Minimal and Orthogonal Residual Methods and their Generalizations for Solving Linear Operator Equations

Habilitationsschrift

zur Erlangung des akademischen Grades
doctor rerum naturalium habilitatus
(Dr. rer. nat. habil.)

vorgelegt

der Fakultät für Mathematik und Informatik
der TU Bergakademie Freiberg

von Dr. Oliver G. Ernst (Ph.D. Stanford University)
geboren am 11. April 1963 in Offenburg

Freiberg, den 14. März 2000

dedicated to the memory of

Dr. Rüdiger Weiss
1954–1999

Acknowledgments

Of the many who have helped me along in the work contained this thesis there are some I wish to thank expressly.

Gene Golub brought me to Stanford for graduate study, where, besides his nurturing, I enjoyed four sunny and formative years in the best learning environment a student could wish for. Gene has since then then been an unfaltering source of advice, assistance and hospitality.

I also thank my colleagues, in particular Werner Queck, Wolfgang Mönch and Olaf Schneider, at the Institut für Angewandte Mathematik II of the TU Bergakademie Freiberg, where I have led a happy existence as wissenschaftlicher Assistent for the past six years.

Howard Elman invited me for a one-year stay at the University of Maryland, where he, Dianne O'Leary and I collaborated on solving discrete Helmholtz equations. I benefitted greatly from the perfect working conditions there, as well as from many discussions with him and his colleagues, Dianne O'Leary and Pete Stewart.

The person to whom I feel most indebted is Michael Eiermann, who introduced me to the field of computational mathematics back when I was an undergraduate student at the Universität Karlsruhe. Since then he has, in the roles of mentor, co-author and co-commuter generously provided support, guidance and inspiration.

Financial support from the National Science Foundation, the University of Maryland Institute for Advanced Computer Studies, the Office of Naval Research, the Deutsche Forschungsgemeinschaft, and the Michael Jürgen Leisler Kiep Endowment are gratefully acknowledged.

Contents

Chapter 1

Introduction

This thesis is concerned with the approximate solution of linear operator equations

$$A\boldsymbol{x} = \boldsymbol{b} \tag{1.1}$$

by two specific classes of projection methods: the *minimal residual (MR)* and *orthogonal residual (OR)* methods. The right hand side $\boldsymbol{b}$ and the unknown $\boldsymbol{x}$ lie in a Hilbert space $\mathscr{H}$ with inner product $(\cdot,\cdot)$ and associated norm $\|\cdot\|$, and $A : \mathscr{H} \to \mathscr{H}$ is a bounded and invertible operator. Both methods can be viewed as *subspace correction methods* in the sense that an approximate solution of (1.1) is sought by adding a correction $\boldsymbol{c}$ from a finite-dimensional *correction space* $\mathscr{C}$ to a given initial approximation $\boldsymbol{x}_0$. The corrections are determined by imposing constraints on the *residual vector* $\boldsymbol{r} = \boldsymbol{b} - A(\boldsymbol{x}_0 + \boldsymbol{c})$, the number of which coincides with the dimension of $\mathscr{C}$.

The MR approximation $\boldsymbol{x}^{\mathrm{MR}}$ relative to the correction space $\mathscr{C}$ is defined by

$$\|\boldsymbol{b} - A\boldsymbol{x}^{\mathrm{MR}}\| = \min_{\boldsymbol{c}\in\mathscr{C}} \|\boldsymbol{b} - A(\boldsymbol{x}_0 + \boldsymbol{c})\| = \min_{\boldsymbol{c}\in\mathscr{C}} \|\boldsymbol{r}_0 - A\boldsymbol{c}\|, \tag{1.2}$$

where $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$ is the residual of the initial approximation. In other words, the residual of the MR method is simply the error of the best approximation to $\boldsymbol{r}_0$ from the space $\mathscr{W} := A\mathscr{C}$, the image of the correction space under A. For this reason we shall call $\mathscr{W}$ the *approximation space*. As is well known, the unique solution $\boldsymbol{c}^{\mathrm{MR}}$ to this problem of best approximation is such that $A\boldsymbol{c}^{\mathrm{MR}} = P_{\mathscr{W}}\boldsymbol{r}_0$, where $P_{\mathscr{W}} : \mathscr{H} \to \mathscr{W}$ denotes the orthogonal projection onto $\mathscr{W}$, and the MR residual $\boldsymbol{r}^{\mathrm{MR}} := \boldsymbol{r}_0 - A\boldsymbol{c}^{\mathrm{MR}}$ is characterized by the constraint

$$\boldsymbol{r}^{\mathrm{MR}} \perp \mathscr{W}.$$

The OR approximation is obtained by imposing, in place of (1.2), the constraint

$$\boldsymbol{r}^{\mathrm{OR}} \perp \mathscr{V}, \tag{1.3}$$

where $\mathscr{V}$ is a suitable test space. In this case, $\boldsymbol{r}_0$ is being approximated by $P^{\mathscr{V}}_{\mathscr{W}}\boldsymbol{r}_0$, where $P^{\mathscr{V}}_{\mathscr{W}} : \mathscr{H} \to \mathscr{W}$ is the *oblique projection* onto $\mathscr{W}$ orthogonal to $\mathscr{V}$. In contrast with the MR approximation, the OR approximation need not exist for all choices of $\mathscr{C}$ and $\mathscr{V}$.

These approaches of approximating elements of a large space from a smaller, more manageable one is encountered throughout applied mathematics and numerical analysis,

and such Galerkin or Least-Squares approximations as the OR and MR approximations above e.g. constitute fundamental frameworks for the approximate solution of boundary value problems (see for instance Aubin (1972), Brezzi & Fortin (1991) or Kress (1989)).

Our primary concern here is the solution of finite dimensional operator equations, i.e., linear systems of equations. We choose an abstract, possibly infinite dimensional setting, at least initially, for two reasons. First, for the algorithms to be considered, it makes no essential difference whether or not the underlying spaces have finite dimension. Second, the dimension of the correction spaces $\mathscr{C}$ to be used in practical applications is always much smaller than that of the space $\mathscr{H}$, so, in comparison, $\dim \mathscr{H}$ may as well be infinite. Furthermore, a large class of linear systems arises from discretization of operators between infinite dimensional spaces, in which case a sequence of problems corresponding to a sequence of discretization parameters is the natural object of study, and the later elements of such a sequence inherit many important properties of the infinite-dimensional problem under approximation.

By far the most widely used type of correction spaces for linear systems are *Krylov spaces*, which are subspaces spanned by successive powers of A applied to a fixed starting vector, in this case the initial residual. Historically, this predominance is explained by the ease with which such spaces can be generated using only matrix-vector products with A, particularly when A can be represented as a sparse or structured matrix, or when the action of A on a vector can be implemented as a specialized subroutine which is more efficient than standard matrix-vector multiplication.

The success of Krylov subspace methods as general-purpose iterative solution methods began with a paper by Reid (1971), in which the conjugate gradient method, a projection method for solving symmetric positive definite problems which had been introduced and analyzed much earlier by Hestenes & Stiefel (1952), is proposed as an iterative solution method for solving well-conditioned linear systems. Other key developments such as preconditioning and the extension of the conjugate gradient method to indefinite and nonsymmetric problems led to a surge of interest in and publications on Krylov subspace methods. Except for some special cases, and these will be described in the following, there is no a priori reason to expect Krylov subspaces to contain particularly good approximations to the solution of (1.1). Moreover, many Krylov subspace methods incur computational work which grows as the iteration proceeds, so that slow convergence can be doubly problematic. For this reason, the basic formulations of Krylov subspace methods are often modified to make them less expensive, resulting in *truncated* or *restarted* formulations, which, however, often converge less rapidly and less reliably than their unmodified counterparts. Several approaches have been recently proposed to enhance the approximation qualities of Krylov spaces, mainly with the goal of compensating for the deterioration due to truncation or restarting.

In this thesis we propose a new approach to deriving Krylov subspace methods based on the well-known basic scheme of MR and OR approximations. This approach has the advantage of simplicity and generality: we can express the basic idea independently of any linear systems as an approximation scheme in a Hilbert space $\mathscr{H}$ based on orthogonal and oblique projection onto a finite dimensional subspace $\mathscr{W}$. By considering this scheme on a nested sequence of subspaces $\{\mathscr{W}_m\}$, a simple assumption relating the test space $\mathscr{V}_m$ of the oblique projections with $\mathscr{W}_m$ already allows one to derive the basic relations on Krylov subspace MR and OR methods for solving (1.1). A key role is played by

the angles between the spaces $\mathscr{W}_m$ and $\mathscr{V}_m$. Moreover, upon introducing bases, abstract counterparts of well known Krylov subspace methods result as the basic alternatives for formulating these projection methods depending on the choice of bases. Specialization of this abstract approach to solving operator equations and to Krylov spaces yields the well-known Krylov subspace methods and their recent generalizations. We believe that this framework offers a natural setting in which to derive and relate the many methods which have been proposed in the past 50 years. Furthermore, as we demonstrate in Chapter 6, the angles formulation offers a new approach for generating error and residual bounds for MR and OR methods.

The thesis is organized as follows: Chapter 2 presents the abstract framework of MR and OR methods as approximation schemes in a Hilbert space. In Chapter 3 we formulate the MR and OR methods for solving equations and obtain known methods for general correction spaces such as FGMRES and GCR. We then derive formulations for successive MR approximations on two spaces and how to relax orthogonality constraints in an optimal way. Chapter 4 further specializes to Krylov spaces, and we discuss the implications of the polynomial representation which these spaces possess and give an overview of the development of the most important methods beginning with their ancestor, the conjugate gradient method. We conclude the chapter on Krylov subspace methods with a look at short recurrences, the role of the inner product and how Krylov spaces are modified by preconditioning. Chapter 5 treats truncated and restarted MR methods. After some results on possible strategies for augmenting Krylov spaces we study and compare the recently proposed methods of deflation by preconditioning, deflation by augmentation, and optimal truncation. Chpater 6 treats error and convergence bounds for MR and OR methods After a survey of the classical approaches for deriving bounds, we show how this can be done using the angle formulation. We close with some consequences on equivalent problems, for which MR and OR methods behave identically, and on the relevance of singular values for convergence.

