Chapter 1

Introduction

1.1 Background

Physical structures and processes are modeled by dynamical systems in a wide range of application areas. In structural analysis, for instance, one is interested in the natural frequency of a structure, that is, the frequency at which the system prefers to vibrate. The natural frequencies are of importance for the design and construction of bridges and large buildings, because precautions must be taken to prevent the structure from resonating with vibrations due to external forces such as pedestrian movements, wind, and earthquakes. Some electrical circuits, on the other hand, are designed to amplify or filter signals at specific frequencies. In chemical engineering, dynamical systems are used to describe heat transfer and convection of chemical processes and reactions. It is often cheaper and safer to simulate the system before actually realizing it, because simulation may give insight in the behavior of the system and may help to check if the design meets the requirements, and to adapt the design if needed.

Of paramount importance in all applications is that the models accurately enough describe the underlying systems. The increasing demand for complex components such as high-speed computer chips with more transistors and decreasing size, and structures such as large airplanes, together with an increasing demand for detail and accuracy, makes the models larger and more complicated. Although computer resources are also getting faster and bigger, direct simulation of the system is often not feasible because the required storage and computer time is proportional to the square and cube, respectively, of the number of elements of the model, which may easily exceed one million. To be able to simulate these large-scale systems, there is need for reduced-order models of much smaller size, that approximate the behavior of the original model and preserve the important characteristics, at least for the frequency or time range of interest.

One can see model order reduction as an application of Ockham's razor. William of Ockham, a medieval logician, stated that in the explanation of a phenomenon,

as few assumptions as possible should be made, and that the simplest explanation is to be preferred. In model order reduction, Ockham's razor is used to "shave off" the less important characteristics of a model. The same goal can be achieved by computing the important characteristics instead.

This thesis presents algorithms for the computation of dominant eigenvalues and corresponding eigenvectors of eigenproblems that are related to large-scale dynamical systems. Here, the interpretation of the adjective *dominant* depends on the application from which the eigenproblem arises. In stability analysis of steady state solutions of discretized Navier-Stokes equations, for example, the dominant eigenvalues are the rightmost eigenvalues, since eigenvalues with positive real part correspond to unstable steady state solutions. In electrical circuit simulation and other applications, one is also interested in the rightmost eigenvalues, that correspond to unstable or slowly damped modes. In structural analysis and engineering, the dominant eigenvalues are the natural frequencies. For general linear time invariant dynamical systems, the dominant eigenvalues are the poles of the transfer function that contribute significantly to the frequency response. In the design of stabilizers and controllers for large-scale systems, the dominant eigenvalues are the zeros of the transfer function that damp unwanted modes.

The dominant eigenvalues and corresponding eigenvectors give information about the behavior of the system or solution, and hence function as a reduced-order model in some sense. The dominant eigenvalues and corresponding eigenvectors, for instance, can be used to construct a reduced-order model for the original system. A large-scale dynamical system can be reduced to a much smaller reduced-order model, which is also a dynamical system, by projecting the state-space on the dominant eigenspace. In stability analysis, insight in the oscillatory or unstable behavior of steady state solutions is obtained by studying the modes corresponding to the rightmost eigenvalues. The construction of the reduced-order model varies with the interpretation of the dominant eigenvalues and the application, but in all cases the quality of the reduced-order model is determined by the degree in which it reflects the characteristic (dominant) behavior of the underlying system or solution. The algorithms presented in this thesis are specialized eigenvalue methods that compute dominant eigenvalues and corresponding eigenvectors for several types of applications.

The remainder of this chapter is organized as follows. Section 1.2 gives a brief introduction to eigenvalue problems. In Section 1.3, essential concepts and results from system and control theory are summarized. Section 1.5 describes methods for eigenvalue problems and the related model order reduction methods. An overview of the contributions of this thesis is given in Section 1.6.

1.2 Eigenvalue problems

This section briefly describes the various types of eigenvalue problems (eigenproblems) that are related to topics discussed in this thesis. For more details the reader is referred to, for instance, [8, 64].

1.2.1 The standard eigenvalue problem

The standard eigenvalue problem is to find $\lambda \in \mathbb{C}$ and $\mathbf{x} \in \mathbb{C}^n$ that satisfy

$$A\mathbf{x} = \lambda\mathbf{x}, \quad \mathbf{x} \neq 0,$$

where A is a complex $n \times n$ matrix. The scalar $\lambda \in \mathbb{C}$, that satisfies $\det(A-\lambda I) = 0$, is an eigenvalue, and the nonzero vector $\mathbf{x} \in \mathbb{C}^n$ is a (right) eigenvector for λ. The pair $(\lambda, \mathbf{x})$ is also referred to as an eigenpair of A. A nonzero vector $\mathbf{y} \in \mathbb{C}^n$ that satisfies $\mathbf{y}^*A = \lambda\mathbf{y}^*$ is a left eigenvector for λ, and $(\lambda, \mathbf{x}, \mathbf{y})$ is called an eigentriplet of A. The set of all eigenvalues of A is called the spectrum of A, denoted by $\Lambda(A)$. Symmetric matrices $A = A^T \in \mathbb{R}^{n\times n}$ have real eigenvalues and an orthonormal basis of real eigenvectors. Hermitian matrices $A = A^* \in \mathbb{C}^{n\times n}$ have real eigenvalues and an orthonormal basis of eigenvectors. If $A \in \mathbb{C}^{n\times n}$ is normal ($AA^* = A^*A$), then there also exists an orthonormal basis of eigenvectors. For symmetric, Hermitian and normal matrices, the right eigenvector for an eigenvalue is also a left eigenvector for the same eigenvalue. For a general matrix $A \in \mathbb{C}^{n\times n}$ there exists a Schur decomposition

$$Q^*AQ = T,$$

where $Q \in \mathbb{C}^{n\times n}$ is unitary ($Q^*Q = I$) and $T \in \mathbb{C}^{n\times n}$ is upper triangular with the eigenvalues of A on its diagonal. A general matrix A is diagonalizable (or nondefective) if and only if there exists a nonsingular $X \in \mathbb{C}^{n\times n}$ such that $X^{-1}AX = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, where λ_i $(i = 1, \ldots, n)$ are the eigenvalues of A and the columns of X are eigenvectors of A. If all λ_i are distinct, then there are n independent eigenvectors. A matrix that is not diagonalizable is called defective, and an eigenvalue λ with algebraic multiplicity (multiplicity of root λ of $\det(A - \lambda I)$) greater than its geometric multiplicity (dimension of the corresponding eigenspace) is called defective. For a general matrix A, there exists a Jordan decomposition $X^{-1}AX = \mathrm{diag}(J_1, \ldots, J_s)$, where the Jordan blocks J_i are $m_i \times m_i$ upper triangular matrices with its single eigenvalue λ_i on the diagonal and ones on the first superdiagonal, and all other elements zero. Each block J_i has a single independent left and right eigenvector, and $m_1 + \ldots + m_s = n$, and each block with $m_i > 1$ corresponds to a defective eigenvalue.

1.2.2 The generalized eigenvalue problem

The generalized eigenvalue problem is of the form

$$A\mathbf{x} = \lambda B\mathbf{x}, \quad \mathbf{x} \neq 0,$$

where A and B are $n \times n$ complex matrices, and reduces to the standard eigenvalue problem if $B = I$. An eigentriplet $(\lambda, \mathbf{x}, \mathbf{y})$ of the pair (A, B) (or pencil $A - \lambda B$) consists of an eigenvalue $\lambda \in \mathbb{C}$ that satisfies $\det(A - \lambda B) = 0$, a nonzero (right) eigenvector $\mathbf{x} \in \mathbb{C}^n$ that satisfies $A\mathbf{x} = \lambda B\mathbf{x}$, and a nonzero left eigenvector $\mathbf{y} \in \mathbb{C}^n$ that satisfies $\mathbf{y}^*A = \lambda\mathbf{y}^*B$. The set of all eigenvalues of (A, B) is called the

spectrum of (A, B), denoted by $\Lambda(A, B)$. If A and B are Hermitian and B is positive definite, then all eigenvalues of (A, B) are real and there exists a nonsingular $X \in \mathbb{C}^{n\times n}$ such that $X^*AX = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $X^*BX = I$, where λ_i $(i = 1, \ldots, n)$ are the eigenvalues of (A, B). The columns $\mathbf{x}_i$ of X are B-orthogonal ($\mathbf{x}_i^* B\mathbf{x}_j = 0$ if $i \neq j$), and are both right and left eigenvectors for the same eigenvalue λ_i. If in addition A and B are real, the eigenvectors can be chosen to be real as well. For a general matrix pair (A, B) there exists a generalized Schur decomposition

$$Q^*AZ = T, \quad Q^*BZ = S,$$

where $Q, Z \in \mathbb{C}^{n\times n}$ are unitary ($Q^*Q = Z^*Z = I$) and $S, T \in \mathbb{C}^{n\times n}$ are upper triangular with $t_{ii}/s_{ii} = \lambda_i$ for $s_{ii} \neq 0$. If $s_{ii} = 0$, then $\lambda_i = \infty$, and if both t_{ii} and s_{ii} are zero, then the spectrum $\Lambda(A, B) = \mathbb{C}$. A general matrix pair (A, B) is called diagonalizable (or nondefective) if there are n independent right eigenvectors $\mathbf{x}_i$ and n independent left eigenvectors $\mathbf{y}_i$: if $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, $X = [\mathbf{x}_1, \ldots, \mathbf{x}_n]$ and $Y = [\mathbf{y}_1, \ldots, \mathbf{y}_n]$, then it follows from $\mathbf{y}_i^* B\mathbf{x}_j = 0$ for $i \neq j$ that $Y^*AX = \Lambda_A$ and $Y^*BX = \Lambda_B$, with $\Lambda_A\Lambda_B^{-1} = \Lambda$ (provided Λ_B is nonsingular). If all λ_i are distinct, then there are n independent eigenvectors. Analogous to the Jordan decomposition for the standard eigenvalue problem, there are Weierstrass and Weierstrass-Schur decompositions [8, p. 30] for the generalized eigenvalue problem.

If A (B) is nonsingular, the generalized eigenproblem can be transformed to a standard eigenproblem by multiplying with A^{-1} (B^{-1}), but practically speaking this is often necessary nor advisable.

