of the Hankel operator determines the Hankel norm of system Σ (1.1) [12]; i.e.,

$$\|\Sigma\|_{\mathcal{H}} \triangleq \sigma_{\max}(\Sigma). \tag{1.12}$$

Note that the norm $\mathcal{H}_\infty$ of system Σ is

$$\|\Sigma\|_{\mathcal{H}_\infty} \triangleq \sup_{\omega\in\mathbb{R}} \sigma_{\max}\left(\mathbf{G}(j\omega)\right) \tag{1.13}$$

and the $\mathcal{H}_2$ norm is

$$\|\Sigma\|_{\mathcal{H}_2} \triangleq \left(\int_{-\infty}^{\infty} \mathrm{tr}(\mathbf{G}^*(j\omega)\mathbf{G}(j\omega))d\omega\right)^{0.5}. \tag{1.14}$$

Here, the symbol * denotes conjugation.

2. KRYLOV SUBSPACES

The Krylov subspaces method is widely used for solving various computational problems and, first of all, in iterative solving high-dimensional matrix equations of the form [17]

$$\mathbf{Ax} = \mathbf{b}, \tag{2.1}$$

where $\mathbf{A}$ and $\mathbf{b}$ are given $n \times n$ matrix and $n \times 1$-dimensional vector, respectively.

Let $\mathbf{x}_0 \in \Re^n$ be an initial approximation of a solution to equation (2.1), where $\Re = \mathbb{R}$ or $\Re = \mathbb{C}$,

$$\mathbf{r}_0 = \mathbf{b} - \mathbf{Ax}_0 \tag{2.2}$$

be an initial residual, and

$$\mathcal{H}_m(\mathbf{A}, \mathbf{r}_0) = \mathrm{span}\left\{\mathbf{r}_0, \mathbf{Ar}_0, \mathbf{A}^2\mathbf{r}_0, \ldots, \mathbf{A}^{m-1}\mathbf{r}_0\right\} \tag{2.3}$$

be a subspace of dimension m determined by the matrix $\mathbf{A}$ and vector $\mathbf{r}_0$. For subspace (2.3), the inclusion

$$\mathcal{H}_m \subseteq \mathcal{H}_{m+1}$$

holds.

The methods relying on subspaces $\mathcal{H}$ are iterative ones [18]. By means of these methods, the mth iteration of a solution to equation (2.1) is sought in the form

$$\mathbf{x}_m = \mathbf{x}_0 + q_{m-1}(\mathbf{A})\mathbf{r}_0, \tag{2.4}$$

where q_{m-1} is a matrix polynomial in $\mathbf{A}$ of the degree less than or equal to $m - 1$. If the matrix equation (2.1) is defined over $\mathbb{R}$, then the coefficients of q_{m-1} are also real.

According to (2.4), the residual of a solution to equation (2.1) on the mth step can be written as

$$\mathbf{r}_m = \mathbf{b} - \mathbf{Ax}_m = \mathbf{r}_0 - \mathbf{A}q_{m-1}(\mathbf{A})\mathbf{r}_0 = p_m(\mathbf{A})\mathbf{r}_0, \tag{2.5}$$

where p_m is a remainder polynomial. Similarly, if $\mathbf{x}_*$ is a solution to (2.1), then

$$\mathbf{x}_m - \mathbf{x}_* = p_m(\mathbf{A})(\mathbf{x}_0 - \mathbf{x}_*).$$

In the course of the iterative construction of basis $\mathcal{H}_m$, any method based on subspaces (2.3) uses on each iteration one or two multiplications of the matrix by a vector in the form $\mathbf{z} = \mathbf{Av}$ (or $\mathbf{y} = \mathbf{A}^{\mathrm{T}}\mathbf{w}$). Therefore, the corresponding algorithms can be used both for solving problems with explicit system matrices and for solving problems in which the matrix is available only through the multiplication by a vector [16]. It should be noted that the convergence of the iteration process is ensured by using an appropriate algorithm.

In all methods based on the subspaces $\mathcal{K}_m$, the iteration procedures are initialized by specifying an initial approximation and the corresponding initial residual (2.2). The mth approximation is found in the form

$$\mathbf{x}_0 + \mathcal{K}_m(\mathbf{A}, \mathbf{r}_0).$$

Without loss of generality, we may assume that $\mathbf{x}_0 = 0$ and $\mathbf{r}_0 = \mathbf{b}$. Based on the above, we introduce the following definition.

Definition 1. For a system Σ (1.1) given by a pair $(\mathbf{A}, \mathbf{b})$, where $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{b} \in \mathbb{R}^n$, the kth Krylov sequence is formed by the following k column vectors:

$$\mathcal{K}_k(\mathbf{A}, \mathbf{b}) \equiv \left\{\mathbf{b}, \mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}, \ldots, \mathbf{A}^{k-1}\mathbf{b}\right\}. \tag{2.6}$$

The linear space spanned by the column vectors (2.6), i.e.,

$$\mathcal{K}_k(\mathbf{A}, \mathbf{b}) = \operatorname{span}\left\{\mathbf{b}, \mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}, \ldots, \mathbf{A}^{k-1}\mathbf{b}\right\}, \tag{2.7}$$

is called the kth Krylov subspace.

By the Hamilton–Cayley theorem, the Krylov subspaces satisfy the condition

$$\forall\, k > n : \mathcal{K}_k(\mathbf{A}, \mathbf{b}) = \mathcal{K}_n(\mathbf{A}, \mathbf{b}). \tag{2.8}$$

Condition (2.8) yields direct relationship of the Krylov subspaces with the controllability and observability matrices $\mathcal{C}$ (1.4) and $\mathcal{O}$ (1.6) of system Σ (1.1).

The approach is extended to the so-called partial reachability matrix

$$\mathcal{R}_k(\mathbf{A}, \mathbf{B}, \sigma) = ((\sigma\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{B} \,|\, (\sigma\mathbf{I}_n - \mathbf{A})^{-2}\mathbf{B} \,|\cdots|\, (\sigma\mathbf{I}_n - \mathbf{A})^{-k}\mathbf{B}) \tag{2.9}$$

and partial observability matrix

$$\mathcal{O}_k(\mathbf{C}, \mathbf{A}, \sigma) = ((\sigma\mathbf{I}_n - \mathbf{A}^{\mathrm{T}})^{-1}\mathbf{C}^{\mathrm{T}} \,|\cdots|\, (\sigma\mathbf{I}_n - \mathbf{A}^{\mathrm{T}})^{-k}\mathbf{C}^{\mathrm{T}}). \tag{2.10}$$

3. PROJECTORS

A projector $\mathbf{P} \in \mathbb{C}^{n\times n}$ onto a subspace $\mathcal{M} \subseteq \mathbb{C}^n$ is defined to be a linear idempotent mapping in $\mathbb{C}^n$ [19]:

$$\mathbf{P}^2 = \mathbf{P}, \quad \mathbf{P}\mathbf{x} \in \mathcal{M}, \quad \forall \mathbf{x} \in \mathbb{C}^n. \tag{3.1}$$

From the idempotency of the operator $\mathbf{P}$, it follows that

$$\ker(\mathbf{P}) = \operatorname{range}(\mathbf{I} - \mathbf{P}); \tag{3.2}$$

i.e., the kernel of the projector $\mathbf{P}$ coincides with the range of the mapping $(\mathbf{I} - \mathbf{P})$.

Any vector $\mathbf{x} \in \mathbb{C}^n$ can be represented as

$$\mathbf{x} = \mathbf{P}\mathbf{x} + (\mathbf{I} - \mathbf{P})\mathbf{x}; \tag{3.3}$$

hence, the space $\mathbb{C}^n$ can be decomposed into the direct sum of the subspaces:

$$\mathbb{C}^n = \ker(\mathbf{P}) \oplus \operatorname{range}(\mathbf{P}). \tag{3.4}$$

From (3.4), it follows that, for any pair of subspaces $\mathcal{N} \subset \mathbb{C}^n$, $\mathcal{M} \subset \mathbb{C}^n$ whose direct sum of the form (3.4) is $\mathbb{C}^n$ and for any vector $\mathbf{x} \in \mathbb{C}^n$, vector $\mathbf{P}\mathbf{x}$ satisfies the conditions

$$\mathbf{P}\mathbf{x} \in \mathcal{M}, \tag{3.5}$$

$$\mathbf{x} - \boldsymbol{P}\mathbf{x} \in \mathcal{N}. \tag{3.6}$$

In this case, we say that the linear mapping $\mathbf{P} \in \mathbb{C}^{n\times n}$ projects $\mathbf{x} \in \mathbb{C}^n$ into $\mathcal{M} \subset \mathbb{C}^n$ in the direction of (or parallel to) the subspace $\mathcal{N} \subset \mathbb{C}^n$.