Chapter 2

Projection Methods in Hilbert Space

As stated in the Introduction, the MR and OR approximations of the solution of (1.1) can be viewed as approximations of the initial residual $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$ by its orthogonal or oblique projection, respectively, onto an approximation space $\mathscr{W} \subset \mathscr{H}$. In this chapter we shall analyze these approximations without reference to any operator equation, only with increasingly specific choices of approximation spaces $\mathscr{W}$ and test spaces $\mathscr{V}$. More specifically, Section 2.1 treats general spaces, followed by nested sequences of spaces in Section 2.2, in which the test spaces $\mathscr{V}$ also have a specific dependence on $\mathscr{W}$. Section 2.3 is concerned with the coordinate calculations necessary to compute MR and OR approximations with regard to various convenient choices of bases of $\mathscr{W}$ and $\mathscr{V}$. So as not to lose sight of how these approximations fit into the equation-solving context, we shall refer to the best approximation of an element $\boldsymbol{r} \in \mathscr{W}$ as its *MR approximation* and its approximation by oblique projection as its *OR approximation*.

2.1 Angles and Approximation

Given an arbitrary finite dimensional subspace $\mathscr{W} \subset \mathscr{H}$ and an element $\boldsymbol{r} \in \mathscr{H}$, we define its MR approximation $\boldsymbol{w}^{\mathrm{MR}}$ as the best approximation of $\boldsymbol{r}$ from $\mathscr{W}$ and denote by $\boldsymbol{d}^{\mathrm{MR}}$ the associated approximation error:

$$\boldsymbol{w}^{\mathrm{MR}} := P_{\mathscr{W}}\boldsymbol{r}, \qquad \boldsymbol{d}^{\mathrm{MR}} := \boldsymbol{r} - \boldsymbol{w}^{\mathrm{MR}} = (I - P_{\mathscr{W}})\boldsymbol{r} \perp \mathscr{W}.$$

The distance between $\boldsymbol{r}$ and its best approximation $P_{\mathscr{W}}\boldsymbol{r}$ in $\mathscr{W}$ can be described in terms of angles between vectors and subspaces of $\mathscr{H}$. The angle $\measuredangle(\boldsymbol{x}, \boldsymbol{y})$ between two nonzero vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathscr{H}$ is defined by the relation

$$\cos \measuredangle(\boldsymbol{x}, \boldsymbol{y}) := \frac{|(\boldsymbol{x}, \boldsymbol{y})|}{\|\boldsymbol{x}\| \, \|\boldsymbol{y}\|},$$

which, in view of the Cauchy-Schwarz inequality, uniquely determines $\measuredangle(\boldsymbol{x}, \boldsymbol{y})$ as a number in $[0, \pi/2]$. We note here that the natural definition of the angle would replace the modulus in the numerator by its real part; our definition, however, is more appropriate for comparing subspaces (see also the discussion of this issue in Davis & Kahan (1970, p. 9)). Similarly, we define the angle between a nonzero vector $\boldsymbol{x} \in \mathscr{H}$ and a subspace

$\mathscr{U} \subset \mathscr{H}$, $\mathscr{U} \neq \{\mathbf{0}\}$, by

$$\measuredangle(\boldsymbol{x}, \mathscr{U}) := \inf_{\mathbf{0}\neq\boldsymbol{u}\in\mathscr{U}} \measuredangle(\boldsymbol{x}, \boldsymbol{u}), \quad \text{i.e.,} \quad \cos\measuredangle(\boldsymbol{x}, \mathscr{U}) = \sup_{\mathbf{0}\neq\boldsymbol{u}\in\mathscr{U}} \cos\measuredangle(\boldsymbol{x}, \boldsymbol{u}). \tag{2.1}$$

Correspondingly, the sine of this angle is $\sin\measuredangle(\boldsymbol{x}, \mathscr{U}) = \sqrt{1 - \cos^2\measuredangle(\boldsymbol{x}, \mathscr{U})}$.

The angle between a vector and a subspace may be expressed in terms of the orthogonal projection onto that space, as shown in the following Lemma.

Lemma 2.1.1. *Let $\mathscr{U}$ be a finite dimensional subspace of $\mathscr{H}$ and denote by $P_{\mathscr{U}}$ the orthogonal projection onto $\mathscr{U}$. For each $\boldsymbol{x} \in \mathscr{H}$ there holds*

$$\measuredangle(\boldsymbol{x}, \mathscr{U}) = \measuredangle(\boldsymbol{x}, P_{\mathscr{U}}\boldsymbol{x}) \tag{2.2}$$

and, as a consequence,

$$\|P_{\mathscr{U}}\boldsymbol{x}\| = \|\boldsymbol{x}\| \cos\measuredangle(\boldsymbol{x}, \mathscr{U}), \tag{2.3}$$
$$\|(I - P_{\mathscr{U}})\boldsymbol{x}\| = \|\boldsymbol{x}\| \sin\measuredangle(\boldsymbol{x}, \mathscr{U}). \tag{2.4}$$

Proof. For any $\boldsymbol{u} \in \mathscr{U}$, we have $|(\boldsymbol{x}, \boldsymbol{u})| = |(P_{\mathscr{U}}\boldsymbol{x}, \boldsymbol{u})| \leq \|P_{\mathscr{U}}\boldsymbol{x}\|\,\|\boldsymbol{u}\|$, and therefore, for $\boldsymbol{u} \neq \mathbf{0}$,

$$\frac{|(\boldsymbol{x}, \boldsymbol{u})|}{\|\boldsymbol{x}\|\|\boldsymbol{u}\|} \leq \frac{\|P_{\mathscr{U}}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = \frac{|(P_{\mathscr{U}}\boldsymbol{x}, P_{\mathscr{U}}\boldsymbol{x})|}{\|\boldsymbol{x}\|\|P_{\mathscr{U}}\boldsymbol{x}\|} = \frac{|(\boldsymbol{x}, P_{\mathscr{U}}\boldsymbol{x})|}{\|\boldsymbol{x}\|\|P_{\mathscr{U}}\boldsymbol{x}\|},$$

from which (2.2) follows after taking suprema. Equation (2.3) is merely a reformulation of (2.2), and (2.4) follows from $\|\boldsymbol{x}\|^2 = \|P_{\mathscr{U}}\boldsymbol{x}\|^2 + \|(I - P_{\mathscr{U}})\boldsymbol{x}\|^2$. □

In light of this fact, the distance between $\boldsymbol{r}$ and and its MR approximation $\boldsymbol{w}^{\mathrm{MR}}$ may be expressed as

$$\|\boldsymbol{d}^{\mathrm{MR}}\| = \|\boldsymbol{r} - \boldsymbol{w}^{\mathrm{MR}}\| = \|(I - P_{\mathscr{W}})\boldsymbol{r}\| = \|\boldsymbol{r}\| \sin\measuredangle(\boldsymbol{r}, \mathscr{W}). \tag{2.5}$$

To define the OR approximation in this abstract setting we require a further finite dimensional subspace $\mathscr{V} \subset \mathscr{H}$ to formulate the orthogonality constraint. The OR approximation $\boldsymbol{w}^{\mathrm{OR}} \in \mathscr{W}$ of $\boldsymbol{r}$ is then defined by the requirement that it satisfy the orthogonality or *Galerkin* condition

$$\boldsymbol{w}^{\mathrm{OR}} \in \mathscr{W}, \qquad \boldsymbol{r} - \boldsymbol{w}^{\mathrm{OR}} \perp \mathscr{V}. \tag{2.6}$$

Of course, since choosing $\mathscr{V} = \mathscr{W}$ yields the MR approximation, the latter is just a special case of the OR approximation. We choose nonetheless to distinguish the two, both for historical reasons and for ease of exposition. Existence and uniqueness of $\boldsymbol{w}^{\mathrm{OR}}$ are summarized in

Proposition 2.1.2. *If $\mathscr{V}$ and $\mathscr{W}$ are subspaces of the Hilbert space $\mathscr{H}$ and $\boldsymbol{r} \in \mathscr{H}$, then:*

1. *There exists $\boldsymbol{w} \in \mathscr{W}$ such that $\boldsymbol{r} - \boldsymbol{w} \perp \mathscr{V}$ if and only if $\boldsymbol{r} \in \mathscr{W} + \mathscr{V}^{\perp}$.*
2. *Such a $\boldsymbol{w}$ is unique if and only if $\mathscr{W} \cap \mathscr{V}^{\perp} = \{\mathbf{0}\}$.*

Thus, a unique OR approximation is defined whenever $\mathscr{H} = \mathscr{W} \oplus \mathscr{V}^\perp$. In this case $\boldsymbol{w}^{\mathrm{OR}}$ is the *oblique* projection of $\boldsymbol{r}$ onto $\mathscr{W}$ *orthogonal to* $\mathscr{V}$ (or, equivalently, along $\mathscr{V}^\perp$), which we denote by $P_{\mathscr{W}}^{\mathscr{V}} : \mathscr{H} \to \mathscr{W}$, and $\boldsymbol{w}^{\mathrm{OR}}$ is characterized by

$$\boldsymbol{w}^{\mathrm{OR}} = P_{\mathscr{W}}^{\mathscr{V}} \boldsymbol{r}, \qquad \boldsymbol{d}^{\mathrm{OR}} = \boldsymbol{r} - \boldsymbol{w}^{\mathrm{OR}} = (I - P_{\mathscr{W}}^{\mathscr{V}})\boldsymbol{r} \perp \mathscr{V}.$$

When it exists, the oblique projection $P_{\mathscr{W}}^{\mathscr{V}}$, coincides with the Moore-Penrose pseudoinverse $(P_{\mathscr{V}} P_{\mathscr{W}})^+$ of $P_{\mathscr{V}} P_{\mathscr{W}}$ (cf. Wedin (1983)). This is established in the following two lemmas, the first of which describes the mapping properties of $(P_{\mathscr{V}} P_{\mathscr{W}})^+$.