1.2.3 The polynomial eigenvalue problem

The polynomial eigenvalue problem is of the form

$$(\lambda^p A_p + \lambda^{p-1} A_{p-1} + \cdots \lambda A_1 + A_0)\mathbf{x}, \quad \mathbf{x} \neq 0,$$

where the A_i are $n \times n$ complex matrices, and is a generalization of the standard and generalized eigenvalue problem. Definitions of eigenvalues and left and right eigenvectors follow from the definitions for the generalized eigenvalue problem. Although there are some similarities between polynomial and standard/generalized eigenproblems, the major difference is that the polynomial eigenproblem has np eigenvalues with up to np left and np right eigenvectors, that, if there are more than n eigenvectors, are not independent. If $p = 2$, the polynomial eigenvalue problem is a quadratic eigenvalue problem. See [152] and references therein for more details on quadratic and polynomial eigenproblems.

1.2.4 The singular value problem

For an $m\times n$ matrix A, the singular value problem consists of finding $\sigma \in \mathbb{R}$ $(\sigma \geq 0)$, nonzero $\mathbf{u} \in \mathbb{C}^m$ and nonzero $\mathbf{v} \in \mathbb{C}^n$ with $\|\mathbf{u}\|_2 = \|\mathbf{v}\|_2 = 1$ that satisfy

$$\begin{aligned} A\mathbf{v} &= \sigma\mathbf{u}, \\ A^*\mathbf{u} &= \sigma\mathbf{v}. \end{aligned}$$

The singular value σ and corresponding left singular vector $\mathbf{u}$ and right singular vector $\mathbf{v}$ form a singular triplet $(\sigma, \mathbf{u}, \mathbf{v})$. The nonzero singular values are the square roots of the nonzero eigenvalues of AA^* or A^*A, and the left (right) singular vectors of A are the eigenvectors of AA^* (A^*A). Alternatively, the absolute values of the eigenvalues of

$$\begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}$$

are the singular values of A, and the left (right) singular vectors can be extracted from the first (second) part of the corresponding eigenvectors. The decomposition $A = U\Sigma V^*$, with U and V unitary, is called a singular value decomposition (SVD) of A if the columns of $U \in \mathbb{C}^{m\times m}$ are the left singular vectors, the columns of $V \in \mathbb{C}^{n\times n}$ are the right singular vectors, and $\Sigma \in \mathbb{C}^{m\times n}$ is a diagonal matrix with the singular values on its diagonal.

1.3 System and control theory

This section describes the concepts from system and control theory that are needed in this thesis. Most of the material is adapted from [27, 78, 118, 145], that give detailed introductions to system theory, and [5], that provides a good introduction to system theory from the view point of model order reduction and numerical linear algebra.

1.3.1 State space systems

The internal description of a linear time invariant (LTI) system $\Sigma = (A, B, C, D)$ is

$$\left\{ \begin{array}{rcl} \dot{\mathbf{x}}(t) & = & A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) & = & C^*\mathbf{x}(t) + D\mathbf{u}(t), \end{array} \right. \tag{1.3.1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{n\times p}$, $\mathbf{x}(t) \in \mathbb{R}^n$, $\mathbf{u}(t) \in \mathbb{R}^m$, $\mathbf{y}(t) \in \mathbb{R}^p$, and $D \in \mathbb{R}^{p\times m}$. The matrix A is called the state-space matrix, the matrices B and C are called the input and output map, respectively, and D is the direct transmission map; they are also referred to as system matrices. The vector $\mathbf{u}(t)$ is called the input or control, $\mathbf{x}(t)$ is called the state vector, and $\mathbf{y}(t)$ is called the output of the system. The order of the system is n. If $m, p > 1$, the system is called multi-input multi-output (MIMO). If $m = p = 1$ the system is called single-input single-output (SISO) and reduces to $\Sigma = (A, \mathbf{b}, \mathbf{c}, d)$:

$$\left\{ \begin{array}{rcl} \dot{\mathbf{x}}(t) & = & A\mathbf{x}(t) + \mathbf{b}u(t) \\ y(t) & = & \mathbf{c}^*\mathbf{x}(t) + du(t), \end{array} \right. \tag{1.3.2}$$

where $A \in \mathbb{R}^{n\times n}$, $\mathbf{b}, \mathbf{c}, \mathbf{x}(t) \in \mathbb{R}^n$, and $u(t), y(t), d \in \mathbb{R}$. In a more general setting the matrices A, B, C, D may also be time dependent, and the relations between input, state and output may be nonlinear. The dynamical systems (1.3.1) and (1.3.2) can come directly from linear models, or can be linearizations of nonlinear models.

The transfer function $H : \mathbb{C}^m \longrightarrow \mathbb{C}^p$ of (1.3.1),

$$H(s) = C^*(sI - A)^{-1}B + D, \tag{1.3.3}$$

can be obtained by applying the Laplace transform to (1.3.1) under the condition $\mathbf{x}(0) = 0$. The transfer function relates outputs to inputs in the frequency domain via $Y(s) = H(s)U(s)$, where $Y(s)$ and $U(s)$ are the Laplace transforms of $\mathbf{y}(t)$ and $\mathbf{u}(t)$, respectively, and is important in many engineering applications. It is well known that the transfer function is invariant under state-space transformations $\mathbf{x} \mapsto T\mathbf{x}$ where $T \in \mathbb{R}^{n\times n}$ is nonsingular. Consequently, there exist infinitely many realizations $(TAT^{-1}, TB, T^{-*}C, D)$ of the same LTI transfer function. There exist also realizations of order $\tilde{n}$ with $\tilde{n} > n$ (that are in general not of interest). The lower limit $\hat{n} < n$ on the order of the system is called the McMillan degree of the system and a realization of order $\hat{n}$ is called a minimal realization.

A pole of transfer function $H(s)$ is a $p \in \mathbb{C}$ for which $\lim_{s\to p} \|H(s)\|_2 = \infty$. The set of poles of $H(s)$ is a subset of the eigenvalues $\lambda_i \in \mathbb{C}$ of the state-space matrix A. A (transmission) zero of transfer function $H(s)$ is a $z \in \mathbb{C}$ for which $H(s)$ drops in rank if $s = z$; in the SISO case this means that $H(z) = 0$ for a zero $z \in \mathbb{C}$.

Let $(\lambda_i, \mathbf{x}_i, \mathbf{y}_i)$ $(i = 1, \ldots, n)$ be eigentriplets of A, with all λ_i distinct and the eigenvectors scaled so that $\mathbf{y}_i^*\mathbf{x}_i = 1$ and $\mathbf{y}_i^*\mathbf{x}_j = 0$ for $i \neq j$. The transfer function $H(s)$ can be expressed as a sum of residue matrices $R_j \in \mathbb{C}^{p\times m}$ over the poles [78]:

$$H(s) = D + \sum_{j=1}^{n} \frac{R_j}{s - \lambda_j}, \tag{1.3.4}$$

where

$$R_j = (C^*\mathbf{x}_j)(\mathbf{y}_j^*B).$$

If there are nondefective multiple eigenvalues, then the eigenvectors can be identified by their components in B and C. In the SISO case, for instance, decompose $\mathbf{b}$ as

$$\mathbf{b} = \sum_{i=1}^{k} \sum_{j=1}^{m_i} \beta_i^j \mathbf{x}_i^j,$$

where k is the number of distinct eigenvalues, m_i is the multiplicity of λ_i, β_i^j are coefficients, and $\mathbf{x}_i^j$ are eigenvectors. The right eigenvector $\mathbf{x}_i$ of a pole λ_i of multiplicity m_i is then identified by $\mathbf{x}_i = \sum_{i=1}^{m_i} \beta_i^j \mathbf{x}_i^j$. Note that the summation in (1.3.4) then consists of $k \leq n$ terms. Poles λ_j with large $\|R_j\|_2/|\mathrm{Re}(\lambda_j)|$ (or $\|R_j\|_2$) are called dominant poles and are studied in Chapters 2–6.

A system (A, B, C, D) is called (asymptotically) stable if all eigenvalues of A have strictly negative real parts (A is also called Hurwitz then). This can also be seen in the solution

$$\mathbf{y}(t) = C^*\mathbf{x}(t) + D\mathbf{u}(t), \quad \mathbf{x}(t) = e^{A(t-t_0)}\mathbf{x}_0 + \int_{t_0}^{t} e^{A(t-\tau)}B\mathbf{u}(\tau)d\tau \tag{1.3.5}$$

of system (1.3.1) with initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$: $\mathbf{y}(t)$ becomes unbounded if A has eigenvalues with positive real part. A square $(m = p)$ system is passive if and

only if its transfer function $H(s)$ is positive real, that is, $H(s)$ is analytic for all s with $\mathrm{Re}(s) > 0$, $H(\bar{s}) = H^*(s)$ for all $s \in \mathbb{C}$, and $H(s) + H^*(s) = \mathrm{Re}(H(s)) \geq 0$ for all s with $\mathrm{Re}(s) > 0$ (in the MIMO case, $>$ ($\geq$) denotes positive (semi-)definite). Hence, a real stable system is passive if and only if $\mathrm{Re}(H(s)) \geq 0$ for all s with $\mathrm{Re}(s) > 0$. Passivity [54] means that the system does not generate energy and only absorbs energy from the sources used to excite it.

A system is called causal if for any t, the output $\mathbf{y}(t)$ at time t depends only on inputs $\mathbf{u}(\tilde{t})$ with $\tilde{t} \leq t$. A system is called controllable if for any t_0, starting from zero initial state $\mathbf{x}(t_0) = 0$, every state $\mathbf{x}(t)$ can be reached via a suitable input $\mathbf{u}(t)$. It follows from (1.3.5) that controllability means that the range of $e^{At}B$ is $\mathbb{R}^n$, and by expansion of e^{At} and the Cayley-Hamilton theorem[1] a system (A, B, C, D) is controllable if and only if the controllability matrix

$$C(A, B) = \begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix}$$

has full rank n. Dually, observability means that the initial state can be uniquely determined from the input and the output, or equivalently, that with $\mathbf{u}(t) = 0$, a zero output $\mathbf{y}(t) = 0$ implies that the initial state is zero. Hence, a system is called observable if and only if the observability matrix

$$O(A, C) = \begin{bmatrix} C & A^*C & \dots & (A^{n-1})^*C \end{bmatrix}^*$$

has full rank n. A system is called complete if it is both controllable and observable, and a realization is minimal if and only if it is both controllable and observable. The set of poles of a minimal realization coincides with the set of eigenvalues of A.