An orthogonal projector $\mathbf{P}$ onto a subspace $\mathcal{N} \subseteq \mathbb{C}^n$ is also defined as a linear mapping of $\mathbb{C}^n$ satisfying condition (3.1) and the additional requirement [19]

$$(\mathbf{I} - \mathbf{P})\mathbf{x} \in \mathcal{N}^{\perp}. \tag{3.7}$$

According to equations (3.5), (3.6), matrix representation of a projector needs two bases:

a basis for the subspace $\mathcal{M} = \mathrm{range}(\boldsymbol{P})$ with the matrix

$$\mathbf{V} = (\mathbf{v}_1 | \mathbf{v}_2 | \mathbf{v}_3 | \cdots | \mathbf{v}_m) \tag{3.8}$$

and a basis for the subspace $\mathcal{N}^{\perp}$ with the matrix

$$\mathbf{W} = (\mathbf{w}_1 | \mathbf{w}_2 | \mathbf{w}_3 | \cdots | \mathbf{w}_m). \tag{3.9}$$

If matrices (3.8) and (3.9) are biorthogonal [20], i.e.,

$$\mathbf{W}^*\mathbf{V} = \mathbf{I}_m, \tag{3.10}$$

then the projector $\mathbf{P}$ can be represented as

$$\mathbf{P} = \mathbf{V}\mathbf{W}^*. \tag{3.11}$$

Otherwise, the following more general representation holds:

$$\mathbf{P} = \mathbf{V}(\mathbf{W}^*\mathbf{V})^{-1}\mathbf{W}^*. \tag{3.12}$$

The set of projectors can be partitioned into two classes: orthogonal (satisfying the condition $\mathcal{M} = \mathcal{N}^{\perp}$) and oblique projectors. The orthogonal projectors satisfy the conditions [20]

$$\ker(\mathbf{P}) = \mathrm{range}(\mathbf{P})^{\perp}, \tag{3.13}$$

$$(\mathbf{I} - \mathbf{P})\mathbf{x} \in \mathcal{N}^{\perp}, \tag{3.14}$$

$$\mathbf{P}\mathbf{x} \in \mathcal{M}, \tag{3.15}$$

$$\mathbf{x} - \mathbf{P}\mathbf{x} \perp \mathcal{M}. \tag{3.16}$$

All other projectors are classified as nonorthogonal (oblique) ones.

Returning to the iterative solving equation (2.1), we note that, in the given case, a problem of minimization of an additional condition is solved as a rule [18].

Let

$$\mathbf{r}_m = \mathbf{b} - \mathbf{A}\mathbf{x}_m \tag{3.17}$$

be a residual on the mth step. Then, for an additional condition, we may consider [17, 18]

(1) the Petrov–Galerkin orthogonality condition

$$\mathbf{r}_m \perp \mathcal{L}_m, \tag{3.18}$$

where $\mathcal{L}_m$ is an m-dimensional subspace; if $\mathcal{L}_m = \mathcal{K}_m$, then (3.18) is called

(2) the Galerkin orthogonality condition

$$\mathbf{r}_m \perp \mathcal{K}_m, \tag{3.19}$$

or

(3) the minimum residual condition

$$\|\mathbf{r}_m\| = \min_{\mathbf{x}_0 + \mathcal{K}_m} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|. \tag{3.20}$$

If $\mathcal{L}_m = \mathbf{A}\mathcal{K}_m$, then condition (3.20) is equivalent to the Petrov–Galerkin orthogonality condition (3.18) [18].

4. MODEL REDUCTION

Let us return to the linear time-invariant system Σ of form (1.1) and its transfer function (matrix) of form (1.3) given by

$$\mathbf{F}(s) = \mathbf{C}\left(s\mathbf{I}_n - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D}. \tag{4.1}$$

Let us rewrite the transfer function (4.1) as

$$\mathbf{F}(s) = \mathbf{C}\mathbf{F}_B(s) = \mathbf{F}_C^{\mathrm{T}}(s)\mathbf{B}, \tag{4.2}$$

where $\mathbf{F}_B(s)$ and $\mathbf{F}_C(s)$ are solutions to the linear parametric equations

$$\left(s\mathbf{I}_n - \mathbf{A}\right)\mathbf{F}_B(s) = \mathbf{B}, \tag{4.3}$$

$$\mathbf{F}_C^{\mathrm{T}}(s)\left(s\mathbf{I}_n - \mathbf{A}\right) = \mathbf{C}. \tag{4.4}$$

We consider the model reduction (order reduction) problem as search for appropriate approximations $\mathbf{F}_{B,m}(s)$ for $\mathbf{F}_B(s)$ and $\mathbf{F}_{C,m}^{\mathrm{T}}(s)$ for $\mathbf{F}_C^{\mathrm{T}}(s)$ in the sense of the satisfaction of the Petrov–Galerkin orthogonality condition (3.18). In this case, the following reduced model can be obtained:

$$\Sigma_r : \begin{cases} \dot{\mathbf{x}}_r(t) = \mathbf{A}_r\mathbf{x}_r(t) + \mathbf{B}_r\mathbf{u}(t) \\ \mathbf{y}_r(t) = \mathbf{C}_r\mathbf{x}_r(t) + \mathbf{D}_r\mathbf{u}(t) \end{cases} \quad \Leftrightarrow \quad \Sigma_r = \left(\begin{array}{c|c} \mathbf{A}_r & \mathbf{B}_r \\ \hline \mathbf{C}_r & \mathbf{D}_r \end{array}\right), \tag{4.5}$$

where

$$\mathbf{A}_r = \mathbf{W}^{\mathrm{T}}\mathbf{A}\mathbf{V}, \quad \mathbf{B}_r = \mathbf{W}^{\mathrm{T}}\mathbf{B}, \quad \mathbf{C}_r = \mathbf{C}\mathbf{V}, \quad \mathbf{D}_r = \mathbf{D} \tag{4.6}$$

and

$$\mathbf{W}^{\mathrm{T}}\mathbf{V} = \mathbf{I}_m.$$

Without loss of generality, we assume in what follows that $\mathbf{D} = 0$ (which is typical of the majority of practical problems).

Let us introduce the following definition [21].

Definition 2. Let a transfer function of the original LTI system be given as

$$\mathbf{F}(s) = \mathbf{C}\left(s\mathbf{I}_n - \mathbf{A}\right)^{-1}\mathbf{B}, \tag{4.7}$$

and let its transfer function $\mathbf{G}(s)$ be expanded into the Laurent series in a neighborhood of a given point $\sigma \in \mathbb{C}$:

$$\begin{aligned} &\mathbf{F}(s) = \mathbf{C}\left(s\mathbf{I}_n - \mathbf{A}\right)^{-1}\mathbf{B} = \mathbf{C}\left(s\mathbf{I}_n + \sigma\mathbf{I}_n - \sigma\mathbf{I}_n - \mathbf{A}\right)^{-1}\mathbf{B} \\ &= \mathbf{C}(\mathbf{I}_n + \left(\sigma - s\right)\left(\sigma\mathbf{I}_n - \mathbf{A}\right)^{-1})^{-1}\left(\sigma\mathbf{I}_n - \mathbf{A}\right)^{-1}\mathbf{B} = m_0 + m_1\left(\sigma - s\right) \\ &+ m_2\left(\sigma - s\right)^2 + m_3\left(\sigma - s\right)^3 + \cdots = \sum_{i=0}^{\infty} m_i\left(\sigma - s\right)^i. \end{aligned} \tag{4.8}$$

The coefficients m_i are called moments of the transfer function of the system at point σ. It can be shown that the moments m_i are equal to the values of the transfer function $\mathbf{F}(s)$ and its derivatives at the point $\sigma \in \mathbb{C}$. The moments m_{r_i} of the transfer function of the reduced system are introduced in a similar way.

The basic idea of the model reduction is to establish correspondence between the moments m_i of the transfer function of the original system and the moments m_{r_i} of the transfer function of the reduced system for a given $i = \overline{1,l}$, where $l \ll n$, i.e.,

$$m_i = m_{r_i}.$$

Let us consider several typical cases.