Lemma 2.1.3. *Given two finite dimensional subspaces $\mathscr{V}$ and $\mathscr{W}$ of the Hilbert space $\mathscr{H}$, let $S := (P_{\mathscr{V}} P_{\mathscr{W}})^+$ denote the Moore-Penrose pseudoinverse of the product of the orthogonal projections onto $\mathscr{W}$ and $\mathscr{V}$. Then S is a projection and there holds*

$$\mathscr{R}(S) = \mathscr{W} \cap (\mathscr{V} + \mathscr{W}^\perp) \tag{2.7}$$

$$\mathscr{N}(S) = \mathscr{V}^\perp + (\mathscr{W}^\perp \cap \mathscr{V}). \tag{2.8}$$

Proof. The proof of the projection property $S^2 = S$ follows along the lines of Greville (1974, Theorem 1). First, since $P := P_{\mathscr{V}} P_{\mathscr{W}}$ has finite rank, its Moore-Penrose pseudoinverse $P^+ = S$ exists and satisfies $\mathscr{R}(P^+) = \mathscr{R}(P^*)$—the asterisk denoting the Hilbert space adjoint—, from which it follows that

$$\mathscr{R}(S) \subset \mathscr{R}(P_{\mathscr{W}}), \qquad \mathscr{R}(S^*) \subset \mathscr{R}(P_{\mathscr{V}}),$$

and hence, by the idempotency of $P_{\mathscr{V}}$ and $P_{\mathscr{W}}$, there holds

$$P_{\mathscr{W}} S = S, \qquad S P_{\mathscr{V}} = S,$$

which together imply $S^2 = S P_{\mathscr{V}} P_{\mathscr{W}} S = P^+ P P^+ = P^+ = S$.

Since $\mathscr{R}(S) = \mathscr{R}((P_{\mathscr{V}} P_{\mathscr{W}})^*) = \mathscr{R}(P_{\mathscr{W}} P_{\mathscr{V}})$ and analogously for the null space, it is sufficient to show (2.7) and (2.8) for the operator $P_{\mathscr{W}} P_{\mathscr{V}}$ in place of S. To derive (2.7), note that $\boldsymbol{w} \in \mathscr{W}$ lies in $\mathscr{R}(P_{\mathscr{W}} P_{\mathscr{V}})$ if and only if there exists $\boldsymbol{v} \in \mathscr{V}$ such that $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{w}_\perp$ for some $\boldsymbol{w}_\perp \in \mathscr{W}^\perp$, which in turn is equivalent with $\boldsymbol{w} \in \mathscr{W} \cap (\mathscr{V} + \mathscr{W}^\perp)$. To see (2.8), note that any $\boldsymbol{x} \in \mathscr{H}$ may be written as $\boldsymbol{x} = (\boldsymbol{w} + \boldsymbol{w}_\perp) + \boldsymbol{v}_\perp$ with $\boldsymbol{w} \in \mathscr{W}$, $\boldsymbol{w}_\perp \in \mathscr{W}^\perp$, $\boldsymbol{v}_\perp \in \mathscr{V}^\perp$, and $\boldsymbol{w} + \boldsymbol{w}_\perp \in \mathscr{V}$. Thus, $\boldsymbol{x} \in \mathscr{N}(P_{\mathscr{W}} P_{\mathscr{V}})$ if and only if $\boldsymbol{w} = \boldsymbol{0}$ or, equivalently, $\boldsymbol{x} \in \mathscr{V}^\perp + (\mathscr{W}^\perp \cap \mathscr{V})$. □

Lemma 2.1.4. *Given two finite dimensional subspaces $\mathscr{V}$ and $\mathscr{W}$ of the Hilbert space $\mathscr{H}$ such that $\mathscr{H} = \mathscr{W} \oplus \mathscr{V}^\perp$, the oblique projection $P_{\mathscr{W}}^{\mathscr{V}}$ onto $\mathscr{W}$ orthogonal to $\mathscr{V}$ is given by*

$$P_{\mathscr{W}}^{\mathscr{V}} = (P_{\mathscr{V}} P_{\mathscr{W}})^+. \tag{2.9}$$

Proof. By the previous lemma, S is a projection and, in view of $\mathscr{V} + \mathscr{W}^\perp = (\mathscr{V}^\perp \cap \mathscr{W})^\perp$, $\mathscr{W}^\perp \cap \mathscr{V} = (\mathscr{W} + \mathscr{V}^\perp)^\perp$, and $\mathscr{H} = \mathscr{W} \oplus \mathscr{V}^\perp$, its range is $\mathscr{W}$ and its null space is $\mathscr{V}^\perp$, which characterizes it as $P_{\mathscr{W}}^{\mathscr{V}}$. □

Remark 2.1.5. We note that, while the left hand side of (2.9) exists only under the condition $\mathscr{H} = \mathscr{W} \oplus \mathscr{V}^\perp$, the right hand side is always defined. Thus, $(P_{\mathscr{V}} P_{\mathscr{W}})^+ \boldsymbol{r}$ can be viewed as a natural way of defining a *generalized* OR approximation in those cases where this direct sum condition fails to hold. This situation, in which the Galerkin condition (2.6) fails to specify a unique approximation, is referred to as a *Galerkin breakdown* in the literature on Krylov subspace methods.

As is to be expected, the error $\boldsymbol{d}^{\mathrm{OR}} = \boldsymbol{r} - P_{\mathscr{W}}^{\mathscr{V}}\boldsymbol{r}$ depends on the angles between the subspaces $\mathscr{V}$ and $\mathscr{W}$, which we introduce as follows (cf. Golub & van Loan (1989)): Given two finite dimensional subspaces $\mathscr{V}$ and $\mathscr{W}$ of $\mathscr{H}$, let $m := \min(\dim\mathscr{V}, \dim\mathscr{W})$. The *canonical angles* $\{\theta_j\}_{j=1}^m$ between $\mathscr{V}$ and $\mathscr{W}$ are defined recursively by

$$\cos\theta_j := \max_{\mathbf{0}\neq\boldsymbol{v}\in\mathscr{V}} \max_{\mathbf{0}\neq\boldsymbol{w}\in\mathscr{W}} \frac{|(\boldsymbol{v},\boldsymbol{w})|}{\|\boldsymbol{v}\|\,\|\boldsymbol{w}\|} =: \frac{|(\boldsymbol{v}_j,\boldsymbol{w}_j)|}{\|\boldsymbol{v}_j\|\,\|\boldsymbol{w}_j\|}$$

subject to $\boldsymbol{v}\perp\boldsymbol{v}_1,\ldots,\boldsymbol{v}_{j-1}$ and $\boldsymbol{w}\perp\boldsymbol{w}_1,\ldots,\boldsymbol{w}_{j-1}$. We further define *the* angle between the spaces $\mathscr{V}$ and $\mathscr{W}$ as the largest of the canonical angles

$$\measuredangle(\mathscr{V},\mathscr{W}) := \theta_m.$$

Remark 2.1.6. If $P_{\mathscr{V}}P_{\mathscr{W}} = \sum_{j=1}^m \sigma_j(\cdot,\boldsymbol{w}_j)\boldsymbol{v}_j$ is a singular value decomposition of $P_{\mathscr{V}}P_{\mathscr{W}}$, then the variational characterization of the singular values,

$$\sigma_j(P_{\mathscr{V}}P_{\mathscr{W}}) = \max_{\boldsymbol{v}\in\mathscr{V},\boldsymbol{w}\in\mathscr{W}} \frac{|(P_{\mathscr{V}}P_{\mathscr{W}}\boldsymbol{w},\boldsymbol{v})|}{\|\boldsymbol{w}\|\,\|\boldsymbol{v}\|} =: \frac{|(P_{\mathscr{V}}P_{\mathscr{W}}\boldsymbol{w}_j,\boldsymbol{v}_j)|}{\|\boldsymbol{w}_j\|\,\|\boldsymbol{v}_j\|}$$

subject to $\boldsymbol{v}\perp\boldsymbol{v}_1,\ldots\boldsymbol{v}_{j-1}$, $\boldsymbol{w}\perp\boldsymbol{w}_1,\ldots,\boldsymbol{w}_{j-1}$ for $j=1,\ldots,m$ (cf. Björck & Golub (1973)), shows immediately that $\cos\theta_j = \sigma_j$. Furthermore, we note that, given any two orthonormal bases $\{\boldsymbol{v}_j\}_{j=1}^{\dim\mathscr{V}}$ and $\{\boldsymbol{w}_j\}_{j=1}^{\dim\mathscr{W}}$ of $\mathscr{V}$ and $\mathscr{W}$, then the cosines of the canonical angles are the singular values of the matrix of inner products $[(\boldsymbol{v}_j,\boldsymbol{w}_k)]_{j=1,\ldots,\dim\mathscr{V},k=1\ldots,\dim\mathscr{W}}$ (see for instance Chatelin (1993, p. 5)) or Björck & Golub (1973)).