If $\mathbf{u}(t) = 0$, the output (1.3.5) is called the zero-input or transient response $\mathbf{y}_h(t)$, which depends on the initial conditions only. If $\mathbf{x}(t_0) = 0$ the output is called the zero-state response. If $\mathbf{u}(t) = \delta(t)$ (the unit Kronecker delta distribution) and $\mathbf{x}(t_0) = 0$, $\mathbf{h}(t) = C^*e^{At}B + D\delta(t)$ is called the impulse response. The steady state response is the output for $t \to \infty$. If the system is stable it follows that $\lim_{t\to\infty} \|\mathbf{y}_h(t)\| = 0$ and hence the steady state response is $\mathbf{y}_p(t)$, where $\mathbf{y}_p(t)$ is the unique (particular) solution (1.3.5). If $u(t) = 1(t)$, the unit step function ($u(t) = 1$ for $t \geq 0$ and $u(t) = 0$ for $t < 0$), the output is called the step response, and if $u(t) = t1(t)$, the output is called the ramp response.

For stable systems, the controllability gramian $P \in \mathbb{R}^{n\times n}$ for a linear time-invariant system is defined as

$$P = \int_0^\infty e^{A\tau}BB^*e^{A^*\tau}d\tau,$$

and can be interpreted in the following way: the minimum energy $J(\mathbf{u}) = \int_{-\infty}^0 \mathbf{u}^*(t)\mathbf{u}(t)dt$ of the input to arrive at $\mathbf{x}(0) = \mathbf{x}_0$ is $J(\mathbf{u}) = \mathbf{x}_0^*P^{-1}\mathbf{x}_0$. Hence, states in the span of the eigenvectors corresponding to small eigenvalues of P are difficult to reach. Dually, the observability gramian $Q \in \mathbb{R}^{n\times n}$ is defined as

$$Q = \int_0^\infty e^{A^*\tau}CC^*e^{A\tau}d\tau.$$

[1] The Cayley-Hamilton theorem states that a matrix A satisfies its characteristic equation $p(\lambda) = \det(A - \lambda I) = 0$.

A system released from $\mathbf{x}(0) = \mathbf{x}_0$ with $\mathbf{u}(t) = 0, t \geq 0$ has

$$\int_0^\infty \mathbf{y}^*(t)\mathbf{y}(t)dt = \mathbf{x}_0^* Q \mathbf{x}_0,$$

and it follows that states in the span of the eigenvectors corresponding to small eigenvalues of Q are difficult to observe. A stable system is controllable if and only if P is positive definite, and it is observable if and only if Q is positive definite. Substitution shows that for stable systems, the gramians are solutions of the Lyapunov equations

$$AP + PA^* + BB^* = 0, \quad \text{and} \quad A^*Q + QA + CC^* = 0.$$

Although the gramians are not similar under state-space transformation $\mathbf{x} \mapsto T\mathbf{x}$ with $T \in \mathbb{R}^{n\times n}$ nonsingular, their product PQ is.

The Hankel operator $\mathcal{H}$ maps past inputs $\mathbf{u}(t)$ $(t < 0)$ to future outputs $\mathbf{y}(t)$ $(t > 0)$:

$$\mathcal{H} : L_2^m((-\infty, 0)) \longrightarrow L_2^p((0, \infty)) : u(t) \mapsto \int_{-\infty}^0 Ce^{A(t-\tau)}Bu(\tau)d\tau,$$

where $L_p^n(I)$ denotes the Lebesque space $L_p^n(I) = \{\mathbf{x}(t) : I \to \mathbb{R}^n | \|\mathbf{x}(t)\|_p < \infty\}$, and is of importance in control theory and model order reduction. The singular values of the Hankel operator are the square roots of the eigenvalues of $\mathcal{H}^*\mathcal{H}$, and are called the Hankel singular values. It can be shown that for stable complete systems, the Hankel singular values $\sigma_i^{\mathcal{H}}$ can be computed as the positive square roots of the eigenvalues of the product of the controllability gramian P and observability gramian Q:

$$\sigma_i^{\mathcal{H}} \equiv \sigma_i(\mathcal{H}) = \sqrt{\lambda_i(PQ)}, \quad i = 1, 2, \ldots, n.$$

The fact that the Hankel operator has a finite number n of singular values partly explains its importance and usefulness in control theory and model order reduction, since this allows, amongst others, for the definition of the Hankel norm of a system (A, B, C, D) with transfer function $H(s)$:

$$\|H\|_{\mathcal{H}} \equiv \sup_{\mathbf{u}\in L_2(-\infty,0)} \frac{\|\mathcal{H}\mathbf{u}\|_2}{\|\mathbf{u}\|_2} = \sigma_{\max}(\mathcal{H}) = \lambda_{\max}^{1/2}(PQ).$$

The Hankel singular values for a system and its transfer function are defined to be the same. The Hankel singular values are invariant under state-space transformations, since similarity of PQ is preserved under state-space transformations, and are so called input-output invariants. The Hankel singular values and Hankel norm are of great importance for certain types of model order reduction, as is described in Section 1.4. See [62] for more details on the Hankel operator and Hankel singular values.

1.3.2 Descriptor systems

Linear time invariant descriptor systems (E, A, B, C, D) are of the form

$$\begin{cases} E\dot{\mathbf{x}}(t) &= A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) &= C^*\mathbf{x}(t) + D\mathbf{u}(t), \end{cases} \tag{1.3.6}$$

where $E, A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{n\times p}$, $\mathbf{x}(t) \in \mathbb{R}^n$, $\mathbf{u}(t) \in \mathbb{R}^m$, $\mathbf{y}(t) \in \mathbb{R}^p$, and $D \in \mathbb{R}^{p\times m}$. Terminology is similar as for state-space systems, but here the descriptor matrix E may be singular. If E is nonsingular, the descriptor can be transformed to the state-space system $(E^{-1}A, E^{-1}B, C, D)$, although this usually is not attractive from a computational viewpoint. The transfer function $H(s)$ of (1.3.6) is defined as

$$H(s) = C^*(sE - A)^{-1}B + D. \tag{1.3.7}$$

The transfer function $H(s)$ can be expressed as a sum of residue matrices $R_j \in \mathbb{C}^{p\times m}$ over the finite first order poles (if (A, E) is nondefective) [78], cf. (1.3.4):

$$H(s) = D + R_\infty + \sum_{j=1}^{\tilde{n}} \frac{R_j}{s - p_j}, \tag{1.3.8}$$

where $\tilde{n} \leq n$ is the number of finite poles (eigenvalues), and R_∞ is the contribution due to poles (eigenvalues) at $\pm\infty$ ($R_\infty = 0$ often). Poles p_j with large $\|R_j\|_2/|\mathrm{Re}(p_j)|$ (or $\|R_j\|_2$) are also called dominant poles and are studied in Chapters 2–6.

Concepts such as controllability and observability, the corresponding gramians, and Hankel singular values can be generalized to descriptor systems, but there is no uniform approach in the literature. Interpretation of these concepts in terms of states that are hard to reach and to observe, however, remains valid and is sufficient for the material in this thesis; most of the systems considered in this thesis are descriptor systems. The reader is referred to [149] and references therein for a clear generalization and applications to model order reduction.

1.4 Model order reduction

The model order reduction problem is to find, given an n-th order (descriptor) dynamical system (E, A, B, C, D) (1.3.6), a k-th order system $(\widetilde{E}, \widetilde{A}, \widetilde{B}, \widetilde{C}, D)$:

$$\begin{cases} \widetilde{E}\dot{\tilde{\mathbf{x}}}(t) &= \widetilde{A}\tilde{\mathbf{x}}(t) + \widetilde{B}\mathbf{u}(t) \\ \tilde{\mathbf{y}}(t) &= \widetilde{C}^*\tilde{\mathbf{x}}(t) + D\mathbf{u}(t), \end{cases} \tag{1.4.1}$$

where $k < n$, and $\widetilde{E}, \widetilde{A} \in \mathbb{R}^{k\times k}$, $\widetilde{B} \in \mathbb{R}^{k\times m}$, $\widetilde{C} \in \mathbb{R}^{k\times p}$, $\tilde{\mathbf{x}}(t) \in \mathbb{R}^k$, $\mathbf{u}(t) \in \mathbb{R}^m$, $\tilde{\mathbf{y}}(t) \in \mathbb{R}^p$, and $D \in \mathbb{R}^{p\times m}$. The number of inputs and outputs are the same as for the original system, and the corresponding transfer function becomes

$$\widetilde{H}(s) = \widetilde{C}^*(s\widetilde{E} - \widetilde{A})^{-1}\widetilde{B} + D.$$

The reduced-order model (1.4.1) should satisfy some or all of the following requirements [6]:

1. The approximation error must be small, and a global error bound should exist. Usually this means that the output error $\|\mathbf{y}(t) - \tilde{\mathbf{y}}(t)\|$ should be minimized for some or even all inputs $\mathbf{u}(t)$ in an appropriate norm.
2. The order of the reduced system is much smaller than the order of the original system: $k \ll n$.
3. Preservation of (physical and numerical) properties such as stability and passivity.
4. The procedure must be computationally stable and efficient.
5. The procedure must be based on some error tolerance (automatically) and a cheap measurement for the error is desired.

Depending on the application area, there may be some additional requirements:

- The reduced-order model should preserve the structure of the original model. In many practical situations, the original model is a linearization of a second order system, so that the system matrices (E, A, B, C) have a specific structure that needs to be preserved in the system matrices $(\widetilde{E}, \widetilde{A}, \widetilde{B}, \widetilde{C})$. See also [11, 55].
- A step further, the reduced-order model itself must be realizable, that is, the model should be realized as physical device. In general, reduced-order models have no clear physical meaning.
- The procedure must fit in existing simulation software, for instance in a hierarchical electric circuit simulator.