Table

Approximation	Expansion of the transfer function	Moment
Partial implementation $\sigma = \infty$	$\sum_{i=1}^{\infty} m_{-i} s^{-i}$	$CA^{i-1}B$
Pade approximation $\sigma = 0$	$\sum_{i=1}^{\infty} m_{i-1} s^{i-1}$	$-CA^{-i}$
Biased Pade approximation $\sigma = s$	$\sum_{i=1}^{\infty} m_{i-1} (s-\sigma)^{i-1}$	$-C(\sigma I_n - A)^{-i}B$

Let $\sigma = \infty$. Then, the moments m_i are called *Markov's parameters*. These parameters may be viewed as coordinates of a point in the n-dimensional space that represents the transfer function. Then, instead of (4.8), we can write

$$\mathbf{F}(s) = \mathbf{CB}s^{-1} + \mathbf{CAB}s^{-2} + \mathbf{CA}^2\mathbf{B}s^{-3} + \cdots + \mathbf{CA}^{i-1}\mathbf{B}s^{-i} + \cdots = \sum_{i=1}^{\infty} \mathbf{CA}^{i-1}\mathbf{B}s^{-i}, \tag{4.9}$$

where

$$m_i = \mathbf{CA}^{i-1}\mathbf{B}. \tag{4.10}$$

Further, let $\sigma = 0$. Then, if matrix $\mathbf{A}$ is nonsingular, we have the following expansion:

$$\mathbf{F}(s) = -\mathbf{CA}^{-1}\mathbf{B} - \mathbf{CA}^{-2}\mathbf{B}^{s^{-1}} - \mathbf{CA}^2\mathbf{B}s^{-3} - \cdots - \mathbf{CA}^{-i}\mathbf{B}s^{-i+1} + \cdots = -\sum_{i=1}^{\infty} \mathbf{CA}^{-i}\mathbf{B}s^{-i+1}. \tag{4.11}$$

In this case, the moments of the transfer function are given by

$$m_i = -\mathbf{CA}^{-i}\mathbf{B}. \tag{4.12}$$

In a more general case, when expansions of the form

$$\mathbf{F}(s) = \sum_{i=1}^{\infty} m_{i-1} (s-\sigma)^{i-1} \tag{4.13}$$

are considered, the moments of the transfer function are given by

$$m_i = -\mathbf{C}(\sigma \mathbf{I}_n - \mathbf{A})^{-i}\mathbf{B}. \tag{4.14}$$

Possible results of the approximation are presented in the table.

Consider a direct procedure for constructing a reduced model assuming that it is required to establish correspondence between $2r$ moments of the transfer function, where r is the order of the reduced system. In other words, the frequency characteristic of the reduced system must coincide with that of the original system at r points.

Based on expansion (4.8), we introduce the following two Hankel matrices:

$$\mathcal{H}_k = \begin{pmatrix} m_1 & m_2 & \cdots & m_k \\ m_2 & m_3 & \cdots & m_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_k & m_{k+1} & \cdots & m_{2k-1} \end{pmatrix}, \quad \sigma\mathcal{H}_k = \begin{pmatrix} m_2 & m_3 & \cdots & m_{k+1} \\ m_3 & m_4 & \cdots & m_{k+2} \\ \vdots & \vdots & \ddots & \vdots \\ m_{k+1} & m_{k+2} & \cdots & m_{2k} \end{pmatrix}. \tag{4.15}$$

It can be shown that these matrices satisfy the identities

$$\mathcal{H}_k = \mathcal{O}_k(\mathbf{A}, \mathbf{B}, \sigma)\mathcal{R}_k(\mathbf{A}, \mathbf{B}, \sigma) = \mathcal{O}_k\mathcal{R}_k, \quad \sigma\mathcal{H}_k = \mathcal{O}_k\mathbf{A}\mathcal{R}_k, \tag{4.16}$$

where matrices $\mathcal{H}_k$ and $\mathcal{O}_k$ are defined in (2.9) and (2.10).

To compute matrices in the reduced model representation in the state space, one can apply the Lanczos [22] and Arnoldi [23] methods.

Let us consider these methods in more detail.

L a n c z o s' m e t h o d. Let us assume that

$$\det \mathcal{H}_k \neq 0, \quad i = 1, \ldots, k$$

and apply the LU decomposition to the Hankel matrix $\mathcal{H}_k$:

$$\begin{aligned} &\mathcal{H}_k = \mathbf{L}\mathbf{U}, \\ &\forall\ i < j: \quad \mathbf{L}(i, j) = 0, \\ &\forall\ i > j: \quad \mathbf{U}(i, j) = 0, \end{aligned} \tag{4.17}$$

where matrices $\mathbf{L}$ and $\mathbf{U}$ satisfy the condition

$$\mathbf{L}(i,i) = \pm\mathbf{U}(i,i).$$

Let us introduce the mappings

$$\mathbf{W}^{\mathrm{T}} = \mathbf{L}^{-1}\mathcal{O}_k, \tag{4.18}$$

$$\mathbf{V} = \mathcal{R}_k\mathbf{U}^{-1}, \tag{4.19}$$

assuming that $\mathbf{W}^{\mathrm{T}}\mathbf{V} = \mathbf{I}$. Then, it follows that $\mathbf{V}\mathbf{W}^{\mathrm{T}}$ is an oblique projector. Hence, we can define the reduced system Σ_r in the form

$$\mathbf{A}_r = \mathbf{W}^{\mathrm{T}}\mathbf{A}\mathbf{V}, \quad \mathbf{B}_r = \mathbf{W}^{\mathrm{T}}\mathbf{B}, \quad \mathbf{C}_r = \mathbf{C}\mathbf{V}. \tag{4.20}$$

The following theorem from [22] is valid.

T h e o r e m 1 (L a n c z o s' p r o c e d u r e). For a given full-order model Σ of form (4.7), the reduced-order model Σ_k satisfying the correspondence conditions for $2k$ Markov's parameters

$$\mathbf{C}\mathbf{A}^{i-1}\mathbf{B} = \mathbf{C}_k\mathbf{A}_k^{i-1}\mathbf{B}_k, \quad i = 1, \ldots, 2k,$$

corresponds to the triple $(\mathbf{B}_k, \mathbf{A}_k, \mathbf{C}_k)$, where $\mathbf{A}_k$ is a tridiagonal matrix and multipliers $\mathbf{B}_k$ and $\mathbf{C}_k^{\mathrm{T}}$ coincide with the unit vector $\mathbf{e}_1$ the first entry of which is one and the others are zeros.

Bellow, we consider a nonsymmetrical algorithm from [21] implementing Lanczos' procedure. Originally, this algorithm was designed for computing eigenvalues of symmetric and nonsymmetrical matrices. This algorithm can also be used for constructing controlled and observed subspaces for linear systems [24].

Consider model (1.1) under the condition $\mathbf{D} = 0$. Suppose also that $\mathbf{A}$ is a sparse matrix of great size. The algorithm below constructs matrices $\mathbf{W}_k$ and $\mathbf{V}_k$ in (4.20). In this case, the following relations hold:

$$\mathbf{V}_m = \left(\mathbf{v}_1 \middle| \mathbf{v}_2 \middle| \mathbf{v}_3 \middle| \cdots \middle| \mathbf{v}_m\right), \tag{4.21}$$

$$\mathbf{W}_m = \left(\mathbf{w}_1 \middle| \mathbf{w}_2 \middle| \mathbf{w}_3 \middle| \cdots \middle| \mathbf{w}_m\right), \tag{4.22}$$

$$\mathbf{A}\mathbf{V}_m = \mathbf{V}_m\mathbf{A}_m + \tilde{\mathbf{v}}_{m+1}\tilde{\mathbf{a}}_{mV}^{\mathrm{T}}, \quad \mathbf{B} = \mathbf{V}_m\mathbf{B}_m, \tag{4.23}$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{W}_m = \mathbf{W}_m\mathbf{A}_m^{\mathrm{T}} + \tilde{\mathbf{w}}_{m+1}\tilde{\mathbf{a}}_{mW}^{\mathrm{T}}, \quad \mathbf{C} = \mathbf{W}_m\mathbf{C}_m, \tag{4.24}$$

where

$$\mathbf{A}_m = \begin{pmatrix} \alpha_1 & \beta_2 & & 0 \\ \delta_2 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_m \\ 0 & & \delta_m & \alpha_m \end{pmatrix} \in \mathbb{R}^{m\times m}, \tag{4.25}$$

$$\tilde{\mathbf{a}}_{mV}^{\mathrm{T}} = \delta_{m+1}\mathbf{e}_m^{\mathrm{T}}, \quad \mathbf{B}_m = \mathbf{e}_1\delta_1, \tag{4.26}$$

$$\tilde{\mathbf{a}}_{mW}^{\mathrm{T}} = \beta_{m+1}\mathbf{e}_m^{\mathrm{T}}, \quad \mathbf{C}_m = \mathbf{e}_1\beta_1, \tag{4.27}$$

$\mathbf{e}_m$ is the unit vector with the mth entry equal to one and the other entries equal to zero; and $\alpha_j, \beta_j, \delta_j$ are entries of matrix $\mathbf{A}_m$ that are computed as specified below.

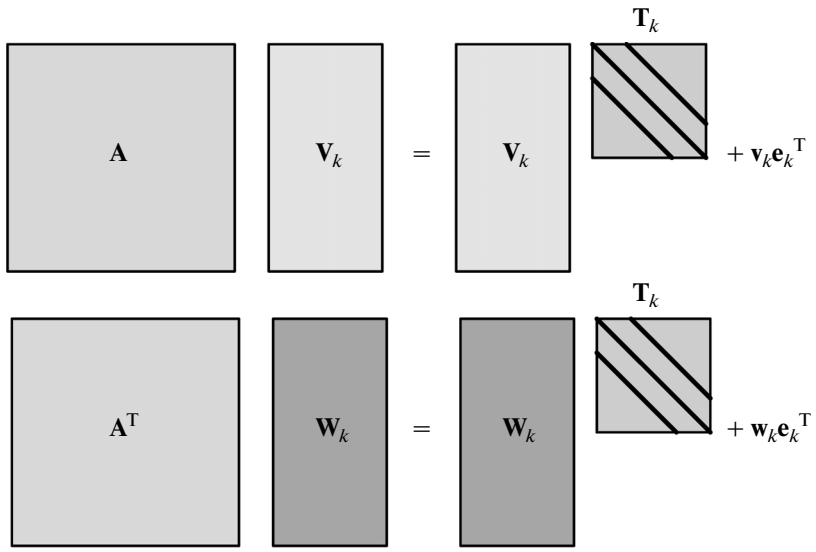

Fig. 1. The *k*th step of the Lanczos algorithm.