Remark 2.1.7. As a consequence of Remark 2.1.6, we see that $S = (P_{\mathscr{V}}P_{\mathscr{W}})^+$ can be written as

$$S = \sum_{j=1}^m \sigma_j^+(\cdot,\boldsymbol{v}_j)\boldsymbol{w}_j \quad \text{with} \quad \sigma_j^+ := \begin{cases} 1/\sigma_j & \text{if } \sigma_j > 0, \\ 0 & \text{otherwise.} \end{cases}$$

In particular, we have

$$\mathscr{R}(S) = \operatorname{span}\{\boldsymbol{w}_j : \sigma_j > 0\}, \qquad \mathscr{N}(S) = \operatorname{span}\{\boldsymbol{v}_j : \sigma_j > 0\}^\perp.$$

Thus $\mathscr{R}(S) = \mathscr{W}$ if and only if $\dim(\mathscr{W}) = m$ and $\sigma_j > 0$ for all $j = 1,\ldots,m$. Similarly, $\mathscr{N}(S) = \mathscr{V}^\perp$ if and only if $\dim(\mathscr{V}) = m$ and $\sigma_j > 0$ for all $j = 1,\ldots,m$. Consequently, the oblique projection $P_{\mathscr{W}}^{\mathscr{V}}$ exists if and only if $\dim(\mathscr{V}) = \dim(\mathscr{W})$ and $\measuredangle(\mathscr{V},\mathscr{W}) < \pi/2$.

Remark 2.1.8. For completeness, we note that the sine of the angle between two equidimensional subspaces $\mathscr{V}$ and $\mathscr{W}$ is also given by $\sin\measuredangle(\mathscr{V},\mathscr{W}) = \|P_{\mathscr{V}} - P_{\mathscr{W}}\|$ (cf. Chatelin (1993)), hence an equivalent characterization of the existence of the oblique projection is $\sin\measuredangle(\mathscr{V},\mathscr{W}) = \|P_{\mathscr{V}} - P_{\mathscr{W}}\| < 1$.

Besides the relative position of the spaces $\mathscr{V}$ and $\mathscr{W}$, the error of the OR approximation also depends on the position of $\boldsymbol{r}$ with respect to $\mathscr{V}$ and $\mathscr{W}$. In this generality, all we can do to bound the OR approximation error is determine the norm of the complementary projection $I - P_{\mathscr{W}}^{\mathscr{V}}$. For simplicity, since $\mathscr{H} = \mathscr{W}\oplus\mathscr{V}^\perp$ for finite dimensional $\mathscr{V}$ and $\mathscr{W}$ implies $\dim\mathscr{V} = \dim\mathscr{W}$, we assume that both spaces have the same dimension $m < \infty$.

Lemma 2.1.9. *Given two m-dimensional subspaces $\mathscr{V}, \mathscr{W} \subset \mathscr{H}$ of the Hilbert space $\mathscr{H}$ such that $\mathscr{H} = \mathscr{W} \oplus \mathscr{V}^\perp$, let $P_{\mathscr{W}}^{\mathscr{V}} : \mathscr{H} \to \mathscr{W}$ denote the oblique projection onto $\mathscr{W}$ orthogonal to $\mathscr{V}$. Then there holds*

$$\|I - P_{\mathscr{W}}^{\mathscr{V}}\| = \frac{1}{\cos \measuredangle(\mathscr{V}, \mathscr{W})}. \tag{2.10}$$

Proof. The proof follows from the fact that $\|P\| = \|I - P\|$ for any projection operator $P \neq O, P \neq I$ in a Hilbert space (see Kato (1960)). Thus,

$$\|I - (P_{\mathscr{V}} P_{\mathscr{W}})^+\| = \|(P_{\mathscr{V}} P_{\mathscr{W}})^+\| = \sigma_{\max}\big((P_{\mathscr{V}} P_{\mathscr{W}})^+\big) = \frac{1}{\sigma_{\min}(P_{\mathscr{V}} P_{\mathscr{W}})}$$

and the assertion follows from Remark 2.1.6. □

Without further assumptions on $\boldsymbol{r}$ and the spaces $\mathscr{V}$ and $\mathscr{W}$, all we can say about the error of the OR approximation is

$$\|\boldsymbol{d}^{\mathrm{OR}}\| = \|\boldsymbol{r} - \boldsymbol{w}^{\mathrm{OR}}\| = \|(I - P_{\mathscr{W}}^{\mathscr{V}})\boldsymbol{r}\| \leq \frac{\|\boldsymbol{r}\|}{\cos \measuredangle(\mathscr{V}, \mathscr{W})}.$$

As an immediate consequence, noting that

$$\boldsymbol{d}^{\mathrm{OR}} = (I - P_{\mathscr{W}}^{\mathscr{V}})\boldsymbol{r} = (I - P_{\mathscr{W}}^{\mathscr{V}})(\boldsymbol{r} - \boldsymbol{w}) \qquad \forall \boldsymbol{w} \in \mathscr{W},$$

we obtain

$$\|\boldsymbol{d}^{\mathrm{OR}}\| \leq \|I - P_{\mathscr{W}}^{\mathscr{V}}\| \inf_{\boldsymbol{w} \in \mathscr{W}} \|\boldsymbol{r} - \boldsymbol{w}\| = \|I - P_{\mathscr{W}}^{\mathscr{V}}\| \, \|\boldsymbol{d}^{\mathrm{MR}}\|,$$

an estimate usually referred to as *Céa's Lemma* in the context of variational discretization methods (see for example Brenner & Scott (1994, Theorem 2.8.2)). Together with (2.10), this implies

$$\cos \measuredangle(\mathscr{V}, \mathscr{W}) \|\boldsymbol{d}^{\mathrm{OR}}\| \leq \|\boldsymbol{d}^{\mathrm{MR}}\| \leq \|\boldsymbol{d}^{\mathrm{OR}}\|.$$

Notes and Remarks

The angles between two subspaces of $\mathbb{R}^n$ were treated systematically as early as 1875 in a paper by Jordan (1875). Abstract treatments of the separation of a pair of subspaces in the context of the perturbation of eigenspaces may be found in Riesz & Sz.-Nagy (1955, Section 136) and Kato (1980, Sections 1.4.6 and 1.6.8). The related concept of direct rotation between two subspaces was introduced by Davis (Davis 1963, Davis 1965) and further developed to study eigenvector perturbations in the celebrated paper by Davis & Kahan (1970). The direct rotation was also a key step in the development of the CS decomposition by Stewart (1977), which has since become an increasingly popular matrix analysis tool, see Paige & Wei (1994) for a survey. The computation of the angles between subspaces using the singular value decomposition was introduced by Björck & Golub (1973). Lemma 2.1.9 was shown by Saad (1982).

2.2 Projections onto Nested Subspaces

Since until now nothing further was assumed to relate $\mathscr{V}$, $\mathscr{W}$ and $\boldsymbol{r}$, we were only able to bound the error of the OR approximation in terms of the largest canonical angle between the spaces $\mathscr{V}$ and $\mathscr{W}$. We will obtain more interesting results in this section by selecting a specific space $\mathscr{V}$ which differs only slightly from $\mathscr{W}$. This choice, however, is still general enough to contain all Krylov subspace methods. Moreover, we now consider the MR and OR approximations on nested sequences of subspaces, which is the setting in which these approximations are used in practical algorithms.

2.2.1 MR Approximations on Nested Subspaces

Consider a sequence of nested subspaces

$$\{\mathbf{0}\} = \mathscr{W}_0 \subseteq \mathscr{W}_1 \subseteq \cdots \subseteq \mathscr{W}_{m-1} \subseteq \mathscr{W}_m \subseteq \cdots \tag{2.11}$$

of $\mathscr{H}$ and assume for simplicity that $\dim \mathscr{W}_m = m$. Throughout this section, $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ will always denote an orthonormal basis of $\mathscr{W}_m$ such that $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{m-1}\}$ forms a basis of $\mathscr{W}_{m-1}$.

In terms of such a basis, the MR approximation, i.e., the best approximation of $\boldsymbol{r} \in \mathscr{H}$ from $\mathscr{W}_m$, can be expressed as the truncated Fourier expansion

$$\boldsymbol{w}_m^{\mathrm{MR}} = P_{\mathscr{W}_m} \boldsymbol{r} = \sum_{j=1}^{m} (\boldsymbol{r}, \boldsymbol{w}_j) \boldsymbol{w}_j,$$

so that the norm of the associated approximation error $\boldsymbol{d}_m^{\mathrm{MR}} = \boldsymbol{r} - \boldsymbol{w}_m^{\mathrm{MR}}$ is given by

$$\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2 = \|(I - P_{\mathscr{W}_m})\boldsymbol{r}\|^2 = \|\boldsymbol{r}\|^2 - \sum_{j=1}^{m} |(\boldsymbol{r}, \boldsymbol{w}_j)|^2. \tag{2.12}$$

To relate the approximations on successive spaces, we note that for every $m \geq 1$ there holds

$$\boldsymbol{w}_m^{\mathrm{MR}} = \sum_{j=1}^{m} (\boldsymbol{r}, \boldsymbol{w}_j) \boldsymbol{w}_j = \boldsymbol{w}_{m-1}^{\mathrm{MR}} + (\boldsymbol{r}, \boldsymbol{w}_m) \boldsymbol{w}_m = \boldsymbol{w}_{m-1}^{\mathrm{MR}} + P_{\mathscr{W}_m} \boldsymbol{d}_{m-1}^{\mathrm{MR}}.$$