All the model order reduction methods discussed in this thesis have the common property that the reduced-order model is constructed via a Petrov-Galerkin type of projection

$$(\widetilde{E}, \widetilde{A}, \widetilde{B}, \widetilde{C}, D) \equiv (Y^*EX, Y^*AX, Y^*B, X^*C, D),$$

where $X, Y \in \mathbb{R}^{n\times k}$ are matrices whose columns form bases for relevant subspaces of the state-space.

Concerning the first requirement, the approximation error $\mathbf{y}(t) - \tilde{\mathbf{y}}(t)$ can be measured in several norms:

- The "eye-norm": a Bode magnitude (phase) plot of a transfer function plots the magnitude (phase) of $H(i\omega)$, usually in decibel, for a number of frequencies ω in the frequency range of interest, see Figure 1.1. If the transfer function of the original system can be evaluated at enough $s = i\omega$ to produce an accurate Bode plot, the original frequency response can be compared

with the frequency response of the reduced model. The degree in which the responses match may give an indication of the quality of the reduced model and may be sufficient for certain applications, but it should be noted that the relative errors may give a different view. In practical situations the experience of a domain specialist, for instance an electrical engineer, may be helpful to judge the quality and usefulness of the reduced-order model.

- The induced $\|\cdot\|_2$ norm or $\|\cdot\|_\infty$ norm: via Parseval's identity, the operator norm induced by the 2-norm in the frequency domain is defined as [62]

$$\|H\|_\infty \equiv \sup_{\omega\in\mathbb{R}} \sigma_{\max}(H(i\omega)),$$

where $\sigma_{\max}$ is the maximum singular value. This gives for the approximation error

$$\|\mathbf{y} - \tilde{\mathbf{y}}\|_2 = \|Y - \widetilde{Y}\|_2 \leq \|H - \widetilde{H}\|_\infty \|\mathbf{u}\|_2,$$

where Y and $\widetilde{Y}$ are the Laplace transforms of $\mathbf{y}$ and $\tilde{\mathbf{y}}$, respectively. For balanced truncation and modal truncation methods, to be described in the next section, there exist upper bounds for this error, although not always easily computable.

- The Hankel norm $\|\cdot\|_{\mathcal{H}}$: it is possible to construct a realization for the error transfer function $H(s) - \widetilde{H}(s)$, see [62], and the error can be measured in the Hankel norm

$$\|H - \widetilde{H}\|_{\mathcal{H}} \leq \|H - \widetilde{H}\|_\infty.$$

Also for Hankel norm approximation there exist upper bounds for this error, although not always easily computable.

1.5 Methods for eigenvalue problems and model order reduction

Roughly speaking, the methods for eigenvalue problems can be divided in two categories (see [156] for a historical background of eigenproblems): *full space* methods based on the QR method [52] for the standard eigenproblem, and the QZ method [101] for the generalized eigenproblem (see [64, Chapter 7] for efficient implementations), and *iterative subspace* methods based on the Lanczos [83], Arnoldi [7], Davidson [37], and Jacobi-Davidson [140] methods. Although the full space methods are sometimes called direct methods, they are in fact iterative as well. The complexity of the full space methods is $O(n^3)$, where n is the order of the matrix, irrespective of the sparsity, and hence they are only applicable to moderately sized problems. The complexity of the iterative subspace methods, on the other hand, usually depends on the number of nonzero elements in A (and B), and are especially applicable to large-scale sparse matrices of practically unlimited order. Full space methods usually compute the complete spectrum (and corresponding

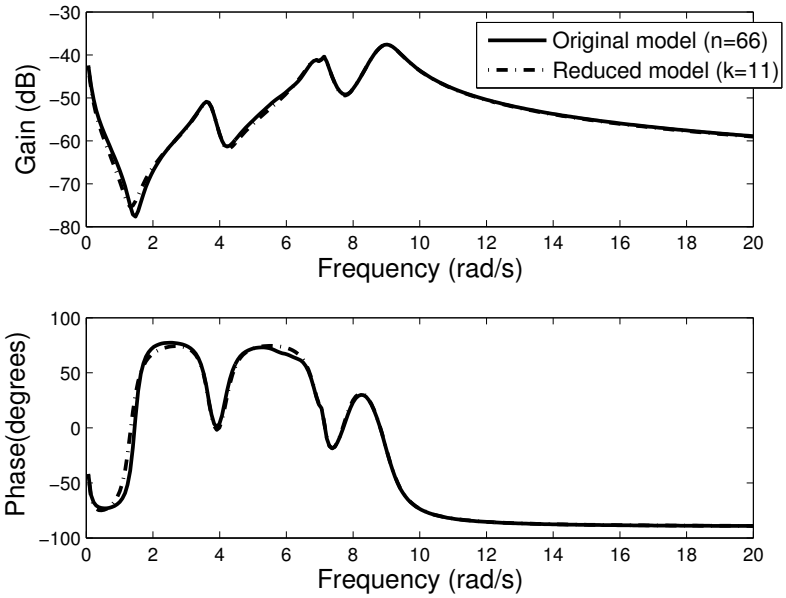

Figure 1.1: Bode magnitude (upper) and phase plot of original 66-th order model (solid), and 11-th order reduced model (dash-dot). The frequency response of the reduced model shows small deviations from the exact frequency response.

eigenvectors), while iterative subspace methods typically focus on the computation of a few specific eigenvalues and eigenvectors.

Basically, there are three main approaches for model order reduction: (1) methods based on balanced truncation [102] and Hankel norm approximation [62], (2) Padé [14] and moment matching [67] type methods, and (3) modal approximation methods [38]. The balanced truncation based methods can also be interpreted as (iterative) full space methods and have complexity $O(n^3)$, although there are developments that make these methods applicable to large sparse systems (see [18] and references therein). Moment matching methods, on the other hand, are usually based on Krylov subspace techniques such as Lanczos and Arnoldi (or rational variants [131]), and applicable to large-scale systems. Modal approximation requires selection of dominant eigenvalues (and eigenvectors) and these can be computed via full space methods (QR, QZ) or iterative subspace methods. One of the contributions of this thesis is an efficient and effective iterative subspace method to compute specifically the dominant poles and corresponding eigenvectors (Chapters 2–6).

Without loss of generality, $D = 0$ in the following, unless stated otherwise. The moments are the coefficients of a power series expansion of $H(s)$ around a finite $s_0 \notin \Lambda(A, E)$:

$$H(s) = C^*(sE - A)^{-1}B = \sum_{i=0}^{\infty} M_i(s - s_0)^i,$$

where $M_i = -C^*((s_0E - A)^{-1}E))^i(s_0E - A)^{-1}B$ for $i \geq 0$ are called the shifted moments, and, if $s_0 = 0$, the $M_i = -C^*(A^{-1}E))^iA^{-1}B$ are simply called the moments of $H(s)$. If E is nonsingular, a Neumann expansion around $s_0 = \infty$ gives $M_i = C^*(E^{-1}A)^{-i-1}E^{-1}B$ for $i < 0$, called the Markov parameters.

For a SISO system $(E, A, \mathbf{b}, \mathbf{c})$, a reduced-order system $(\widetilde{E}, \widetilde{A}, \tilde{\mathbf{b}}, \tilde{\mathbf{c}})$ with transfer function $\widetilde{H}$ and moments $\widetilde{M}$ is called a k-th Padé approximant if [14]

$$H(s) = \widetilde{H}(s) + O((s - s_0)^{2k}), \quad (s \to s_0),$$

with $M_i = \widetilde{M}_i$ for $i = 0, 1, \ldots, 2k-1$. A reduced-order model whose first $2k$ Markov parameters are equal to the first $2k$ Markov parameters of the original system is called a partial realization [67]. A multipoint Padé approximation or rational interpolant is a reduced-order model whose moments $M_i^{(j)}$ match the first $2J_j$ moments at σ_j, with $i = 0, \ldots, J_j-1$ and $J_0+\cdots+J_{l-1} = k$ for l interpolation points σ_j $(j = 0, 1, \ldots, l-1)$ [67, 14]. At the interpolation points σ_j, Padé approximations are exact, but accuracy is lost away from σ_j, even more rapidly if σ_j is close to a pole [14, 30]. Therefore, state-of-the-art moment matching techniques employ multiple interpolation points in order to match the frequency response for a large frequency range of interest [58, 70]. However, the choice of the interpolation points σ_j, and number of moments to match around each point, is usually not easy, except possibly in applications such as electric circuit simulation, where the operating frequency seems to be a good interpolation point [18, 48].

The Arnoldi and Lanczos based methods, and related model order reduction methods, construct bases for Krylov subspaces. A k-th order Krylov subspace for a square matrix A and a vector $\mathbf{v}$ is defined as

$$\mathcal{K}^k(A, \mathbf{v}) = \text{span}(\mathbf{v}, A\mathbf{v}, \ldots, A^{k-1}\mathbf{v}),$$

and can be efficiently computed if applications of A to $\mathbf{v}$ are cheap. Together with the moment matching property (see Section 1.5.3 and Section 1.5.4) this makes Krylov subspace methods successful for large sparse matrices. The methods differ in the way the bases are computed: theoretically the subspaces are the same, but numerically the condition of the bases may vary considerably. Note that Krylov subspaces are invariant under shift and scaling of A, i.e. $\mathcal{K}^k(\alpha A+\sigma I, \mathbf{v}) = \mathcal{K}^k(A, \mathbf{v})$, but that in general $\mathcal{K}^k((\sigma I - A)^{-1}, \mathbf{v}) \neq \mathcal{K}^k(A, \mathbf{v})$. Methods based on rational interpolation construct bases for rational Krylov subspaces [131]. For a regular matrix pencil (A, E), l vectors $\mathbf{v}_j$ and l shifts σ_j, a rational Krylov subspace is defined as [131]

$$\sum_{j=1}^{l} \mathcal{K}^{J_j}((\sigma_j E - A)^{-1}E, \mathbf{v}_j),$$

where $J_1 + \cdots + J_l = k$.

Each of the following subsections briefly describes an eigenvalue method together with (closely) related model order reduction methods, and provides references for further reading. Other overviews of model order reduction methods are given in [5, 18].