N o n s y m m e t r i c a l L a n c z o s' a l g o r i t h m:

1. $\beta_1 = \sqrt{\mathbf{C}^{\mathrm{T}}\mathbf{B}}$; $\delta_1 = \beta_1 \cdot \mathrm{sign}(\mathbf{C}^{\mathrm{T}}\mathbf{B})$;
2. $\mathbf{v}_0 = 0$; $\mathbf{w}_0 = 0$; $\mathbf{v}_1 = \mathbf{B}/\delta_1$; $\mathbf{w}_1 = \mathbf{C}/\beta_1$;
3. **for** $j = 1$ **to** m:

$$\alpha_j = \mathbf{w}_j^{\mathrm{T}}\mathbf{A}\mathbf{v}_j;$$

$$\hat{\mathbf{v}}_{j+1} = \mathbf{A}\mathbf{v}_j - \alpha_j\mathbf{v}_j - \beta_j\mathbf{v}_{j-1};\ \hat{\mathbf{v}}_{j+1} = \hat{\mathbf{v}}_{j+1} - \mathbf{V}_j(\mathbf{W}_j^{\mathrm{T}}\hat{\mathbf{v}}_{j+1});$$

$$\hat{\mathbf{w}}_{j+1} = \mathbf{A}^{\mathrm{T}}\mathbf{w}_j - \alpha_j\mathbf{w}_j - \delta_j\mathbf{w}_{j-1};\ \hat{\mathbf{w}}_{j+1} = \hat{\mathbf{w}}_{j+1} - \mathbf{W}_j(\mathbf{V}_j^{\mathrm{T}}\hat{\mathbf{w}}_{j+1});$$

$$\beta_{j+1} = \sqrt{\hat{\mathbf{w}}_{j+1}^{\mathrm{T}}\hat{\mathbf{v}}_{j+1}};$$

if $\beta_{j+1} \le \varepsilon$:

stop;

end;

4. $\delta_{j+1} = \beta_{j+1} \cdot \mathrm{sign}(\hat{\mathbf{w}}_{j+1}^{\mathrm{T}}\hat{\mathbf{v}}_{j+1})$;
5. $\mathbf{v}_{j+1} = \hat{\mathbf{v}}_{j+1}/\delta_{j+1}$; $\mathbf{w}_{j+1} = \hat{\mathbf{w}}_{j+1}/\beta_{j+1}$.

Functioning of the nonsymmetrical Lanczos' algorithm on the *k*th step is shown schematically in Fig. 1.

A r n o l d i m e t h o d. The Arnoldi method has been known since 1950. However, until 1970, it was seldom used as a computational method. The essence of the Arnoldi method can be explained as follows [25].

Let the reachability matrix Σ for system (1.1) have form (2.9). Then, the identity

$$\mathbf{A}\mathfrak{R}_k = \mathfrak{R}_k\mathbf{E} \tag{4.28}$$

holds, where

$$\mathbf{E} = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \cdots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\alpha_{k-1} \end{pmatrix} \tag{4.29}$$

is the companion Frobenius matrix [26] and

$$\det(s\mathbf{I}_n - \mathbf{A}) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s + \alpha_0.$$

Let us apply the QR decomposition [19]

$$\mathfrak{R}_k = \mathbf{W}\mathbf{R}_A \tag{4.30}$$

to the matrix $\mathfrak{R}_k$, where $\mathbf{W}$ is an orthogonal matrix and $\mathbf{R}_A$ is an upper-triangular nonsingular matrix.

Using (4.28), (4.29), we can write the following chain of assertions:

$$\mathbf{A}\mathbf{W}\mathbf{R}_A = \mathbf{W}\mathbf{R}_A\mathbf{E} \;\Rightarrow\; \mathbf{A}\mathbf{W} = \mathbf{W}\mathbf{R}_A\mathbf{E}\mathbf{R}_A^{-1} \;\Rightarrow\; \mathbf{A}\mathbf{W} = \mathbf{W}\overline{\mathbf{E}}, \tag{4.31}$$

where

$$\overline{\mathbf{E}} = \mathbf{R}_A\mathbf{E}\mathbf{R}_A^{-1}. \tag{4.32}$$

Since $\mathbf{R}_A$ is an upper triangular matrix, $\mathbf{R}_A^{-1}$ is also an upper triangular matrix. Note that (4.32) is an upper Hessenberg matrix.

Recall that a matrix $\mathbf{M}$ is said to have an upper Hessenberg form if it has the following block triangular from [19]:

$$\mathbf{M} = \left(\begin{array}{c|c|c|c} M_{11} & M_{12} & \cdots & M_{1l} \\ \hline 0 & M_{22} & \cdots & M_{2l} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & M_{ll} \end{array}\right),$$

where M_{ii} are square blocks of order 1 or 2.

On the basis of the QR decomposition (4.30), with regard to the condition

$$\mathscr{H}_k = \mathbb{O}_k\mathfrak{R}_k, \tag{4.33}$$

we define the mapping

$$\mathbf{V} = \mathfrak{R}_k\mathbf{R}_A^{-1} \in \mathbb{R}^{n\times k}, \tag{4.34}$$

where

$$\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}. \tag{4.35}$$

Since (4.35) is an orthogonality condition, we have an orthogonal projection of the form

$$\mathbf{A}_r = \mathbf{V}^{\mathrm{T}}\mathbf{A}\mathbf{V}, \quad \mathbf{B}_r = \mathbf{V}^{\mathrm{T}}\mathbf{B}, \quad \mathbf{C}_r = \mathbf{C}\mathbf{V}. \tag{4.36}$$

The following theorem from [21] is valid.

Theorem 2 (the Arnoldi procedure). For a given full-order model Σ of form (4.7), the reduced-order model Σ_k satisfying the correspondence conditions for $2k$ Markov's parameters

$$\mathbf{C}\mathbf{A}^{i-1}\mathbf{B} = \mathbf{C}_k\mathbf{A}_k^{i-1}\boldsymbol{B}_k, \quad i = 1, \ldots, 2k,$$

corresponds to the triple $(\mathbf{B}_k, \mathbf{A}_k, \mathbf{C}_k)$, where $\mathbf{A}_k$ is an upper Hessenberg matrix and multipliers $\mathbf{B}_k$ and $\mathbf{C}_k^{\mathrm{T}}$ coincide with the unit vector $\mathbf{e}_1$ the first entry of which is one and the others are zeros.

An algorithm based on the Arnoldi method (procedure) is described below [21]:

1. $\mathbf{v}_1 = \mathbf{B}/\|\mathbf{B}\|$;
2. **for** $j = 1$ **to** k
   for $i = 1$ **to** j:
   (a) Calculate $h_{ij} = \mathbf{v}_i^{\mathrm{T}}\mathbf{A}\mathbf{v}_j$;
   (b) $\mathbf{w}_j = \mathbf{A}\mathbf{v}_j - \sum_{i=1}^{j} h_{ij}\mathbf{w}_i$;
   (c) $h_{j+1,j} = \|\mathbf{w}_j\|_2$;
   (d) **if** $h_{j+1,j} = 0$ **stop**
   (e) $\mathbf{v}_{j+1} = \mathbf{w}_j/h_{j+1,j}$;
3. **end**.

The orthogonalization coefficients here are indexed by two subscripts, with regard to which the internal loop of the Arnoldi algorithm can be written by the following formula [25]:

$$h_{j+1,j}\mathbf{v}_{j+1} = \mathbf{A}\mathbf{v}_j - \sum_{i=1}^{k} h_{i,j}\mathbf{v}_i. \tag{4.37}$$

The orthogonalization coefficients $h_{i,j}$ on the kth step can be written in the form of the following matrix $\mathbf{H}_k$ if we replace the lacking positions by zeros:

$$\mathbf{H}_k = \begin{pmatrix} h_{11} & h_{12} & h_{13} & \cdots & h_{1k} \\ h_{21} & h_{22} & h_{23} & \cdots & h_{2k} \\ & \ddots & \ddots & \ddots & \vdots \\ & & h_{k-1,k-2} & h_{k-1,k-1} & h_{k-1,k} \\ & 0 & & h_{k,k-1} & h_{kk} \end{pmatrix} \in \mathbb{R}^{k\times k}. \tag{4.38}$$

For a given dimensionality k of the space, $k + 1$ vectors are generated. The last vector $\mathbf{v}_{k+1}$ (possibly, zero one) in the matrix notation means the extension of basis $\mathbf{V}$ by one additional column:

$$\mathbf{V}_k = (\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3\ \cdots\ \mathbf{v}_k),$$
$$\mathbf{V}_{k+1} = (\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3\ \cdots\ \mathbf{v}_k\ |\ \mathbf{v}_{k+1}).$$

The coefficient $h_{k+1,k}$ corresponding to vector $\mathbf{v}_{k+1}$ denotes the extension of matrix $\mathbf{H}$ by one additional row (possibly, zero row). Note that, if vectors

$$\mathbf{v}, \mathbf{A}\mathbf{v}, \mathbf{A}^2\mathbf{v}, \cdots, \mathbf{A}^{k-1}\mathbf{v} \tag{4.39}$$

are linearly dependent, then $h_{k+1,k} \neq 0$; otherwise, $h_{k+1,k} = 0$.