It follows that $\boldsymbol{d}_m^{\mathrm{MR}} = \boldsymbol{d}_{m-1}^{\mathrm{MR}} - P_{\mathscr{W}_m} \boldsymbol{d}_{m-1}^{\mathrm{MR}} = (I - P_{\mathscr{W}_m}) \boldsymbol{d}_{m-1}^{\mathrm{MR}}$ and, since $\boldsymbol{d}_m^{\mathrm{MR}} \perp P_{\mathscr{W}_m} \boldsymbol{d}_{m-1}^{\mathrm{MR}} \in \mathscr{W}_m$,

$$\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2 = \|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2 - \|P_{\mathscr{W}_m} \boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2 = \|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2 - |(\boldsymbol{r}, \boldsymbol{w}_m)|^2. \tag{2.13}$$

Relation (2.13) shows that no improvement is obtained in the MR approximation in step m whenever the direction in which $\mathscr{W}_{m-1}$ is enlarged is orthogonal to $\boldsymbol{r}$. In other words,

$$\|\boldsymbol{d}_m^{\mathrm{MR}}\| < \|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\| \quad \text{if and only if} \quad (\boldsymbol{r}, \boldsymbol{w}_m) \neq 0. \tag{2.14}$$

To express successive approximation errors in terms of angles, we note that $\boldsymbol{d}_m^{\mathrm{MR}} = (I - P_{\mathscr{W}_m}) \boldsymbol{d}_{m-1}^{\mathrm{MR}}$ together with (2.4) yields

$$\|\boldsymbol{d}_m^{\mathrm{MR}}\| = \|(I - P_{\mathscr{W}_m}) \boldsymbol{d}_{m-1}^{\mathrm{MR}}\| = s_m \, \|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|, \tag{2.15}$$

where $s_m := \sin \measuredangle(\boldsymbol{d}_{m-1}^{\mathrm{MR}}, \mathscr{W}_m)$. Note that the sine s_m is also given by (cf. (2.5))

$$s_m = \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|} = \frac{\|(I - P_{\mathscr{W}_m})\boldsymbol{r}\|}{\|(I - P_{\mathscr{W}_{m-1}})\boldsymbol{r}\|} = \frac{\sin \measuredangle(\boldsymbol{r}, \mathscr{W}_m)}{\sin \measuredangle(\boldsymbol{r}, \mathscr{W}_{m-1})}, \tag{2.16}$$

i.e., s_m is the sine of the angle between the previous approximation error and $\mathscr{W}_m$ or, equivalently, the quotient of the sines of the angles between $\boldsymbol{r}$ and the current and previous approximation spaces. In order for the last three terms in (2.16) to make sense, we assume that $\boldsymbol{r} \notin \mathscr{W}_{m-1}$; otherwise, the approximation problem is solved exactly in the space $\mathscr{W}_{m-1}$ and the larger spaces no longer contribute toward improving the approximation.

In view of (2.13), the corresponding cosine is given by

$$c_m := \sqrt{1 - s_m^2} = \sqrt{1 - \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2}} = \frac{|(\boldsymbol{r}, \boldsymbol{w}_m)|}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|} \tag{2.17}$$

and we see that an equivalent statement of (2.14) is

$$\|\boldsymbol{d}_m^{\mathrm{MR}}\| < \|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\| \quad \text{if and only if} \quad c_m \neq 0. \tag{2.18}$$

An obvious induction applied to (2.15) leads to the error formula

$$\|\boldsymbol{d}_m^{\mathrm{MR}}\| = s_1 s_2 \cdots s_m \, \|\boldsymbol{r}\|, \tag{2.19}$$

which shows that the sequence of approximations will converge to $\boldsymbol{r}$ if and only if the product of sines tends to zero. Moreover, whenever the numbers s_m themselves tend to zero the convergence of the approximations is superlinear.

2.2.2 OR Approximations on Nested Subspaces

To define the OR approximations with respect to the nested sequence of spaces $\{\mathscr{W}_m\}_{m\geq 0}$, we fix the sequence $\{\mathscr{V}_m\}_{m\geq 1}$ of test spaces which define the orthogonality condition (2.6) by setting

$$\mathscr{V}_m := \operatorname{span}\{\boldsymbol{r}\} + \mathscr{W}_{m-1}, \qquad m = 1, 2, \ldots. \tag{2.20}$$

Since the m-th OR and MR approximations lie in $\mathscr{W}_m$, the corresponding approximation errors lie in $\mathscr{V}_{m+1}$ as defined by (2.20). For this reason we shall refer to $\{\mathscr{V}_m\}_{m\geq 1}$ as the sequence of *error spaces*.

We first investigate the question of when the OR approximation is well-defined. In view of Remark 2.1.7, this amounts to checking whether the largest canonical angle between $\mathscr{V}_m$ and $\mathscr{W}_m$ is strictly less than $\pi/2$. As a consequence of the special choice (2.20) of $\mathscr{V}_m$, it turns out that this angle is simply that between $\boldsymbol{d}_{m-1}^{\mathrm{MR}}$ and $\mathscr{W}_m$:

Theorem 2.2.1. *If the spaces $\mathscr{V}_m$ and $\mathscr{W}_m$ are related as in* (2.20)*, then the largest canonical angle between them is given by*

$$\measuredangle(\mathscr{V}_m, \mathscr{W}_m) = \measuredangle(\boldsymbol{d}_{m-1}^{MR}, \mathscr{W}_m), \tag{2.21}$$

while the remaining $m-1$ canonical angles between $\mathscr{V}_m$ and $\mathscr{W}_m$ are zero.

Proof. Noting that $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{m-1}, \widehat{\boldsymbol{w}}_m\}$ with $\widehat{\boldsymbol{w}}_m = \boldsymbol{d}_{m-1}^{\mathrm{MR}}/\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|$ is an orthonormal basis of $\mathscr{V}_m$, we obtain the cosine of the largest canonical angle as the smallest singular value of the matrix of inner products

$$\begin{bmatrix} (\boldsymbol{w}_1, \boldsymbol{w}_1) & \ldots & (\boldsymbol{w}_{m-1}, \boldsymbol{w}_1) & (\boldsymbol{w}_m, \boldsymbol{w}_1) \\ \vdots & & \vdots & \vdots \\ (\boldsymbol{w}_1, \boldsymbol{w}_{m-1}) & \ldots & (\boldsymbol{w}_{m-1}, \boldsymbol{w}_{m-1}) & (\boldsymbol{w}_m, \boldsymbol{w}_{m-1}) \\ (\boldsymbol{w}_1, \widehat{\boldsymbol{w}}_m) & \ldots & (\boldsymbol{w}_{m-1}, \widehat{\boldsymbol{w}}_m) & (\boldsymbol{w}_m, \widehat{\boldsymbol{w}}_m) \end{bmatrix} = \begin{bmatrix} I_{m-1} & \boldsymbol{0} \\ \boldsymbol{0} & \frac{(\boldsymbol{w}_m, \boldsymbol{r})}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|} \end{bmatrix}$$

(cf. Remark 2.1.6). We see that the smallest singular value is $|(\boldsymbol{w}_m, \boldsymbol{r})|/\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\| = c_m$ (cf. (2.17)) and that all remaining singular values are equal to one. □

As an immediate consequence of Theorem 2.2.1, we obtain the following characterization of when the oblique projection is defined for our special choice of $\mathscr{V}_m$:

Corollary 2.2.2. *The OR approximation of arbitrary $\boldsymbol{r} \in \mathscr{H}$ with respect to the sequence of spaces $\mathscr{V}_m$ and $\mathscr{W}_m$ as given by* (2.11) *and* (2.20) *is uniquely defined if and only if $(\boldsymbol{r}, \boldsymbol{w}_m) \neq 0$, i.e., if and only if the MR approximation improves as $\mathscr{W}_{m-1}$is enlarged to $\mathscr{W}_m$.*

Thus, the degenerate case, in which the oblique projection $P_{\mathscr{W}_m}^{\mathscr{V}_m}$ is not uniquely defined for all $\boldsymbol{h} \in \mathscr{H}$ is characterized by $c_m = 0$.

For the OR approximation it is not necessary for $P_{\mathscr{W}_m}^{\mathscr{V}_m}$ to be uniquely defined for all $\boldsymbol{h} \in \mathscr{H}$, merely for $\boldsymbol{r}$ occurring in the definition (2.20) of $\mathscr{V}_m$. As shown next, these two requirements are equivalent.

Theorem 2.2.3. *The OR approximation $P_{\mathscr{W}_m}^{\mathscr{V}_m}\boldsymbol{r}$ is uniquely determined if and only if $c_m \neq 0$.*

Proof. Sufficiency of $c_m = 0$ follows from Theorem 2.2.1. To see that uniqueness of the OR approximation implies $c_m = 0$, assume that $\boldsymbol{w}, \widetilde{\boldsymbol{w}} \in \mathscr{W}_m$ such that both $\boldsymbol{r} - \boldsymbol{w} \perp \mathscr{V}_m$ and $\boldsymbol{r} - \widetilde{\boldsymbol{w}} \perp \mathscr{V}_m$. This implies that also $\boldsymbol{w} - \widetilde{\boldsymbol{w}} \perp \mathscr{V}_m$ and, from the orthogonal decomposition $\mathscr{V}_m = \mathscr{W}_{m-1} \oplus \mathrm{span}\{\boldsymbol{d}_{m-1}^{\mathrm{MR}}\}$, we see that $\boldsymbol{w} - \widetilde{\boldsymbol{w}}$ is only guaranteed to be zero if $\boldsymbol{d}_{m-1}^{\mathrm{MR}} \not\perp \mathscr{W}_m$, i.e., if $c_m \neq 0$. To see that existence of the OR approximation implies $c_m \neq 0$, note that the former is equivalent with $\boldsymbol{r} \in \mathscr{W}_m + \mathscr{V}_m^{\perp} = (\mathscr{W}_m^{\perp} \cap \mathscr{V}_m)^{\perp}$. The orthogonal decomposition of $\mathscr{V}_m$ shows that the space in parentheses is trivial—i.e., the OR approximation exists—unless $\boldsymbol{d}_{m-1}^{\mathrm{MR}} \perp \mathscr{W}_m$, which means $c_m = 0$. □

In summary, the OR approximation $P_{\mathscr{W}_m}^{\mathscr{V}_m}\boldsymbol{r}$, with $\mathscr{V}_m$ defined as in (2.20), fails to exist if and only if $c_m = 0$, which, as the preceding section has shown, coincides with the situation in which the associated MR approximation makes no progress when extending $\mathscr{W}_{m-1}$ to $\mathscr{W}_m$. We therefore tacitly assume $(\boldsymbol{r}, \boldsymbol{w}_m) \neq 0$ whenever referring to the OR approximation of index m.