1.5.1 Full space methods and balanced truncation

Full space methods such as QR and QZ can be used to computed complete eigendecompositions of eigenproblems. The SVD [63], that computes the singular value or spectral decomposition $A = U\Sigma V^*$ of a general matrix A, is an important ingredient for balanced truncation based model order reduction methods. Given a state-space system (A, B, C, D), balanced truncation constructs a reduced-order model by truncating a balanced realization of (A, B, C, D). A balanced realization (A_b, B_b, C_b, D_b) is a realization for which the controllability and observability gramians are equal diagonal matrices $P_b = Q_b = \mathrm{diag}(\sigma_1^{\mathcal{H}}, \ldots, \sigma_n^{\mathcal{H}})$, where (for stable systems) P_b and Q_b are the solutions of the Lyapunov equations

$$A_b P_b + P_b A_b^* + B_b B_b^* = 0, \quad \text{and} \quad A_b^* Q_b + Q_b A_b + C_b C_b^* = 0, \tag{1.5.1}$$

and $\sigma_i^{\mathcal{H}} = \sqrt{\lambda_i(P_b Q_b)}$ are the Hankel singular values (see also Section 1.3). Recall that the Hankel singular values are input-output invariants because PQ is similar under state-space transformations $\mathbf{x} \to T\mathbf{x}$. A state-space transformation defined by T that transforms (A, B, C, D) to a balanced realization

$$(TAT^{-1}, TB, T^{-T}C, D) = \left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, D \right),$$

with $A_{11} \in \mathbb{R}^{k\times k}$, $B_1 \in \mathbb{R}^{k\times m}$, $C_1 \in \mathbb{R}^{k\times p}$ and $k < n$, is called a balancing transformation. For minimal (complete) systems there exists such a balancing transformation. Balanced truncation [102] constructs a reduced k-order model of (A, B, C, D) by truncating:

$$(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}) = (A_{11}, B_1, C_1, D).$$

Important properties [62, 102] of this truncated balanced realization are that the gramians $\widetilde{P}$ and $\widetilde{Q}$ are balanced and equal to $\widetilde{P} = \widetilde{Q} = \mathrm{diag}(\sigma_1^{\mathcal{H}}, \ldots, \sigma_k^{\mathcal{H}})$, and that there is the absolute error bound [45, 62]

$$\|H - \widetilde{H}\|_\infty \leq 2 \sum_{i=k+1}^{n} \sigma_i^{\mathcal{H}}. \tag{1.5.2}$$

It follows that the reduced-order model preserves the largest Hankel singular values, and moreover, that k can be chosen such that the error (1.5.2) is smaller than the required tolerance. This all is applicable, however, under the assumption that all Hankel singular values are known, or at least computed in order of decreasing magnitude.

The first step in balanced truncation of a stable minimal system is to compute the unique positive definite solutions P and Q of the Lyapunov equations (1.5.1). Given P and Q, a balancing transformation T can be computed as $T = \Sigma^{\frac{1}{2}} U^T R^{-T}$, where $P = R^T R$ is a Cholesky factorization and $U\Sigma^2 U^T = RQR^T$ is an SVD of RQR^T. The state-space transformation $\mathbf{x} \to T\mathbf{x}$ is a balancing transformation, $P_b = Q_b = \Sigma$, and the balanced system can be truncated up to required accuracy

by using the bound in (1.5.2). Note that this truncated system is not an optimal approximation. In [62] it is shown how, using a balancing transformation, an approximation can be constructed that is optimal in the Hankel norm, that is, $\sigma_{k+1}^{\mathcal{H}} \leq \|H - \widetilde{H}\|_{\mathcal{H}} < \sigma_k^{\mathcal{H}}$ (this approximation also has error bound (1.5.2) and hence is not necessarily more accurate than the balanced truncation, in the $\|\cdot\|_\infty$ norm).

There are several ways to compute full solutions P and Q of the Lyapunov equations, see for instance [5, Chapter 6] and references therein, but these (iterative) full space solution methods have complexity $O(n^3)$ (except in some cases for the ADI iteration [162]). Also the SVD and Cholesky factorization have complexity $O(n^3)$ for general matrices [64]. Hence, when relying on full space methods, balanced truncation is only applicable to systems of moderate order n and not feasible for large-scale sparse systems. The full space balanced truncation methods are also called SVD-type model order reduction methods.

There are methods that compute low-rank solutions of the Lyapunov equations [87, 115] or combine SVD-type methods with (rational) Krylov approaches [71], and these methods make it possible to apply balanced truncation to large sparse systems as well (see [5, 18, 19] for more details and references). Generalizations and methods for Lyapunov equations arising from descriptor systems are described in [149]. Balanced truncation is a concept in the context of proper orthogonal decomposition (POD), also known as the method of empirical eigenfunctions, a method that tries to extract the dominant dynamics from the time response and that is also applicable to nonlinear systems (see [5, Section 9.1]).

1.5.2 The Power method and Asymptotic Waveform Evaluation

Given a matrix $A \in \mathbb{R}^{n\times n}$ and a vector $\mathbf{v}_1$ with $\|\mathbf{v}_1\|_2 = 1$, the power method computes the sequence $\{\mathbf{v}_i\}_{i>0}$ as $\mathbf{v}_1, \mathbf{v}_i = A\mathbf{v}_{i-1}/\|A\mathbf{v}_{i-1}\|_2$ $(i > 1)$, and corresponding approximate eigenvalues via the Rayleigh quotient $\mathbf{v}_i^* A\mathbf{v}_i/\mathbf{v}_i^*\mathbf{v}_i$, as is shown in Algorithm 1.1. It is well known that the sequence converges to the eigenvector corresponding to the dominant (in absolute value) eigenvalue (if it is simple) of A if $\mathbf{v}_1$ has a component in that direction, see for example [156, Chapter 4].

Algorithm 1.1 The Power method

INPUT: $n \times n$ matrix A, initial vector $\mathbf{v}_1 \neq 0$
OUTPUT: Dominant eigenpair $(\lambda_1, \mathbf{x}_1)$

1: $\mathbf{v}_1 = \mathbf{v}_1/\|\mathbf{v}_1\|_2$
2: **for** $i = 1, 2, \ldots,$until convergence **do**
3: $\quad \mathbf{v}_{i+1} = A\mathbf{v}_i$
4: $\quad \theta_i = \mathbf{v}_i^*\mathbf{v}_{i+1}$
5: $\quad \mathbf{v}_{i+1} = \mathbf{v}_{i+1}/\|\mathbf{v}_{i+1}\|_2$
6: **end for**

Asymptotic Waveform Evaluation (AWE) [30, 117] computes the $2k$ moments m_i of a SISO transfer function $H(s)$ explicitly by applying the power method to $(s_0E - A)^{-1}E$ and $\mathbf{v}_1 = \mathbf{b}$, for an expansion point s_0 (typically $s_0 = 0$ or $s_0 = \infty$). In the second step, the transfer function of the approximate k-th order realization is forced to match the $2k$ moments m_i of the original impulse response, which, although not mentioned in the original AWE literature, can be achieved by using a projector based on the vectors generated by the power method [57].

AWE suffers from numerical problems that are caused by the explicit moment matching via the power method: since the $\mathbf{v}_i$ converge to the eigenvector corresponding to the (absolutely) largest eigenvalue, the $\mathbf{v}_i$ and computed moments $m_i = \mathbf{c}^*\mathbf{v}_i$ practically contain dominant information on the largest eigenvalue, resulting in a poor reduced-order model. These effects are already notable for small values of i, and although they can be handled by using several different expansion points s_i, AWE is rarely applicable in practice. For more details on the numerical problems with AWE see [48, 57].

After k iterations, the vectors $\mathbf{v}_i$ generated by the power method span the Krylov subspace

$$\mathcal{K}^k(A, \mathbf{v}_1) = \text{span}(\mathbf{v}_1, A\mathbf{v}_1, \ldots, A^{k-1}\mathbf{v}_1),$$

but from a numerical viewpoint they form an ill-conditioned basis. The Lanczos and Arnoldi methods, to be discussed in the next subsections, are numerically more stable methods for the construction of orthonormal bases for $\mathcal{K}^k(A, \mathbf{v}_1)$.

1.5.3 The Lanczos method and Padé Via Lanczos

Lanczos [83] proposes a numerically more stable method to construct a basis $(\mathbf{v}_1, \ldots, \mathbf{v}_{k+1})$ for the Krylov subspace $\mathcal{K}^{k+1}(A, \mathbf{v}_1)$ for a general matrix A and vector $\mathbf{v}_1$. In fact, the unsymmetric or two-sided Lanczos method computes bi-orthogonal bases $(\mathbf{v}_1, \ldots, \mathbf{v}_{k+1})$ and $(\mathbf{w}_1, \ldots, \mathbf{w}_{k+1})$ for the Krylov subspaces $\mathcal{K}^{k+1}(A, \mathbf{v}_1)$ and $\mathcal{K}^{k+1}(A^*, \mathbf{w}_1)$, respectively, by the following two three term recurrence relations ($\mathbf{v}_0 = \mathbf{w}_0 = 0$):

$$\begin{aligned} \rho_{i+1}\mathbf{v}_{i+1} &= A\mathbf{v}_i - \alpha_i\mathbf{v}_i - \beta_i\mathbf{v}_{i-1}, \quad (i = 1, \ldots, k), \\ \eta_{i+1}\mathbf{w}_{i+1} &= A^*\mathbf{w}_i - \bar{\alpha}_i\mathbf{w}_i - \gamma_i\mathbf{w}_{i-1}, \quad (i = 1, \ldots, k), \end{aligned}$$

as is also shown in Alg. 1.2.