Let $\overline{\mathbf{H}}_k$ be a matrix of orthogonalization coefficients $h_{i,j}$ of size $(k + 1) \times k$ supplemented by the last row at the expense of $h_{k+1,k}$, and let $\mathbf{H}_k$ be the same matrix of size $k \times k$ without the last row. Then, from the description of the Arnoldi algorithm and from (4.37), it follows that $\mathbf{H}_k$ is actually an upper Hessenberg matrix and that it satisfies the equations

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\overline{\mathbf{H}}_k = \mathbf{V}_{k+1}\mathbf{H}_k + \mathbf{w}_k\mathbf{e}_k^{\mathrm{T}}, \tag{4.40}$$

$$\mathbf{V}_k^{\mathrm{T}}\mathbf{A}\mathbf{V}_k = \mathbf{H}_k, \tag{4.41}$$

$$\mathbf{V}^{\mathrm{T}}\mathbf{v}_k = \mathbf{e}_k. \tag{4.42}$$

As an example, we consider the approximation of eigenvalues of matrix $\mathbf{A}$ by means of the Arnoldi procedure [25]. Let $\mathbf{V}_k$, $\mathbf{H}_k$, and $h_{k+1,k}$ satisfy relations (4.40)–(4.42). Let μ be an eigenvalue of matrix $\mathbf{H}_k$ and $\mathbf{x}_H$ be the corresponding eigenvector satisfying the normalization condition $\|\mathbf{x}_H\|_2 = 1$. Let $\mathbf{v} = \mathbf{V}_k\mathbf{x}_H$ and $\|\mathbf{v}\|_2 = 1$. Then,

$$\|\mathbf{A}\mathbf{v} - \mu\mathbf{v}\|_2 = |h_{k+1,k}||x_{H,k}|, \tag{4.43}$$

where $x_{H,m}$ denotes the last (mth) component of vector $\mathbf{x}_H$. The less the orthogonalization coefficient $h_{k+1,k}$, the closer the eigenvalue μ of matrix $\mathbf{H}_k$ to the eigenvalue λ of matrix $\mathbf{A}$. Ideally, instead of (4.43), we have the equation

$$\|\mathbf{A}\mathbf{v} - \lambda\mathbf{v}\|_2 = 0,$$

i.e., solution of equation $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.

Functioning of the Arnoldi algorithm on the kth step is shown schematically in Fig. 2.

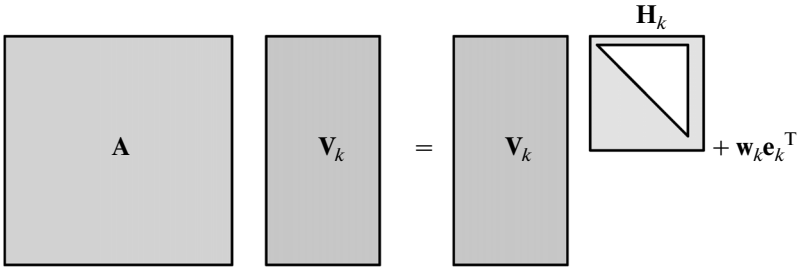

Fig. 2. The kth step of Arnoldi's algorithm.

5. MULTIPOINT RATIONAL INTERPOLATION

Let us return to the definition of moments of the transfer function of an LTI system (1.1) at point $s = \sigma$ and consider generalized reachability and observability matrices (2.9) and (2.10). The assertion below is applicable to SISO systems [21]. It allows one to construct projectors $\mathbf{V}$ and $\mathbf{W}$ ensuring efficient multipoint interpolation based on the generalized reachability and observability matrices.

Theorem 3. If

$$\bigcup_{k=1}^{K} \mathcal{K}_{b_k}\left((\sigma_k\mathbf{I}_n - \mathbf{A})^{-1}, (\sigma_k\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{B}\right) \subseteq \mathcal{V} = \mathrm{Im}(\mathbf{V}) \tag{5.1}$$

and

$$\bigcup_{k=1}^{K} \mathcal{K}_{c_k}((\sigma_k\mathbf{I}_n - \mathbf{A})^{-\mathrm{T}}, (\sigma_k\mathbf{I}_n - \mathbf{A})^{-\mathrm{T}}\mathbf{C}^{\mathrm{T}}) \subseteq \mathcal{W} = \mathrm{Im}(\mathbf{W}), \tag{5.2}$$

where $\mathbf{V}^{\mathrm{T}}\mathbf{W} = \mathbf{I}$ and σ_k is a number such that matrix $\sigma_k\mathbf{I}_n - \mathbf{A}$ is invertible for all $k \in \{\overline{1, K}\}$, then the moments of the LTI systems Σ and $\hat{\Sigma}$ satisfy the equalities

$$\eta_{\sigma_k}^{(j_k)} = \hat{\eta}_{\sigma_k}^{(j_k)}, \quad j_k = 0,1,2,\ldots, b_k + c_k - 1; \quad k = \overline{1, K}, \tag{5.2}$$

and matrices $\sigma_k\mathbf{I} - \mathbf{A}_r$ are invertible.

Based on Theorem 3, the so-called *dual* Arnoldi algorithm can be formulated [27]:

1. initial parameters: $m = 0$, empty matrices $\mathbf{V} = [\,]$, $\boldsymbol{Z} = [\,]$;
2. **for** $j = 1$ **to** J:
   (a): **for** $i = 1$ **to** K:
   (i) **if** $j = 1$,
   $\tilde{\mathbf{v}}_m = (\mathbf{A} - \sigma_k\mathbf{I}_n)^{-1}\mathbf{B}$, $\tilde{\mathbf{z}}_m = (\mathbf{A} - \sigma_k\mathbf{I}_n)^{-\mathrm{T}}\mathbf{C}^{\mathrm{T}}$;
   else,
   $\tilde{\mathbf{v}}_m = (\mathbf{A} - \sigma_k\mathbf{I}_\mathbf{n})^{-1}\tilde{\mathbf{v}}_{m-k}$, $\tilde{\mathbf{z}}_m = (\mathbf{A} - \sigma_k\mathbf{I}_n)^{-\mathrm{T}}\tilde{\mathbf{z}}_{m-k}$;
   end;
   (ii) $\hat{\mathbf{v}}_m = \tilde{\mathbf{v}}_m - \mathbf{V}\mathbf{V}^{\mathrm{T}}\tilde{\mathbf{v}}_m$, $\hat{\mathbf{z}}_m = \tilde{\mathbf{z}}_{m-k} - \mathbf{Z}\mathbf{Z}^{\mathrm{T}}\tilde{\mathbf{z}}_{m-k}$;
   (iii) $\mathbf{V} = [\mathbf{V}\,|\,\hat{\mathbf{v}}_m/\|\hat{\mathbf{v}}_m\|]$, $\mathbf{Z} = [\mathbf{Z}\,|\,\hat{\mathbf{z}}_m/\|\hat{\mathbf{z}}_m\|]$;
   (iv) $m = m + 1$;
3. $\mathbf{Z} \leftarrow \mathbf{Z} = \mathbf{Z}(\mathbf{V}^{\mathrm{T}}\mathbf{Z})^{-1}$.