In many algorithms such as GMRES it is convenient to work also with orthonormal bases $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ of $\mathscr{V}_m$, $m = 1, 2, \ldots$, i.e., $\boldsymbol{v}_m$ is a unit vector in $\mathscr{V}_m \cap \mathscr{V}_{m-1}^{\perp}$. The following result relates $\measuredangle(\boldsymbol{r}, \boldsymbol{w}_m)$ with $\measuredangle(\boldsymbol{v}_{m+1}, \boldsymbol{w}_m)$.

Lemma 2.2.4. *If the subspaces $\mathscr{V}_m$ and $\mathscr{W}_m$ are related by* (2.20) *and $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ and $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ are orthonormal bases of $\mathscr{V}_m$ and $\mathscr{W}_m$, respectively, then there holds*

$$|(\boldsymbol{v}_{m+1}, \boldsymbol{w}_m)| = s_m.$$

In the degenerate case $(\boldsymbol{r}, \boldsymbol{w}_m) = 0$, the vectors $\boldsymbol{v}_{m+1}$ and $\boldsymbol{w}_m$ must be collinear.

Proof. From $\boldsymbol{v}_{m+1} \in \operatorname{span}\{\boldsymbol{r}, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\} \cap \operatorname{span}\{\boldsymbol{r}, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{m-1}\}^\perp$ we conclude $\boldsymbol{v}_{m+1} \in \operatorname{span}\{\boldsymbol{d}_m^{\mathrm{MR}}, \boldsymbol{w}_m\}$ and the assertion follows from the remaining requirements $\|\boldsymbol{v}_{m+1}\| = 1$ and $\boldsymbol{v}_{m+1} \perp \boldsymbol{r}$: Using the notation from the proof of Theorem 2.2.1, we set $\widehat{\boldsymbol{w}}_{m+1} := \boldsymbol{d}_m^{\mathrm{MR}}/\|\boldsymbol{d}_m^{\mathrm{MR}}\|$ and note that $\{\widehat{\boldsymbol{w}}_m, \boldsymbol{w}_m\}$ form an orthonormal basis of $\operatorname{span}\{\boldsymbol{d}_m^{\mathrm{MR}}, \boldsymbol{w}_m\}$, hence $\boldsymbol{v}_{m+1} = \alpha\widehat{\boldsymbol{w}}_m + \beta\boldsymbol{w}_m$ for some coefficients $\alpha, \beta \in \mathbb{C}$. Since $(\boldsymbol{d}_m^{\mathrm{MR}}, \boldsymbol{r}) = \|\boldsymbol{d}_m^{\mathrm{MR}}\|^2$, orthogonality of $\boldsymbol{v}_{m+1}$ and $\boldsymbol{r}$ yields

$$0 = (\boldsymbol{v}_{m+1}, \boldsymbol{r}) = \frac{\alpha}{\|\boldsymbol{d}_m^{\mathrm{MR}}\|}(\boldsymbol{d}_m^{\mathrm{MR}}, \boldsymbol{r}) + \beta(\boldsymbol{w}_m, \boldsymbol{r}) = \alpha\|\boldsymbol{d}_m^{\mathrm{MR}}\| + \beta(\boldsymbol{w}_m, \boldsymbol{r}).$$

The requirement that $\boldsymbol{v}_{m+1}$ have unit norm now gives

$$|\beta|^2 = \left(1 + \frac{|(\boldsymbol{w}_m, \boldsymbol{r})|^2}{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2}\right)^{-1} = \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2}{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2 + |(\boldsymbol{w}_m, \boldsymbol{r})|^2} = \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2} = s_m^2.$$

Since $\boldsymbol{d}_m^{\mathrm{MR}} \perp \boldsymbol{w}_m$, we now obtain $|(\boldsymbol{v}_{m+1}, \boldsymbol{w}_m)| = |\beta| = s_m$.

When $(\boldsymbol{r}, \boldsymbol{w}_m) = 0$, this implies $|(\boldsymbol{v}_{m+1}, \boldsymbol{w}_m)| = 1$, i.e., $|(\boldsymbol{v}_{m+1}, \boldsymbol{w}_m)| = \|\boldsymbol{v}_{m+1}\|\,\|\boldsymbol{w}_m\|$ which means that these two vectors are collinear. □

2.2.3 Relations Between Nested MR and OR Approximations

Recall that $\boldsymbol{r} \in \mathscr{W}_{m-1}$ implies that the MR approximation with respect to $\mathscr{W}_{m-1}$ solves the approximation problem exactly. The same is then true for the OR approximation, since $\boldsymbol{r} \in \mathscr{W}_{m-1}$ implies $\boldsymbol{r} - \boldsymbol{w}_{m-1}^{\mathrm{OR}} \in \mathscr{W}_{m-1} \cap \mathscr{V}_{m-1}^\perp = \{\boldsymbol{0}\}$, so that the OR approximation solves the problem exactly whenever the MR approximation does. In other words, the assumption $\boldsymbol{r} \notin \mathscr{W}_{m-1}$ is *equivalent* with saying that both $\boldsymbol{d}_{m-1}^{\mathrm{MR}}$ and $\boldsymbol{d}_{m-1}^{\mathrm{OR}}$ are not yet zero.

If we define $\widetilde{\boldsymbol{w}}_m := |(\boldsymbol{w}_m, \boldsymbol{r})|/(\boldsymbol{r}, \boldsymbol{w}_m)\widehat{\boldsymbol{w}}_m$, where $\widehat{\boldsymbol{w}}_m = \boldsymbol{d}_{m-1}^{\mathrm{MR}}/\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|$, then $(\boldsymbol{w}_m, \widetilde{\boldsymbol{w}}_m) = c_m$ (cf. the proof of Theorem 2.2.1). Consequently, the sets $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{m-1}, \widetilde{\boldsymbol{w}}_m/c_m\}$ and $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ form a pair of *biorthonormal* bases of $\mathscr{V}_m$ and $\mathscr{W}_m$. This fact allows us to express the oblique projection which determines the OR approximation as the singular value expansion

$$P_{\mathscr{W}_m}^{\mathscr{V}_m} = \sum_{j=1}^{m-1}(\cdot, \boldsymbol{w}_j)\boldsymbol{w}_j + \frac{1}{c_m}(\cdot, \widetilde{\boldsymbol{w}}_m)\boldsymbol{w}_m, \tag{2.22}$$

from which we derive the following expression for the difference of the OR and MR approximations:

$$\begin{aligned}\boldsymbol{w}_m^{\mathrm{OR}} - \boldsymbol{w}_m^{\mathrm{MR}} &= (P_{\mathscr{W}_m}^{\mathscr{V}_m} - P_{\mathscr{W}_m})\boldsymbol{r} = \left[c_m^{-1}(\boldsymbol{r}, \widetilde{\boldsymbol{w}}_m) - (\boldsymbol{r}, \boldsymbol{w}_m)\right]\boldsymbol{w}_m \\ &= \frac{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2 - |(\boldsymbol{r}, \boldsymbol{w}_m)|^2}{(\boldsymbol{w}_m, \boldsymbol{r})}\boldsymbol{w}_m = \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2}{(\boldsymbol{w}_m, \boldsymbol{r})}\boldsymbol{w}_m.\end{aligned} \tag{2.23}$$

In other words, since the spaces $\mathscr{V}_m$ and $\mathscr{W}_m$ are so closely related, the projection $P_{\mathscr{W}_m}^{\mathscr{V}_m}$ is simply a rank-one modification of $P_{\mathscr{W}_m}$, and this is the essential ingredient of the proof of the following familiar relations:

Theorem 2.2.5. *Given an arbitrary element $\boldsymbol{r} \in \mathscr{H}$, a nested sequence of subspaces $\mathscr{W}_m \subset \mathscr{H}$ of dimension m (cf. (2.11)) and a corresponding sequence of error spaces $\mathscr{V}_m$ as defined by (2.20), then the MR and OR approximations of $\boldsymbol{r}$ with respect to $\mathscr{W}_m$ and $\mathscr{V}_m$ satisfy*

$$\|\boldsymbol{d}_m^{MR}\| = s_m \|\boldsymbol{d}_{m-1}^{MR}\|, \tag{2.24}$$

$$\|\boldsymbol{d}_m^{MR}\| = s_1 s_2 \dots s_m \|\boldsymbol{r}\|, \tag{2.25}$$

$$\|\boldsymbol{d}_m^{MR}\| = c_m \|\boldsymbol{d}_m^{OR}\|, \tag{2.26}$$

$$\|\boldsymbol{d}_m^{OR}\| = s_1 s_2 \cdots s_m \|\boldsymbol{r}\| / c_m, \tag{2.27}$$

where $s_m = \sin \measuredangle(\boldsymbol{d}_{m-1}^{MR}, \mathscr{W}_m)$ and $c_m = \cos \measuredangle(\boldsymbol{d}_{m-1}^{MR}, \mathscr{W}_m)$.