Algorithm 1.2 The two-sided Lanczos method

INPUT: $n \times n$ matrix A, nonzero vectors $\mathbf{v}_1, \mathbf{w}_1$ with $\mathbf{w}_1^* \mathbf{v}_1 \neq 0$
OUTPUT: $n \times k$ matrices $V = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$, $W = [\mathbf{w}_1, \ldots, \mathbf{w}_k]$,
$k \times k$ matrices $T = \text{tridiag}(\rho_{2:k}, \alpha_{1:k}, \beta_{2:k})$, $S = \text{tridiag}(\eta_{2:k}, \bar{\alpha}_{1:k}, \gamma_{2:k})$,
vectors $\mathbf{v}_{k+1}, \mathbf{w}_{k+1}$, scalars $t_{k+1,k}, s_{k+1,k}$

1: $\rho_1 = \|\mathbf{v}_1\|_2$, $\eta_1 = \|\mathbf{w}_1\|_2$, $\mathbf{v}_1 = \mathbf{v}_1/\rho_1$, $\mathbf{w}_1 = \mathbf{w}_1/\eta_1$
2: $\mathbf{v}_0 = 0$, $\mathbf{w}_0 = 0$, $\delta_0 = 1$
3: **for** $i = 1, 2, \ldots, k$ **do**
4: $\quad \delta_i = \mathbf{w}_i^* \mathbf{v}_i$
5: $\quad \mathbf{x} = A\mathbf{v}_i$
6: $\quad \alpha_i = \mathbf{w}_i^* \mathbf{x}/\delta_i$
7: $\quad \beta_i = \delta_i \eta_i / \delta_{i-1}$
8: $\quad \gamma_i = \delta_i \rho_i / \delta_{i-1}$
9: $\quad \mathbf{x} = \mathbf{x} - \alpha_i \mathbf{v}_i - \beta_i \mathbf{v}_{i-1}$
10: $\quad \mathbf{y} = A^* \mathbf{w}_i - \bar{\alpha}_i \mathbf{w}_i - \gamma_i \mathbf{w}_{i-1}$
11: $\quad \rho_{i+1} = \|\mathbf{x}\|_2$, $\eta_{i+1} = \|\mathbf{y}\|_2$
12: $\quad \mathbf{v}_{i+1} = \mathbf{x}/\rho_{i+1}$, $\mathbf{w}_{i+1} = \mathbf{y}/\eta_{i+1}$
13: **end for**

The coefficients $\alpha_j, \beta_j, \gamma_j, \delta_j, \eta_j, \rho_j$ are computed such that the basis vectors are biorthogonal, i.e. $\mathbf{w}_i^* \mathbf{v}_j = 0$ if $i \neq j$ and $\mathbf{w}_i^* \mathbf{v}_j = \delta_j$ if $j = i$, for $i, j = 1, \ldots, k+1$. Then the following relations hold:

$$\begin{aligned} AV_k &= V_k T_k + \rho_{k+1} \mathbf{v}_{k+1} \mathbf{e}_k^T, \\ A^* W_k &= W_k S_k + \eta_{k+1} \mathbf{w}_{k+1} \mathbf{e}_k^T, \end{aligned}$$

where $V_k = [\mathbf{v}_1, \ldots, \mathbf{v}_k] \in \mathbb{R}^{n \times k}$ and $W_k = [\mathbf{w}_1, \ldots, \mathbf{w}_k] \in \mathbb{R}^{n \times k}$, and $T_k = \text{tridiag}(\rho_{2:k}, \alpha_{1:k}, \beta_{2:k})$ and $S_k = \text{tridiag}(\eta_{2:k}, \bar{\alpha}_{1:k}, \gamma_{2:k})$ are tridiagonal $k \times k$ matrices. The eigenvalues of T (S) are approximate eigenvalues of A (A^*), also known as Petrov values. The process can suffer from break down if $\mathbf{w}_i^* \mathbf{v}_i = 0$ for nonzero $\mathbf{v}_i$ and $\mathbf{w}_i$, and remedies are described in for example [56]. See [64, 156] and references therein for more details and robust implementations of Lanczos methods.

Padé Via Lanczos (PVL), derived independently in [48, 57], uses the two-sided Lanczos method to deal with the problems of AWE. For a single interpolation point $\sigma_0 \notin \Lambda(A, E)$, the transfer function of the SISO system $(E, A, \mathbf{b}, \mathbf{c})$ can be rewritten as

$$H(s) = \mathbf{c}^*(sE - A)^{-1}\mathbf{b} = \mathbf{c}^*(I + (s - \sigma_0)(\sigma_0 E - A)^{-1}E)^{-1}((\sigma_0 E - A)^{-1}\mathbf{b}).$$

Two-sided Lanczos can be used to compute bi-orthogonal bases $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ and $(\mathbf{w}_1, \ldots, \mathbf{w}_k)$ for the Krylov spaces $\mathcal{K}((s_0 E - A)^{-1}E, (s_0 E - A)^{-1}\mathbf{b})$ and $\mathcal{K}((s_0 E - A)^{-*}E^*, \mathbf{c})$, and the reduced-order model is constructed as $(T_k, \sigma_0 T_k - I, W_k^*(\sigma_0 E - A)^{-1}\mathbf{b}, V_k^* \mathbf{c})$. It can be shown that the reduced-order model preserves $2k$ moments [48, Section 3] and [57, Thm. 1]. In general PVL is not stability preserving, since the Petrov values (eigenvalues of T_k) are not necessarily located in the left half

plane, even if that is the case for the eigenvalues of (E, A). Passivity preserving PVL like methods exist for specific applications as *RLC*-circuits (SyPVL) [53], and there are also PVL methods for MIMO systems (MPVL) [54]. Because of its short recurrences, PVL is an efficient way to create a reduced model, but it may have numerical difficulties due to break down; see [49] for a more robust implementation (via block Lanczos). In [13] an estimate is given for the approximation error of the model computed by PVL.

As mentioned before, single interpolation point methods usually produce reduced-order models that are accurate in the neighborhood of the interpolation point. Rational Krylov methods [131] can be used to construct bases for rational Krylov subspaces, that may lead to improved reduced-order models. In [58, 70], a model order reduction method based on a rational Lanczos scheme is presented. In the SISO case, a reduced-order model is constructed as $(W^*EV, W^*AV, W^*\mathbf{b}, V^*\mathbf{c})$, where the columns of the $n \times k$ matrices V and W form bases for the rational Krylov subspaces

$$\sum_{j=1}^{l} \mathcal{K}^{J_j}((\sigma_j E - A)^{-1}E, (\sigma_j E - A)^{-1}\mathbf{b}),$$

and

$$\sum_{j=1}^{l} \mathcal{K}^{J_j}((\sigma_j E - A)^{-*}E^*, (\sigma_j E - A)^{-*}\mathbf{c}),$$

respectively, with $2J_j$ the number of moments to match around interpolation point σ_j, with $J_1 + \cdots + J_l = k$. The reader is referred to Section 3.8 and [70] for more information about two-sided methods. For a two-sided interpolation approach for MIMO systems, see [59].

1.5.4 The Arnoldi method and PRIMA

Arnoldi [7] computes an orthonormal basis $(\mathbf{v}_1, \ldots, \mathbf{v}_{k+1})$ for the Krylov subspace $\mathcal{K}^{k+1}(A, \mathbf{v}_1)$ for a general matrix A and vector $\mathbf{v}_1$. Given a basis $(\mathbf{v}_1, \ldots, \mathbf{v}_i)$, the next basis vector $\mathbf{v}_{i+1}$ is computed as the normalized orthogonal complement of $A\mathbf{v}_i$ in $\text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_i)$. Algorithm 1.3 shows the Arnoldi process with repeated modified Gram-Schmidt orthogonalization (MGS): if the new vector after orthogonalization is a fraction $\gamma < 1$ of the vector before orthogonalization, it is orthogonalized again to deal with cancellation effects (a practical choice is $\gamma = 0.1$) [36, 64].

Algorithm 1.3 The Arnoldi method

INPUT: $n \times n$ matrix A, nonzero vector $\mathbf{v}_1$
OUTPUT: $n \times k$ matrix $V = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$, $k \times k$ matrix $H = [h_{ij}]$, vector $\mathbf{v}_{k+1}$, scalar $h_{k+1,k}$

$\mathbf{v}_1 = \mathbf{v}_1/\|\mathbf{v}_1\|_2$, $H = 0^{k\times k}$
for $i = 1, 2, \ldots, k$ **do**
  $\mathbf{w} = A\mathbf{v}_i$
  $(\mathbf{w}, \mathbf{g})$ =MGS($[\mathbf{v}_1, \ldots, \mathbf{v}_i], \mathbf{w}$) {Alg. 1.4}
  $h_{ji} = g_j$ $(j = 1, \ldots, i)$
  $h_{i+1,i} = \|\mathbf{w}\|_2$
  $\mathbf{v}_{i+1} = \mathbf{w}/h_{i+1,i}$
end for

Algorithm 1.4 Modified Gram-Schmidt (MGS)

INPUT: $n \times k$ matrix $V = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$, nonzero vector $\mathbf{v}$, $\gamma < 1$
OUTPUT: vector $\mathbf{w} \perp V$, $k \times 1$ vector $\mathbf{h}$

$\mathbf{w} = \mathbf{v}$
$\tau_0 = \|\mathbf{w}\|_2$
for $i = 1, \ldots, k$ **do**
  $h_i = \mathbf{v}_i^*\mathbf{w}$
  $\mathbf{w} = \mathbf{w} - h_i\mathbf{v}_i$
end for
if $\|\mathbf{w}\|_2 \leq \gamma\tau_0$ **then**
  for $i = 1, \ldots, k$ **do**
    $\alpha = \mathbf{v}_i^*\mathbf{w}$
    $\mathbf{w} = \mathbf{w} - \alpha\mathbf{v}_i$
    $h_i = h_i + \alpha$
  end for
end if

The basis vectors are related by

$$AV_k = V_k H_k + h_{k+1,k}\mathbf{v}_{k+1}\mathbf{e}_k^T = V_{k+1}\underline{H}_k, \quad V_{k+1}^T V_{k+1} = I, \tag{1.5.3}$$

where $V_k = [\mathbf{v}_1, \ldots, \mathbf{v}_k] \in \mathbb{R}^{n\times k}$, $H_k \in \mathbb{R}^{k\times k}$, and $\underline{H}_k = [H_k^T, h_{k+1,k}\mathbf{e}_k]^T \in \mathbb{R}^{(k+1)\times k}$ are upper Hessenberg[2]. Relation (1.5.3) characterizes a k-step Arnoldi factorization. Approximate eigenpairs $(\theta_i, V_k\mathbf{y}_i)$, called Ritz pairs, can be computed from eigenpairs $(\theta_i, \mathbf{y}_i)$ of H_k. See Chapter 7, and [64, 156] and references therein for more details and efficient implementations of the Arnoldi method. If A is symmetric, it follows that H_k is a symmetric tridiagonal matrix, and the Arnoldi algorithm reduces (mathematically) to the three term recurrence known as the Lanczos method [83].