In the case of a SISO system, the dual Arnoldi algorithm ensures full rank of matrices $\mathbf{V}$ and $\mathbf{W}$ upon minimal implementation in the system state space (in this case, generalized reachability (2.9) and observability (2.10) matrices have full rank by themselves, which guarantees full rank of matrices $\mathbf{V}$ and $\mathbf{W}$). However, for MIMO systems, trivial readjustment of the Arnoldi algorithm consisting in just replacement of the corresponding matrices does not lead to the desired result. The basic difficulty here consists in

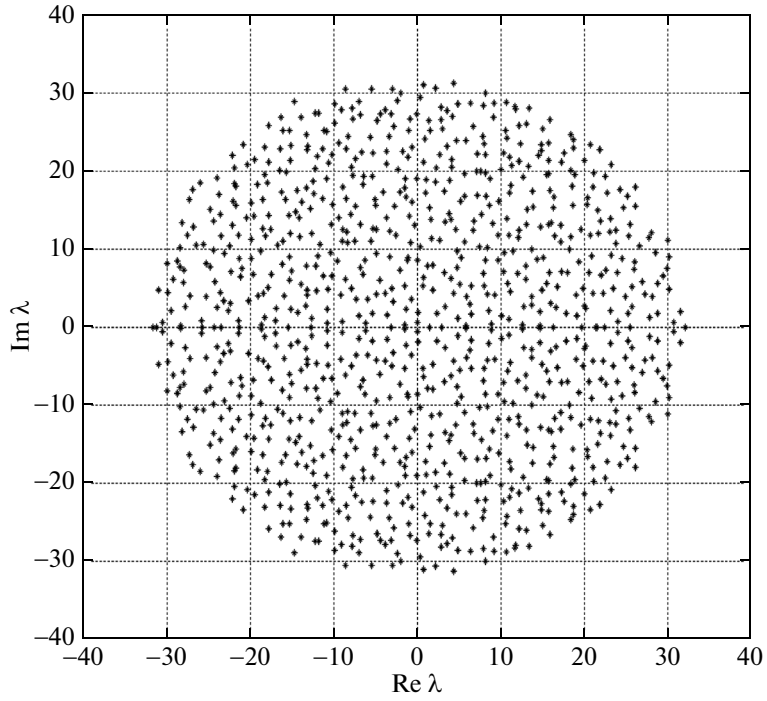

Fig. 3. Eigenvalues of large-scale matrix **A** (6.1).

impossibility (in the general case) to construct systems of linearly independent vectors (bases) associated with the Krylov subspaces $\mathcal{K}_m$. This problem can be solved by applying the following procedure:

"basis generation → completeness check → exclusion of linearly dependent vectors (*deflating vectors*) → generation of new vectors → ... ."

Extensions of the Lanczos and Krylov algorithms to MIMO systems are well studied [28]. To exclude linearly dependent vectors, the QR factorization is used.

Consider an LTI system of form (1.1) with K interpolation points σ_i of multiplicity b_i, where $i \in \{\overline{1, K}\}$. Let us also introduce the following definitions:

$$\text{if} \quad \sigma_i \neq \infty, \quad \text{then} \quad \mathbf{F}_i := (\sigma_i \mathbf{I}_n - \mathbf{A})^{-1}, \quad \mathbf{G}_i := (\sigma_i \mathbf{I}_n - \mathbf{A})^{-1}\mathbf{B},$$
$$\text{if} \quad \sigma_i = \infty, \quad \text{then} \quad \mathbf{F}_i := \mathbf{A}, \quad \mathbf{G}_i := \mathbf{B},$$

where

$$\mathbf{G}_i = (\mathbf{g}_{i_1} | \mathbf{g}_{i_2} | \cdots | \mathbf{g}_{i_m}), \quad \mathbf{V}_i = \mathcal{K}_{b_i}(\mathbf{F}_i, \mathbf{G}_i), \quad \mathbf{V} = (\mathbf{V}_1 | \mathbf{V}_2 | \cdots | \mathbf{V}_K).$$

Then, the Arnoldi algorithm presented in Section 4 can be used for calculating an orthogonal full-rank matrix $\mathbf{V}$ such that

$$\mathrm{Im}(\mathbf{V}) = \mathrm{Im}(\mathbf{V}_1 | \mathbf{V}_2 | \cdots | \mathbf{V}_K).$$

This is true because all vectors $\mathbf{V}_i$ satisfy the equation

$$\mathrm{Im}(\mathbf{V}_i) = \mathcal{K}_{b_i}(\mathbf{F}_i, \mathbf{G}_i).$$

The fulfillment of condition (5.1) is guaranteed in this case; hence, $\mathbf{V}$ is an appropriate matrix. The above-specified algorithm is called vector Arnoldi's algorithm with exclusion [29].

Below is the block Arnoldi algorithm, which can also be used for reduction of a MIMO system [21].

1. **for** $i = 1$ **to** K:

(a) **if** $\sigma_i \neq \infty$,

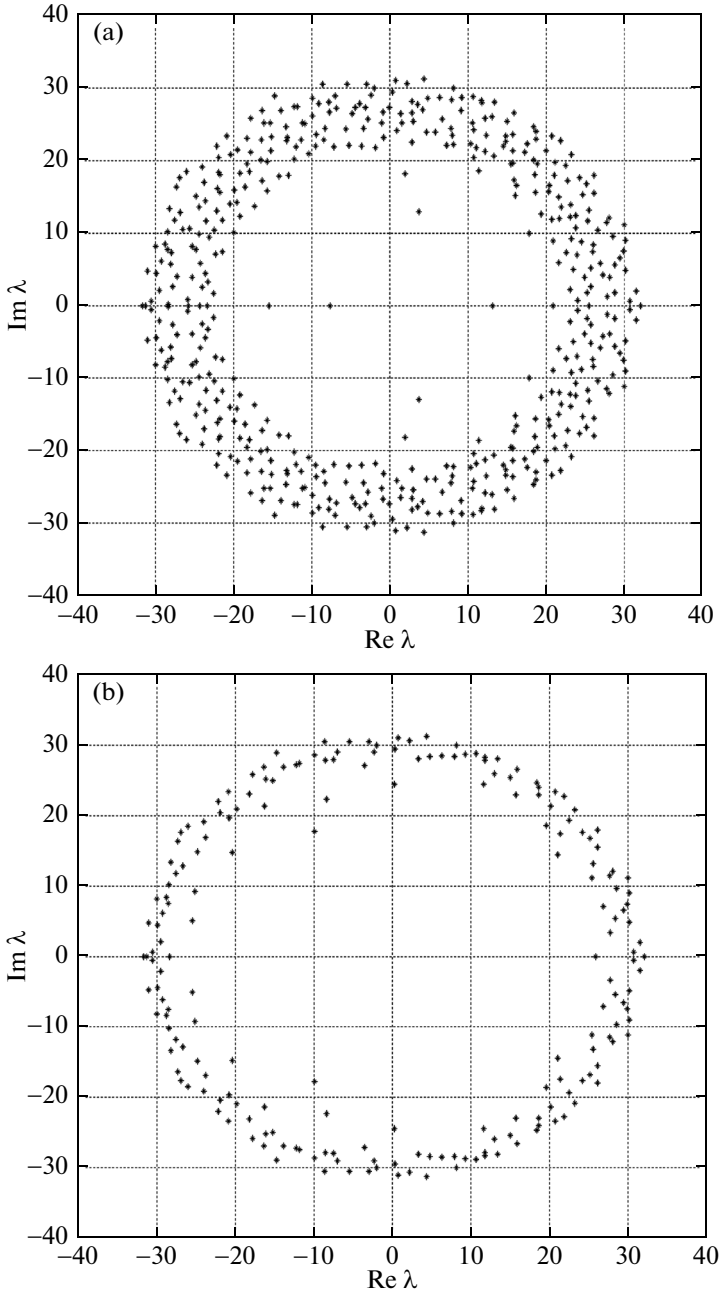

Fig. 4. (a) Approximation of eigenvalues of large-scale matrix **A** (6.1) of size $m \times m = 500 \times 500$, (b) approximation of eigenvalues of large-scale matrix **A** (6.1) of size $m \times m = 200 \times 200$, (c) approximation of eigenvalues of large-scale matrix **A** (6.1) of size $m \times m = 100 \times 100$, (d) approximation of eigenvalues of large-scale matrix **A** (6.1) of size $m \times m = 50 \times 50$.

$$\mathbf{F}_i = (\sigma_i \mathbf{I}_n - \mathbf{A})^{-1}, \mathbf{G}_i = (\sigma_i \mathbf{I}_n - \mathbf{A})^{-1} \mathbf{B};$$

else,

$$\mathbf{F}_i = \mathbf{A}, \mathbf{G}_i = \mathbf{B};$$

(b) $\mathbf{Q}_0\mathbf{R} = qr(\mathbf{G}_i)$ (QR factorization);

(c) **for** $k = 1$ **to** $b_k - 1$:

(i) $\mathbf{Q}_k = \mathbf{F}_i\mathbf{Q}_{k-1}$;

(ii) Gram–Schmidt orthogonalization of $\mathbf{Q}_k$ with respect to $\mathbf{Q}'_j$, $j = 1, \ldots, k-1$;