Proof. Identities (2.24) and (2.25) are merely restatements of (2.16) and (2.19) which have already been proven. Next, from (2.23) there follows

$$\boldsymbol{d}_m^{\mathrm{MR}} - \boldsymbol{d}_m^{\mathrm{OR}} = \boldsymbol{w}_m^{\mathrm{OR}} - \boldsymbol{w}_m^{\mathrm{MR}} \in \operatorname{span}\{\boldsymbol{w}_m\}, \tag{2.28}$$

and

$$\boldsymbol{w}_m^{\mathrm{OR}} = \boldsymbol{w}_m^{\mathrm{MR}} + \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2}{(\boldsymbol{w}_m, \boldsymbol{r})} \boldsymbol{w}_m = \boldsymbol{w}_{m-1}^{\mathrm{MR}} + \frac{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2}{(\boldsymbol{w}_m, \boldsymbol{r})} \boldsymbol{w}_m, \tag{2.29}$$

where we have used (2.13) for the last equality. Since $\boldsymbol{d}_m^{\mathrm{MR}} \perp \boldsymbol{w}_m$, the Pythagorean identity and (2.13) yield

$$\|\boldsymbol{d}_m^{\mathrm{OR}}\|^2 = \left(1 + \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2}{|(\boldsymbol{w}_m, \boldsymbol{r})|^2}\right) \|\boldsymbol{d}_m^{\mathrm{MR}}\|^2 = \frac{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2}{|(\boldsymbol{w}_m, \boldsymbol{r})|^2} \|\boldsymbol{d}_m^{\mathrm{MR}}\|^2,$$

which, in view of (2.17), gives the error formula

$$\|\boldsymbol{d}_m^{\mathrm{OR}}\| = \frac{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\| \, \|\boldsymbol{d}_m^{\mathrm{MR}}\|}{|(\boldsymbol{w}_m, \boldsymbol{r})|} = \frac{1}{c_m} \|\boldsymbol{d}_m^{\mathrm{MR}}\|$$

for the OR approximation, establishing (2.26) and (2.27). □

In addition to the norm identities contained in Theorem 2.2.5, it is also possible to relate the MR and OR approximations and their errors, as the next theorem shows.

Theorem 2.2.6. *Under the assumptions of Theorem 2.2.5, the MR and OR approximations and errors satisfy*

$$\boldsymbol{w}_m^{MR} = s_m^2 \boldsymbol{w}_{m-1}^{MR} + c_m^2 \boldsymbol{w}_m^{OR}, \tag{2.30}$$

$$\boldsymbol{d}_m^{MR} = s_m^2 \boldsymbol{d}_{m-1}^{MR} + c_m^2 \boldsymbol{d}_m^{OR}, \tag{2.31}$$

$$\frac{1}{\|\boldsymbol{d}_m^{MR}\|^2} \boldsymbol{w}_m^{MR} = \sum_{j=0}^{m} \frac{1}{\|\boldsymbol{d}_j^{OR}\|^2} \boldsymbol{w}_j^{OR}, \tag{2.32}$$

$$\frac{1}{\|\boldsymbol{d}_m^{MR}\|^2} \boldsymbol{d}_m^{MR} = \sum_{j=0}^{m} \frac{1}{\|\boldsymbol{d}_j^{OR}\|^2} \boldsymbol{d}_j^{OR}, \tag{2.33}$$

$$\frac{1}{\|\boldsymbol{d}_m^{MR}\|^2} = \sum_{j=0}^{m} \frac{1}{\|\boldsymbol{d}_j^{OR}\|^2} = \frac{1}{\|\boldsymbol{d}_{m-1}^{MR}\|^2} + \frac{1}{\|\boldsymbol{d}_m^{OR}\|^2}, \tag{2.34}$$

Proof. From (2.29) and $\boldsymbol{h}_m^{\mathrm{MR}} - \boldsymbol{h}_{m-1}^{\mathrm{MR}} = (\boldsymbol{r}, \boldsymbol{w}_m)\boldsymbol{w}_m$ we obtain

$$\begin{aligned}
\boldsymbol{w}_m^{\mathrm{OR}} &= \boldsymbol{w}_m^{\mathrm{MR}} + \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2}{(\boldsymbol{w}_m, \boldsymbol{r})} \frac{1}{(\boldsymbol{r}, \boldsymbol{w}_m)} (\boldsymbol{w}_m^{\mathrm{MR}} - \boldsymbol{w}_{m-1}^{\mathrm{MR}}) \\
&= \boldsymbol{w}_m^{\mathrm{MR}} + \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2} \frac{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2}{|(\boldsymbol{r}, \boldsymbol{w}_m)|^2} (\boldsymbol{w}_m^{\mathrm{MR}} - \boldsymbol{w}_{m-1}^{\mathrm{MR}}) \\
&= \boldsymbol{w}_m^{\mathrm{MR}} + \frac{s_m^2}{c_m^2} (\boldsymbol{w}_m^{\mathrm{MR}} - \boldsymbol{w}_{m-1}^{\mathrm{MR}})
\end{aligned}$$

(cf. (2.24) and (2.17)), which implies the relationship (2.30) between the OR and MR approximations and, by way of $s_m^2 + c_m^2 = 1$, the corresponding relationship (2.31) between their errors.

Repeated application of these two formulas leads to

$$\boldsymbol{w}_m^{\mathrm{MR}} = \sum_{j=0}^{m} \tau_{m,j}^2 \boldsymbol{w}_j^{\mathrm{OR}} \qquad \text{and} \qquad \boldsymbol{d}_m^{\mathrm{MR}} = \sum_{j=0}^{m} \tau_{m,j}^2 \boldsymbol{d}_j^{\mathrm{OR}},$$

where $\tau_{m,0} := s_1 s_2 \dots s_m$ and $\tau_{m,j} := c_j s_{j+1} \dots s_m$ $(1 \le j \le m)$. Using (2.24) and (2.26) this can be simplified to

$$\tau_{m,j} = c_j \frac{\|\boldsymbol{d}_{j+1}^{\mathrm{MR}}\|}{\|\boldsymbol{d}_j^{\mathrm{MR}}\|} \frac{\|\boldsymbol{d}_{j+2}^{\mathrm{MR}}\|}{\|\boldsymbol{d}_{j+1}^{\mathrm{MR}}\|} \cdots \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|} = c_j \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|}{\|\boldsymbol{d}_j^{\mathrm{MR}}\|} = \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|}{\|\boldsymbol{d}_j^{\mathrm{OR}}\|}$$

and we obtain (2.32) as well as (2.33). Finally, since the errors $\boldsymbol{d}_j^{\mathrm{OR}}$ are orthogonal, we have

$$\frac{1}{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2} = \sum_{j=0}^{m} \frac{1}{\|\boldsymbol{d}_j^{\mathrm{OR}}\|^2} = \sum_{j=0}^{m-1} \frac{1}{\|\boldsymbol{d}_j^{\mathrm{OR}}\|^2} + \frac{1}{\|\boldsymbol{d}_m^{\mathrm{OR}}\|^2} = \frac{1}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2} + \frac{1}{\|\boldsymbol{d}_m^{\mathrm{OR}}\|^2},$$

which proves (2.34). Strictly speaking, this proves

$$\frac{1}{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2} = \frac{1}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2} + \frac{1}{\|\boldsymbol{d}_m^{\mathrm{OR}}\|^2}$$

only under the assumption that all OR approximations $\boldsymbol{h}_1^{\mathrm{OR}}, \ldots, \boldsymbol{h}_m^{\mathrm{OR}}$ exist. But this last equation is merely a reformulation of the Pythagorean identity,

$$1 = s_m^2 + c_m^2 = \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2} + \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2}{\|\boldsymbol{d}_m^{\mathrm{OR}}\|^2}$$

(cf. (2.24) and (2.26)), requiring only the existence of $\boldsymbol{w}_m^{\mathrm{OR}}$ (besides $\boldsymbol{r} \notin \mathscr{W}_m$). □

Corollary 2.2.7. *In view of*

$$s_m = \frac{\|\boldsymbol{d}_m^{MR}\|}{\|\boldsymbol{d}_{m-1}^{MR}\|}, \qquad \textit{i.e.,} \quad c_m = \sqrt{1 - \frac{\|\boldsymbol{d}_m^{MR}\|^2}{\|\boldsymbol{d}_{m-1}^{MR}\|^2}},$$

an angle-free formulation of (2.26), (2.30) *and* (2.31) *reads*

$$\begin{aligned}
\|\boldsymbol{d}_m^{MR}\| &= \sqrt{1 - \frac{\|\boldsymbol{d}_m^{MR}\|^2}{\|\boldsymbol{d}_{m-1}^{MR}\|^2}}\, \|\boldsymbol{d}_m^{OR}\|, \\
\boldsymbol{w}_m^{MR} &= \boldsymbol{w}_m^{OR} + \frac{\|\boldsymbol{d}_m^{MR}\|^2}{\|\boldsymbol{d}_{m-1}^{MR}\|^2}(\boldsymbol{w}_{m-1}^{MR} - \boldsymbol{w}_m^{OR}), \\
\boldsymbol{d}_m^{MR} &= \boldsymbol{d}_m^{OR} + \frac{\|\boldsymbol{d}_m^{MR}\|^2}{\|\boldsymbol{d}_{m-1}^{MR}\|^2}(\boldsymbol{d}_{m-1}^{MR} - \boldsymbol{d}_m^{OR}).
\end{aligned}$$

Of course, the first of these identities, or its reformulation

$$\|\boldsymbol{d}_m^{\mathrm{OR}}\| = \left(1 - \frac{\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2}\right)^{-1/2} \|\boldsymbol{d}_m^{\mathrm{MR}}\|,$$

only makes sense if $\boldsymbol{w}_m^{\mathrm{OR}}$ is defined, which is equivalent to $\|\boldsymbol{d}_m^{\mathrm{MR}}\| < \|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|$. Whenever $\|\boldsymbol{d}_m^{\mathrm{MR}}\| \approx \|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|$ then the factor $(1 - \|\boldsymbol{d}_m^{\mathrm{MR}}\|^2/\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2)^{-1/2}$ will be large and, consequently, $\|\boldsymbol{d}_m^{\mathrm{OR}}\| \gg \|\boldsymbol{d}_m^{\mathrm{MR}}\|$. Conversely, if the MR approximation makes considerable progress in step m then $(1 - \|\boldsymbol{d}_m^{\mathrm{MR}}\|^2/\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2)^{-1/2} \approx 1$ and $\|\boldsymbol{d}_m^{\mathrm{OR}}\| \approx \|\boldsymbol{d}_m^{\mathrm{MR}}\|$. In the context of Krylov subspace methods, this observation is sometimes referred to as the *peak/plateau-phenomenon* of OR/MR approximations (see e.g. Cullum & Greenbaum (1996)).