[2]Barred identifiers $\underline{H}_k$ are elements of $\mathbb{R}^{(k+1)\times k}$, whereas $H_k \in \mathbb{R}^{k\times k}$.

Consider a single interpolation point $\sigma_0 \notin \Lambda(A, E)$ and rewrite the transfer function of the square (p inputs - p outputs) system (E, A, B, C) as

$$H(s) = C^*(sE - A)^{-1}B = C^*(I + (s - \sigma_0)(\sigma_0 E - A)^{-1}E)^{-1}((\sigma_0 E - A)^{-1}B).$$

Most of the Arnoldi-based model order reduction methods use a (block) Arnoldi method [132] to construct an orthonormal basis $(\mathbf{v}_1, \ldots, \mathbf{v}_{kp})$ for the (block) Krylov space $\mathcal{K}((s_0E - A)^{-1}E, (s_0E - A)^{-1}B)$, but they differ in how the reduced-order system matrices $(\widetilde{E}, \widetilde{A}, \widetilde{B}, \widetilde{C})$ are computed. In [136] the reduced-order model is constructed as $(H_k, \sigma_0 H_k - I, V_k^* B, V_k^* C)$. This reduced-order model matches k moments, but does not preserve passivity. The Passive Reduced-Order Interconnect Macromodeling Algorithm (PRIMA) [105] constructs the reduced model as $(V_k^* E V_k, V_k^* A V_k, V_k^* B, V_k^* C)$ and is passivity preserving (if E is symmetric semi-positive definite and $B = C$, as is often the case in simulation of RLC-circuits). In [73] efficient techniques for building the bases for the block Krylov subspaces are described. Structure-preserving variants are described in [11, 55].

In [70], a model order reduction method based on a two-sided rational Arnoldi scheme is presented. In the SISO case, a reduced-order model is constructed as $(W^*EV, W^*AV, W^*\mathbf{b}, V^*\mathbf{c})$, where the columns of the $n \times k$ matrices V and W form bases for the rational Krylov subspaces

$$\sum_{j=1}^{l} \mathcal{K}^{J_j}((\sigma_j E - A)^{-1}E, (\sigma_j E - A)^{-1}\mathbf{b}),$$

and

$$\sum_{j=1}^{l} \mathcal{K}^{J_j}((\sigma_j E - A)^{-*}E^*, (\sigma_j E - A)^{-*}\mathbf{c}),$$

respectively, with $2J_j$ the number of moments to match around interpolation point σ_j, with $J_1 + \cdots + J_l = k$. Note that in contrast with the standard (single-sided) Arnoldi methods such as PRIMA, here twice as many moments are matched (k vs. $2k$), at the costs of (in total) k additional matrix vector products with $(\sigma_j E - A)^{-*}E^*$. The reader is referred to Section 3.8 of Chapter 3 and [70] for more information on two-sided methods.

1.5.5 Jacobi-Davidson, the Dominant Pole Algorithm and modal approximation

The Jacobi-Davidson method [140] combines two principles to compute eigenpairs of eigenvalue problems $A\mathbf{x} = \lambda\mathbf{x}$. The first (Davidson) principle is to apply a Ritz-Galerkin approach with respect to the search space spanned by the orthonormal columns of $V_k = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$:

$$AV_k\mathbf{s} - \theta V_k\mathbf{s} \perp \{\mathbf{v}_1, \ldots, \mathbf{v}_k\},$$

which leads to k Ritz pairs $(\theta_i, \mathbf{q}_i = V_k\mathbf{s}_i)$, where $(\theta_i, \mathbf{s}_i)$ are eigenpairs of $V_k^* A V_k$. The second (Jacobi) principle is to compute a correction $\mathbf{t}$ orthogonal to the selected

eigenvector approximation $\mathbf{q}$ (for instance, corresponding to the largest Ritz value θ) from the Jacobi-Davidson correction equation

$$(I - \mathbf{q}\mathbf{q}^*)(A - \theta I)(I - \mathbf{q}\mathbf{q}^*)\mathbf{t} = -(A\mathbf{q} - \theta\mathbf{q}).$$

The search space is expanded with (an approximation of) $\mathbf{t}$. A Ritz pair is accepted if $\|\mathbf{r}\|_2 = \|A\mathbf{q} - \theta\mathbf{q}\|_2$ is smaller than a given tolerance. A two-sided variant, called two-sided Jacobi-Davidson [74, 148], is shown in Alg. 1.5 (see Sections 3.4–3.6 of Chapter 3 for more details).

Algorithm 1.5 The two-sided Jacobi-Davidson method

INPUT: $n \times n$ matrix A, initial vectors $\mathbf{v}_1, \mathbf{w}_1$ ($\mathbf{w}_1^*\mathbf{v}_1 \neq 0$), tolerance ϵ
OUTPUT: approximate eigentriplet $(\theta, \mathbf{v}, \mathbf{w})$ with
$\min(\|A\mathbf{v} - \theta\mathbf{v}\|, \|A^*\mathbf{w} - \theta^*\mathbf{w}\|) \leq \epsilon$

1: $\mathbf{s} = \mathbf{v}_1, \mathbf{t} = \mathbf{w}_1$
2: $U_0 = V_0 = W_0 = H_0 = []$
3: **for** $i = 1, 2, \ldots$ **do**
4: $\quad \mathbf{v}_i = \mathrm{MGS}(V_{i-1}, \mathbf{s})$, $\mathbf{w}_i = \mathrm{MGS}(W_{i-1}, \mathbf{t})$ {Alg. 1.4}
5: $\quad \mathbf{v}_i = \mathbf{v}_i / \|\mathbf{v}_i\|_2$, $V_i = [V_{i-1}, \mathbf{v}_i]$
6: $\quad \mathbf{w}_i = \mathbf{w}_i / \|\mathbf{w}_i\|_2$, $W_i = [W_{i-1}, \mathbf{w}_i]$
7: $\quad \mathbf{u}_i = A\mathbf{v}_i$, $U_i = [U_{i-1}, \mathbf{u}_i]$
8: $\quad H_i = \begin{bmatrix} H_{i-1} & W_{i-1}^*\mathbf{u}_i \\ \mathbf{w}_i^*U_{i-1} & \mathbf{w}_i^*\mathbf{u}_i \end{bmatrix}$
9: $\quad$ Select suitable eigentriplet $(\theta, \mathbf{s}, \mathbf{t})$ of H_i
10: $\quad \mathbf{v} = V_i\mathbf{s}/\|V_i\mathbf{s}\|_2$, $\mathbf{w} = W_i\mathbf{t}/\|W_i\mathbf{t}\|_2$
11: $\quad \mathbf{r}_v = A\mathbf{v} - \theta\mathbf{v}$, $\mathbf{r}_w = A^*\mathbf{w} - \theta^*\mathbf{w}$
12: $\quad$ **if** $\min(\|\mathbf{r}_v\|_2, \|\mathbf{r}_w\|_2) \leq \epsilon$ **then**
13: $\quad\quad$ Stop
14: $\quad$ **end if**
15: $\quad$ Solve (approximately) $\mathbf{s} \perp \mathbf{v}$, $\mathbf{t} \perp \mathbf{w}$ from correction equations

$$\begin{aligned} (I - \frac{\mathbf{v}\mathbf{w}^*}{\mathbf{w}^*\mathbf{v}})(A - \theta I)(I - \mathbf{v}\mathbf{v}^*)\mathbf{s} &= -\mathbf{r}_v \\ (I - \frac{\mathbf{w}\mathbf{v}^*}{\mathbf{v}^*\mathbf{w}})(A - \theta I)^*(I - \mathbf{w}\mathbf{w}^*)\mathbf{t} &= -\mathbf{r}_w \end{aligned}$$

16: **end for**

The search space V_k in general is *not* a Krylov subspace, but for certain choices it can be shown to be a rational Krylov subspace (see Chapter 3 and [131]). If the correction equations are solved exactly, Jacobi-Davidson is an exact Newton method [139, 140], but one of the properties that makes Jacobi-Davidson powerful is that it is often sufficient for convergence to solve the correction equation up to moderate accuracy only, using, for instance, a (preconditioned) linear solver such as GMRES [133]. Jacobi-Davidson is especially effective for computing a number of eigenvalues near a target in the complex plane, or that satisfy a certain criterion

(such as lying in the right half-plane in stability analysis). The fact that no exact solves are required at all makes Jacobi-Davidson well-suited to large-scale standard and generalized eigenproblems. Jacobi-Davidson QR (QZ) methods, that compute partial (generalized) Schur forms for standard (generalized) eigenproblems, are described in [51, 139]. In Chapter 7 a scheme based on Jacobi-Davidson QZ is presented for the computation of right half-plane eigenvalues in the presence of eigenvalues at infinity (stability analysis), and in Chapter 8 a Jacobi-Davidson QZ variant for the computation of a partial generalized *real* Schur form is described.

The Dominant Pole Algorithm (DPA) [91] is a specialized eigenvalue method for the computation of dominant poles of a transfer function $H(s) = \mathbf{c}^*(sE - A)^{-1}\mathbf{b}$. Here, a dominant pole is a pole with a large residue (observable via peaks in the Bode plot). DPA uses the Newton method to compute the poles λ_i as zeros of the function $1/H(s)$. Starting with initial estimate s_1, a sequence of estimates is computed via

$$s_{k+1} = s_k - \frac{\mathbf{c}^*(s_kE - A)^{-1}\mathbf{b}}{\mathbf{c}^*(s_kE - A)^{-1}E(s_kE - A)^{-1}\mathbf{b}}.$$

The DPA is shown in Alg. 2.1 and its typical convergence to more dominant poles, even for initial estimates in the neighborhood of less dominant poles, is studied in detail in Chapter 2. An initial MIMO variant of DPA (MDP) is described in [93], and a block variant (DPSE) for the computation of more than one pole is presented in [90].