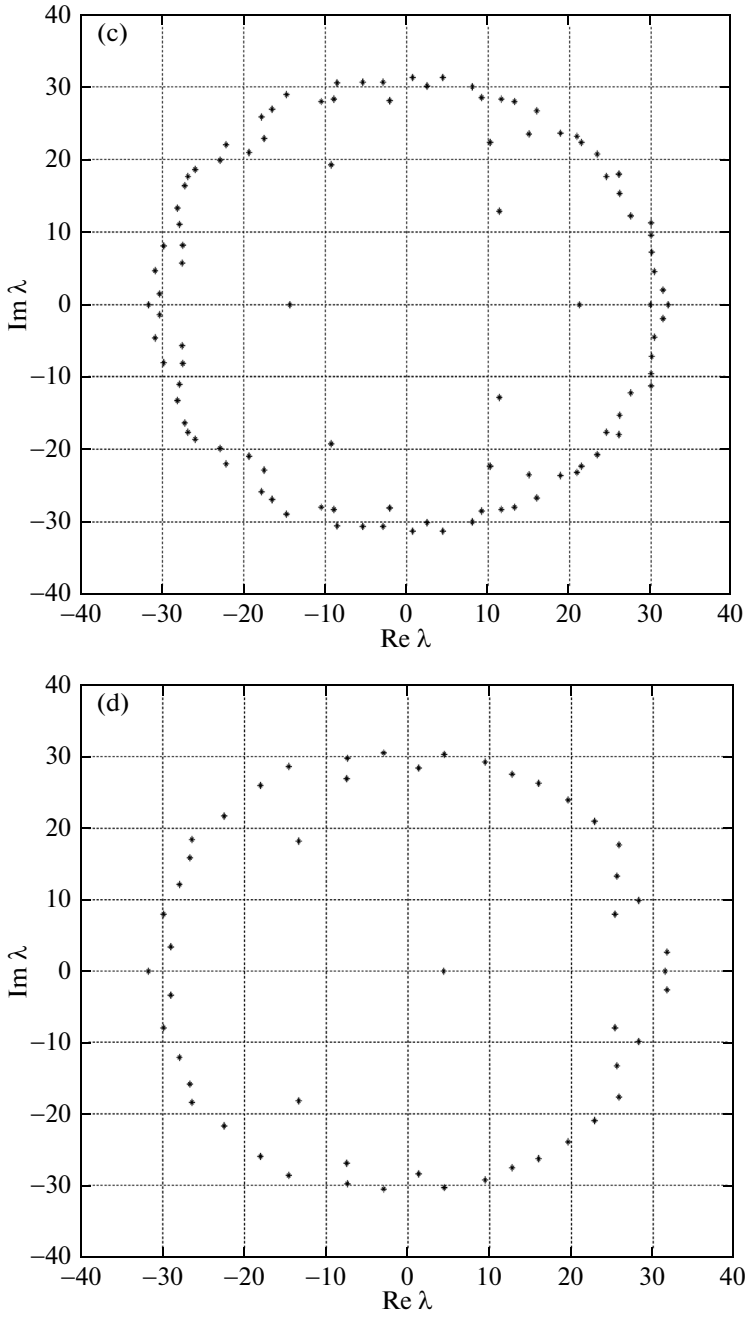

Fig. 4. (Contd.)

(iii) $\mathbf{Q}_k\mathbf{R} = qr(\mathbf{Q}_k)$;

(iv) $r_k = \mathrm{rank}(\mathbf{Q}_k)$;

(v) **if** $r_k \neq 0$,

$$\mathbf{V}_i = (\mathbf{V}_i | \mathbf{Q}_k);$$

else,

$k = b_k - 1$ (the kth block is filled);

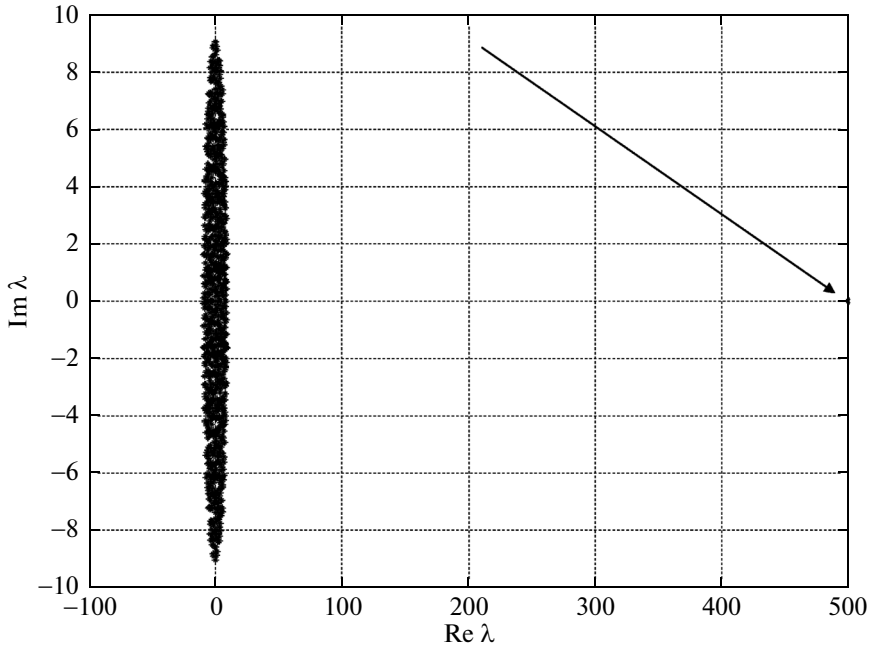

Fig. 5. Eigenvalues of large-scale matrix **A** (6.3).

2. **for** $i = 1$ **to** K (if $\mathbf{V}_i$ is linearly independent) perform Gram–Schmidt orthogonalization $\{\mathbf{V}_1 | \mathbf{V}_2 | \cdots | \mathbf{V}_K\}$;

3. matrix V found is a full-rank projector.

A similar computational process can be implemented by means of Lanczos' procedure; in this case, it is possible to estimate $\mathcal{H}_2$-norm of the approximation error by applying the following lemma [30].

L e m m a. Let a reduced Markov's chain Σ_r be obtained on the r th step of Lanczos' procedure. Then, the norm of the error $\Sigma_e = \Sigma - \Sigma_r$ is given by

$$\|\Sigma_e\|^2_{\mathcal{H}_2} = \sum_{i=1}^{n} \phi_i(\mathbf{G}(-\lambda_i^*) - \widehat{\mathbf{G}}(-\lambda_i^*)) + \sum_{j=1}^{r} \widehat{\phi}_j(\widehat{\mathbf{G}}(-\widehat{\lambda}_j^*) - \mathbf{G}(-\widehat{\lambda}_j^*)), \tag{5.3}$$

where

$$\phi_i = \mathbf{F}(s)(s - \lambda_i)\big|_{s=\lambda_i}, \quad i = 1, \ldots, n,$$

$$\widehat{\phi}_j = \widehat{\mathbf{F}}(s)(s - \widehat{\lambda}_j)\big|_{s=\lambda_j}, \quad j = 1, \ldots, r,$$

$\mathbf{F}(s)$ is the transfer matrix of the original system and $\widehat{\mathbf{F}}(s)$ is the transfer matrix of the reduced system.

6. NUMERICAL EXAMPLES

Consider reduction of a dynamical system with the large matrix

$$\mathbf{A} = \mathrm{randn}(1000 \times 1000). \tag{6.1}$$

The matrix is dense (all 10^6 entries of the matrix are nonzero), and its entries are uniformly distributed random numbers (obtained by the generator **randn**).

In the given case, vector **y** of the LTI system (1.1) coincides with the state vector **x** of dimension $n = 1000$, **C** is the identity matrix, and **B** is selected to be the column vector

$$\mathbf{B} = \mathrm{randn}(1000 \times 1). \tag{6.2}$$

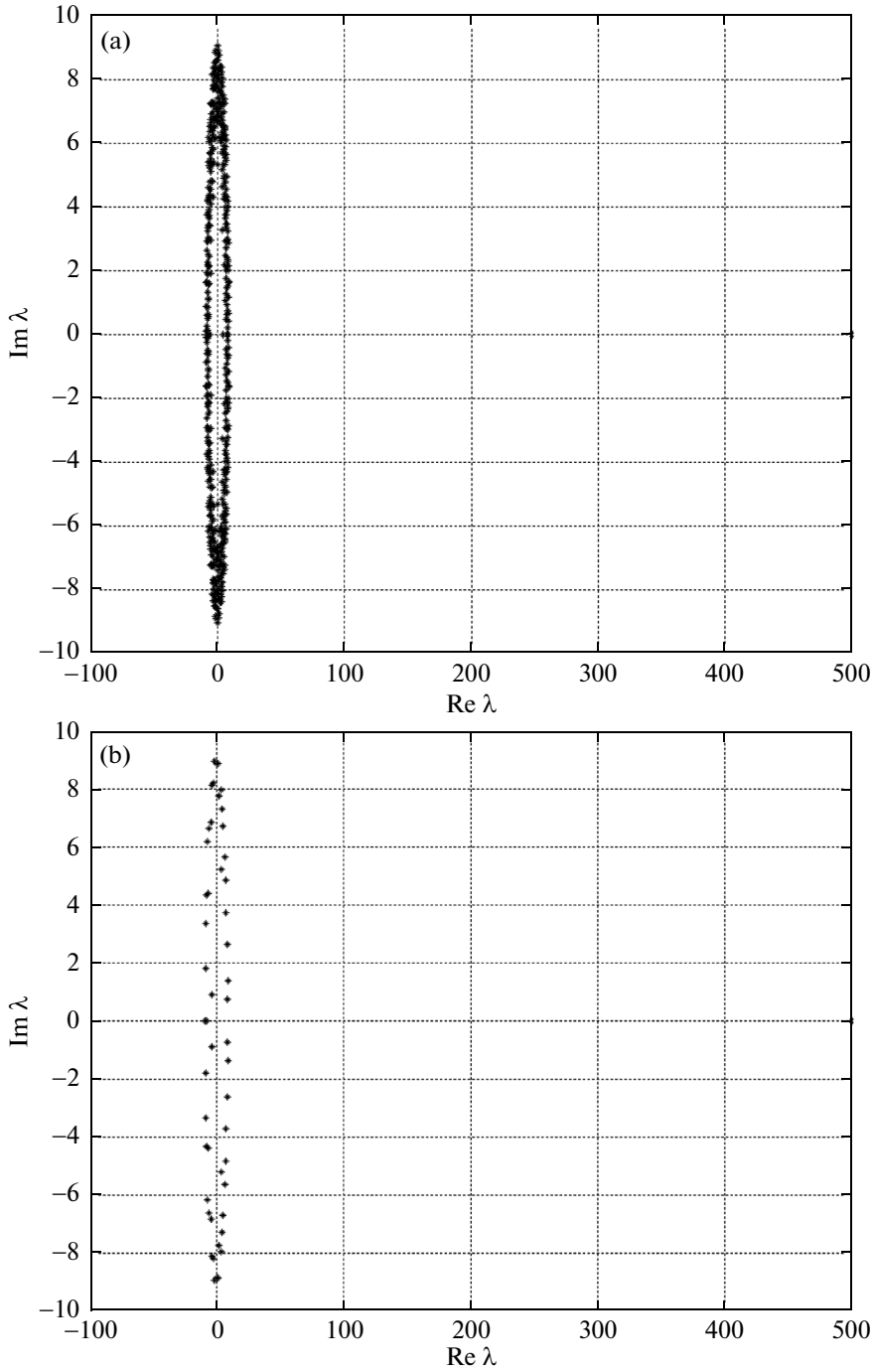

Fig. 6. (a) Approximation of eigenvalues of large-scale matrix **A** (6.3) of size $m \times m = 500 \times 500$, (b) approximation of eigenvalues of large-scale matrix **A** (6.1) of size $m \times m = 50 \times 50$.

The eigenvalues of this matrix obey the circle Girko law, and their distribution on the complex plane is shown in Fig. 3. The radius of the circle in which all eigenvalues of the matrix lie is equal to $\sqrt{n} = \sqrt{1000} \approx 31.6$.

The approximation (reduction) of the system with matrices (6.1) and (6.2) by the Krylov subspaces method implemented by the Arnoldi procedure yields the diagrams depicted in Fig. 4.

As can be seen from these diagrams, the reduced systems very accurately approximate the boundary of the localization of the eigenvalues of the original matrix. Figure 4d demonstrates that even 20-fold (!) reduction of the order of the system still provides acceptable accuracy.

Now consider system (1.1) with matrices

$$\mathbf{A} = \mathrm{rand}(1000 \times 1000), \quad \mathbf{B} = \mathrm{rand}(1000 \times 1). \tag{6.3}$$

Generator **rand** ensures strict positiveness of all entries of matrix **A**.

The eigenvalues of matrix (6.3) obey partially the circle Girko law, to be more specific, all, but one, eigenvalues of the matrix lie in the ellipse with the major semiaxis equal to $\sqrt{n}$, and one eigenvalue necessarily lies at infinity (Fig. 5).

Reducing the system with matrices (6.3) by means of the Arnoldi procedure, we obtain diagrams shown in Fig. 6.

This example demonstrates exceptional capability of the Krylov subspaces method to approximate "regular" and "abnormal" subsets of eigenvalues.

ACKNOWLEDGMENTS

This work was supported by the Russian Research Foundation (project no. 14-11-00046).

REFERENCES

1. A. N. Krylov, "On numerical solution of equation governing frequencies of small oscillations of material systems,'' Izv. Akad. Nauk SSSR, Ser. OMEN., No. 4, 491–539 (1931).
2. Yu. A. Andreev, *Control of Finite-Dimensional Linear Objects* (Nauka, Moscow, 1976) [in Russian].
3. A. A. Voronov, *Introduction to Dynamics of Complex Controlled Systems* (Nauka, Moscow, 1985) [in Russian].
4. M. Sh. Misrikhanov, *Invariant Control of Multidimensional Systems: An Algebraic Approach* (Nauka, Moscow, 2007) [in Russian].
5. M. Sh. Misrikhanov, *Classical and New Methods for Analysis of Dynamical Multidimensional Systems* (Energoatomizdat, Moscow, 2004) [in Russian].
6. M. Sh. Misrikhanov, "Band controllability and observability tests," Autom. Remote Control, No. 12, 1953–1963 (2005).
7. M. Sh. Misrikhanov and V. N. Ryabchenko, "Algebraic and matrix methods in the theory of linear MIMO systems," Vestn. IGEU, No. 5, 196–240 (2005).
8. M. Sh. Misrikhanov and V. N. Ryabchenko, "The quadratic eigenvalue problem in electric power systems," Autom. Remote Control, No. 5, 698–720 (2006).
9. M. Sh. Misrikhanov and V. N. Ryabchenko, "The band formula for A.N. Krylov's problem," No. 12, 2142–2157 (2007).
10. M. Sh. Misrikhanov and V. N. Ryabchenko, "Reduction of the Rosenbrock matrix in analysis of invariant zeros of the linear MIMO-system," Autom. Remote Control, No. 12, 1675–1691 (2008).
11. A. V. Bogachev, E. A. Vorob'eva, N. E. Zubov, E. A. Mikrin, M. Sh. Misrikhanov, V. N. Ryabchenko, and S. N. Timakov, "Unloading angular momentum for inertial actuators of a spacecraft in the pitch channel," J. Comput. Syst. Sci. Int. **50**, 483–490 (2011).
12. *Essentials of Robust Control*, Ed. by K. Zhou (Prentice Hall, Upper Saddle River, New Jersey, 1998).
13. T. Kailath, *Linear Systems* (Prentice Hall, Upper Saddle River, New Jersey, 1990).
14. N. E. Zubov, E. A. Mikrin, M. Sh. Misrikhanov, and V. N. Ryabchenko, "Solution of linear matrix equations and Lyapunov's inequalities by the Krylov subspaces method," Vestn. MGTU Im. N.E. Baumana, No. 2, 103–119 (2014).
15. *Handbook of Automatic Control Theory*, Ed. by A. A. Krasovskii (Nauka, Moscow, 1987) [in Russian].
16. G. H. Golub and C. F. van Loan, *Matrix Computations* (John Hopkins Univ. Press, 1989; Mir, Moscow, 1999).
17. M. Yu. Balandin and E. P. Shurina, *Methods for Solving Systems of Linear Algebraic Equations of High Dimension* (Izd-vo NGTU, Novosibirsk, 2000) [in Russian].
18. Y. Saad, *Iterative Methods for Sparse Linear Systems* (SIAM, Philadelphia, 2003).
19. V. V. Voevodin and Yu. A. Kuznetsov, *Matrices and Calculations* (Nauka, Moscow, 1984) [in Russian].
20. D. S. Bernstein, *Matrix Mathematics* (Princeton Univ. Press, Princeton, 2005).

21. A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems* (SIAM, Philadelphia, 2005).
22. C. Lanczos, "An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators," J. Res. Nat. Bur. Stand. **45**, 255–282 (1950).
23. W. E. Arnoldi, "The Principle of Minimized Iteration in the Solution of the Matrix Eigenvalue Problem," Quart. Appl. Math. **9**, 17–29 (1951).
24. D. Boley and G. Golub, "The Nonsymmetric Lanczos Algorithm and Controllability," Syst. Control Lett. **16**, 97–105 (1991).
25. D. S. Watkins, *Fundamentals of Matrix Computations* (Wiley, 2002; BINOM. Laboratoriya znanii, Moscow, 2006).
26. Gantmakher, F.R., *Theory of Matrices* (Nauka, Moscow, 1988).
27. E. Grimme, D. Sorensen, and P. van Dooren, "Model reduction of state-space systems via an implicitly restarted Lanczos method," Numerical Algorithms **12**, 1–31 (1995).
28. A. C. Antoulas, "A new result on positive real interpolation and model reduction," Systems Control Lett. **54**, 361–374 (2005).
29. R. W. Freund, "Model reduction methods based on Krylov subspaces," Acta Numerica **12**, 267–319 (2003).
30. S. Gugercin and A. C. Antoulas, "An H2 error expression for the Lanczos procedure," Proc. of the 42nd IEEE Conf. on Decision Procedure, 2003.

Translated by A. Pesterev