Although DPA is a single-pole algorithm (it computes one pole per run), different initial estimates can be used to find different dominant poles and corresponding left and right eigenvectors (with the risk of finding duplicate poles, see Chapter 3 for the multi-pole variant SADPA). Similarly, two-sided Jacobi-Davidson [74, 148] can be used to compute dominant poles and corresponding left and right eigenvectors, as is also described in Chapter 3. These dominant eigenspaces can be used for the construction of reduced-order models in the form of modal approximations. Note that the idea is to compute only the dominant eigentriplets, and not to compute a complete eigendecomposition (which is not feasible for large-scale systems). In principle, implicitly restarted Arnoldi [85, 146] and Lanczos methods can also be used for the computation of (dominant) eigentriplets. Their typical convergence to well-separated eigenvalues at the outer-edge of the spectrum, however, makes these methods less efficient than Jacobi-Davidson methods for selection criteria other than closest to a specific target [51, 139, 140].

Let (A, E) be stable and diagonalizable, and let $R_\infty = 0$. If $Y = [Y_1, Y_2] \in \mathbb{C}^{n\times n}$ and $X = [X_1, X_2] \in \mathbb{C}^{n\times n}$ are such that $Y^*AX = \Lambda = \mathrm{diag}(\Lambda_1, \Lambda_2)$ and $Y^*EX = I$ are diagonal, and ordered such that $\Lambda_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_k)$ has the k most dominant poles on its diagonal, then a modal approximation can be constructed as $(Y_1^*EX_1 = I, Y_1^*AX_1 = \Lambda_1, Y_1^*B, X_1^*C, D)$. If H_k is the transfer function of the reduced-order

model, then the truncation error becomes [68, lemma 9.2.1]

$$\begin{aligned} \|H - H_k\|_\infty &= \| \sum_{j=k+1}^{n} \frac{R_j}{s - \lambda_j} \|_\infty \\ &\leq \sum_{j=k+1}^{n} \frac{\|R_j\|_2}{|\mathrm{Re}(\lambda_j)|}, \end{aligned}$$

where $R_j = (C^* X_j)(Y_j^* B)$ are the residues. Note that

$$\|R_j\|_2 \leq \|X\|_2 \|Y\|_2 \|B\|_2 \|C\|_2,$$

and in general *all* eigenvalues and eigenvectors are needed to compute the error. However, this information is usually not available for large-scale systems. The advantages of modal truncation are that it is conceptually simple, and that the poles of the reduced-order model are also poles of the original system, so that they keep their physical interpretation as, for instance, resonance frequencies [68, p. 317] and stability is preserved.

The Dominant Pole Algorithm (DPA) plays an important role in Chapter 2 to Chapter 6, where it is studied in more detail and extended to SADPA (SAMDP) for the efficient computation of specifically the dominant spectra (and zeros) of SISO (MIMO) transfer functions, and for the construction of modal approximations.

1.6 Overview

Eigenvalues play an important role throughout this thesis. In each chapter algorithms are presented for the computation of specific eigenvalues and corresponding right (and left) eigenvectors of large sparse matrices. Furthermore, there is always a connection with model order reduction in some sense, varying from the reduction of large-scale dynamical systems, where the dominant poles are of interest, to the stability analysis of steady state solutions of discretized Navier-Stokes equations, where the rightmost eigenvalues are of importance.

Chapter 2 gives a detailed analysis of the convergence behavior of the dominant pole algorithm (DPA). For symmetric matrices, a similar method was already considered briefly by Ostrowski in 1958 [106], for the computation of a single eigenvalue, but Ostrowski brought up this iteration primarily as an introduction to the well-known Rayleigh quotient iteration (RQI), which has cubic instead of quadratic rate of convergence in the neighborhood of an eigenvalue. Both iterations use a new shift (the eigenvalue estimate) every iteration, but DPA keeps the right-hand side fixed, while RQI updates the right-hand side every iteration. Keeping the right-hand side fixed forces convergence to eigenvalues with eigenvectors in the direction of the right-hand side. The asymptotic convergence rate drops to quadratic for DPA, but in practice this only costs one or two additional iterations. The convergence behavior of DPA makes it an effective method for the computation of dominant poles of large-scale dynamical systems, and the significantly better

convergence (with respect to dominance) compared to two-sided Rayleigh quotient iteration is illustrated by numerical examples. This chapter is also available as [126]:

> Joost Rommes and Gerard L. G. Sleijpen, *Convergence of the dominant pole algorithm and Rayleigh quotient iteration*, Preprint 1356, Utrecht University, 2006,

and has been submitted for publication.

DPA is a single-pole algorithm. Given an initial estimate, it computes one dominant pole using Newton's method. In Chapter 3, DPA is extended with subspace acceleration to obtain better global convergence, and with deflation to compute more than one pole without recomputing already computed poles: subspace accelerated DPA (SADPA). Deflation can be implemented very efficiently via the fixed right-hand sides of the DPA iteration, so no explicit orthogonalizations of the search space expansion vectors are needed. Under certain conditions, SADPA is equivalent to two-sided Jacobi-Davidson and rational Krylov methods, and if the exact solves of linear systems are not feasible, inexact two-sided Jacobi-Davidson can be used to compute dominant poles. The modal approximations that can be constructed with the right and left eigenvectors of the dominant poles are compared to reduced-order models computed by rational Krylov methods, and it is shown how both methods can improve each other. This chapter is based on [123]:

> Joost Rommes and Nelson Martins, *Efficient computation of transfer function dominant poles using subspace acceleration*, IEEE Transactions on Power Systems **21** (2006), no. 3, 1218–1226,

and [120]:

> Joost Rommes, *Modal Approximation and Computation of Dominant Poles*, Model Order Reduction: Theory, Research Aspects and Applications, (H. A. van der Vorst and W. H. A. Schilders, eds.), Springer, To Appear,

but is almost completely rewritten, extended with new results, relations to Jacobi-Davidson and rational Krylov, additional numerical experiments and reflections to rational Krylov based model order reduction methods.

In Chapter 4, DPA and SADPA are generalized to algorithms for the computation of dominant poles of multi-input multi-output (MIMO) transfer functions: subspace accelerated MIMO dominant pole algorithm (SAMDP). The basic idea is the same as for (SA)DPA, but the fact that MIMO transfer functions are square or non-square matrix valued functions leads to additional algorithmic and numerical considerations. This chapter is published as [122]:

> Joost Rommes and Nelson Martins, *Efficient computation of multivariable transfer function dominant poles using subspace acceleration*, IEEE Transactions on Power Systems **21** (2006), no. 4, 1471–1483.

Chapter 5 focuses on the computation of dominant zeros of transfer functions. Like dominant poles, the dominant zeros of a transfer function are of importance for stability analysis and design of large-scale control systems. The zeros are generalized eigenvalues of a matrix pair related to the system matrices, and the dominant zeros cause the dips in the Bode magnitude plot of the transfer function. If the inverse of the transfer function exists, the zeros are equal to the poles of the inverse transfer function. Because the system matrices of the inverse transfer function are closely related to the system matrices of the original transfer function, SADPA and SAMDP can be used to compute the dominant zeros via the dominant poles of the inverse transfer function. This chapter is also available as [125]:

> Joost Rommes, Nelson Martins, and Paulo C. Pellanda, *Efficient computation of large-scale transfer function dominant zeros*, Preprint 1358, Utrecht University, 2006,

and has been submitted for publication.

In Chapter 6, the Dominant Pole Algorithm is generalized to an algorithm for the computation of dominant poles of transfer functions of second-order dynamical systems: Quadratic DPA (QDPA). In this case, the dominant poles are specific eigenvalues of a quadratic eigenvalue problem. Since QDPA works with the original system matrices, no linearization to a generalized eigenproblem is required. Furthermore, the modal approximations that are constructed using the dominant eigenspaces preserve the second-order structure of the original system. It is shown that the dominant poles can be used to improve reduced-order models computed by second-order Krylov methods [11, 12]. Generalizations to higher-order systems and MIMO systems are described, and QDPA can also be used for the computation of dominant zeros. This chapter is available as [124]:

> Joost Rommes and Nelson Martins, *Efficient computation of transfer function dominant poles of large second-order dynamical systems*, Preprint 1360, Utrecht University, 2007,

and has been submitted for publication.

In Chapter 7 algorithms based on Arnoldi and Jacobi-Davidson methods are considered for the computation of the rightmost eigenvalues of large-scale generalized eigenproblems $A\mathbf{x} = \lambda B\mathbf{x}$ with singular B. These eigenproblems arise, for instance, in stability analysis of discretized Navier-Stokes equations and large-scale power systems. Eigenvalues with positive real parts imply instability of the steady state solution and are therefore of importance. The computation of these rightmost eigenvalues is, however, complicated by the presence of eigenvalues at infinity caused by the singularity of B. These eigenvalues at infinity have no physical relevance. Standard Arnoldi and Jacobi-Davidson approaches may fail because they may interpret approximations of eigenvalues at infinity as approximations to finite eigenvalues. In this chapter, strategies and algorithms are presented for the successful computation of the finite rightmost eigenvalues. This chapter (without appendix) is also available as [121]:

> Joost Rommes, *Arnoldi and Jacobi-Davidson methods for generalized eigenvalue problems $A\mathbf{x} = \lambda B\mathbf{x}$ with singular B*, Preprint 1339, Utrecht University, 2005 (revised 2007),

and has been accepted for publication in Mathematics of Computation.

Chapter 8 presents a variant of the Jacobi-Davidson method that is specifically designed for real unsymmetric matrix pencils: real Jacobi-Davidson QZ (RJDQZ). Because the search and test space are kept purely real, this method has lower memory and computational costs than standard JDQZ. Numerical experiments confirm that also the convergence is accelerated. This chapter is published as [158]:

> Tycho van Noorden and Joost Rommes, *Computing a partial generalized real Schur form using the Jacobi-Davidson method*, Numerical Linear Algebra with Applications **14** (2007), no. 3, 197–215.

Except for Chapter 3, all chapters are available as separate papers. The notation for this thesis has been made uniform, and addenda have been added with additional remarks. Chapter 7 has been extended with an appendix.