Remark 2.2.8. We conclude this section by reconsidering the issue of the Galerkin breakdown (see Remark 2.1.5) for the case of approximation from nested subspaces. Using the biorthonormal bases introduced in (2.22), we obtain the singular value expansion of $P_{\mathscr{V}_m} P_{\mathscr{W}_m}$ as

$$P_{\mathscr{V}_m} P_{\mathscr{W}_m} = \sum_{j=1}^{m-1} (\cdot, \boldsymbol{w}_j)\boldsymbol{w}_j + c_m(\cdot, \boldsymbol{w}_m)\widetilde{\boldsymbol{w}}_m. \tag{2.35}$$

In Remark 2.1.5 we noted that the natural generalization of the OR approximation in the case of a Galerkin breakdown ($c_m = 0$) is given by $\boldsymbol{w}_m^{\mathrm{OR}} := (P_{\mathscr{V}_m} P_{\mathscr{W}_m})^+ \boldsymbol{r}$. With the singular value expansion (2.35) this leads to simply $\boldsymbol{w}_m^{\mathrm{OR}} = \boldsymbol{w}_{m-1}^{\mathrm{MR}} = \boldsymbol{w}_m^{\mathrm{MR}}$.

2.2.4 Smoothing Algorithms

A *smoothing algorithm* transforms a given sequence $\{\boldsymbol{h}_m\} \subset \mathscr{W}_m$ of approximations to $\boldsymbol{r}$ into a new sequence $\{\widehat{\boldsymbol{h}}_m\} \subset \mathscr{W}_m$, according to

$$\widehat{\boldsymbol{h}}_m := (1-\alpha_m)\widehat{\boldsymbol{h}}_{m-1} + \alpha_m \boldsymbol{h}_m, \quad \text{i.e.,} \quad \widehat{\boldsymbol{d}}_m = (1-\alpha_m)\widehat{\boldsymbol{d}}_{m-1} + \alpha_m \boldsymbol{d}_m, \tag{2.36}$$

($m = 1, 2, \ldots$, $\widehat{\boldsymbol{h}}_0 := \boldsymbol{h}_0 = \boldsymbol{0}$). The intention is that the approximation errors of the transformed sequence should decrease "more smoothly" than those associated with the original sequence. Ideally, we would like to have $\widehat{\boldsymbol{h}}_m = \boldsymbol{w}_m^{\mathrm{MR}}$, and we shall discuss two smoothing procedures which achieve this goal when applied to the sequence of OR approximations.

In *minimal residual smoothing* (Weiss (1994), Zhou & Walker (1994), Gutknecht (1997, Section 17)) the parameter α_m in (2.36) is chosen to minimize the norm of the error $\widehat{\boldsymbol{d}}_m = \widehat{\boldsymbol{d}}_{m-1} - \alpha_m(\widehat{\boldsymbol{d}}_{m-1} - \boldsymbol{d}_m)$ as a function of α_m. In other words, this determines the best approximation $\alpha_m(\widehat{\boldsymbol{d}}_{m-1} - \boldsymbol{d}_m)$ to $\widehat{\boldsymbol{d}}_{m-1}$ from $\operatorname{span}\{\widehat{\boldsymbol{d}}_{m-1} - \boldsymbol{d}_m\}$, which is obtained for

$$\alpha_m^{\mathrm{MR}} := \frac{(\widehat{\boldsymbol{d}}_{m-1}, \widehat{\boldsymbol{d}}_{m-1} - \boldsymbol{d}_m)}{\|\widehat{\boldsymbol{d}}_{m-1} - \boldsymbol{d}_m\|^2}.$$

In an alternative smoothing procedure known as *quasi-minimal residual smoothing* (Zhou & Walker (1994), Gutknecht (1997, Section 17)) the parameter α_m is chosen as

$$\alpha_m^{\mathrm{QMR}} := \frac{\tau_m^2}{\|\boldsymbol{d}_m\|^2} \quad \text{with } \tau_m \text{ such that} \quad \frac{1}{\tau_m^2} = \frac{1}{\tau_{m-1}^2} + \frac{1}{\|\boldsymbol{d}_m\|^2},\ \tau_0 = \|\boldsymbol{r}\|.$$

It is easy to see by induction that

$$\tau_m^2 = \frac{1}{\sum_{j=0}^m 1/\|\boldsymbol{d}_j\|^2}, \quad \text{i.e.,} \quad \alpha_m^{\mathrm{QMR}} = \frac{1/\|\boldsymbol{d}_m\|^2}{\sum_{j=0}^m 1/\|\boldsymbol{d}_j\|^2},$$

and therefore

$$\widehat{\boldsymbol{h}}_m = \frac{\sum_{j=0}^m \boldsymbol{h}_j/\|\boldsymbol{d}_j\|^2}{\sum_{j=0}^m 1/\|\boldsymbol{d}_j\|^2} \quad \text{as well as} \quad \widehat{\boldsymbol{d}}_m = \frac{\sum_{j=0}^m \boldsymbol{d}_j/\|\boldsymbol{d}_j\|^2}{\sum_{j=0}^m 1/\|\boldsymbol{d}_j\|^2}.$$

The last formula for $\widehat{\boldsymbol{h}}_m$ reveals the strategy behind quasi-minimal residual smoothing: $\widehat{\boldsymbol{h}}_m$ is a weighted sum of all previous approximations $\boldsymbol{h}_0, \boldsymbol{h}_1, \ldots, \boldsymbol{h}_m$ with weights $(1/\|\boldsymbol{d}_k\|^2)/(\sum_{j=0}^m (1/\|\boldsymbol{d}_j\|^2))$ which are (relatively) large if $\boldsymbol{h}_k$ approximates $\boldsymbol{r}$ well, and (relatively) small whenever $\boldsymbol{h}_k$ is a poor approximation of $\boldsymbol{r}$.

In general, i.e., for an arbitrary sequence $\{\boldsymbol{h}_m\} \subset \mathscr{W}_m$, minimal and quasi-minimal residual smoothing will generate different "smoothed" approximations $\widehat{\boldsymbol{h}}_m$. In the case of $\boldsymbol{h}_m = \boldsymbol{w}_m^{\mathrm{OR}}$, however, the two schemes are equivalent:

Proposition 2.2.9. *If either the minimal residual or the quasi-minimal residual smoothing algorithm is applied to the sequence of OR approximations $\{\boldsymbol{w}_m^{OR}\}$ for an element $\boldsymbol{r} \in \mathscr{H}$ with respect to the sequence of approximation spaces $\{\mathscr{W}_m\}$ and error spaces $\{\mathscr{V}_m\}$, then the resulting smoothed sequence consists of the MR approximations $\boldsymbol{w}_m^{MR}$ for $\boldsymbol{r}$ and the associated smoothing parameters are $\alpha_m^{MR} = \alpha_m^{QMR} = c_m^2$.*

Proof. An induction shows that minimal residual smoothing applied to $\boldsymbol{h}_m = \boldsymbol{w}_m^{\mathrm{OR}}$ yields $\widehat{\boldsymbol{h}}_m = \boldsymbol{w}_m^{\mathrm{MR}}$ and that, in this case, $\alpha_m^{\mathrm{MR}} = c_m^2$: The assertion is trivial for $m = 1$. Assuming $\widehat{\boldsymbol{h}}_{m-1} = \boldsymbol{w}_{m-1}^{\mathrm{MR}}$ for some $m \geq 2$, we see from (2.29)

$$\widehat{\boldsymbol{d}}_{m-1} - \boldsymbol{d}_m = \boldsymbol{d}_{m-1}^{\mathrm{MR}} - \boldsymbol{d}_m^{\mathrm{OR}} = \frac{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2}{(\boldsymbol{w}_m, \boldsymbol{r})} \boldsymbol{w}_m$$

and consequently, noting that $\boldsymbol{d}_{m-1}^{\mathrm{MR}} = \boldsymbol{d}_m^{\mathrm{MR}} + (\boldsymbol{r}, \boldsymbol{w}_m)\boldsymbol{w}_m$ and $\boldsymbol{d}_m^{\mathrm{MR}} \perp \boldsymbol{w}_m$,

$$(\widehat{\boldsymbol{d}}_{m-1}, \widehat{\boldsymbol{d}}_{m-1} - \boldsymbol{d}_m) = \left((\boldsymbol{r}, \boldsymbol{w}_m)\boldsymbol{w}_m, \frac{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2}{(\boldsymbol{w}_m, \boldsymbol{r})} \boldsymbol{w}_m \right) = \|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2.$$

From (2.17) and (2.30) it then follows that

$$\alpha_m^{\mathrm{MR}} = \|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2 \frac{|(\boldsymbol{w}_m, \boldsymbol{r})|^2}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^4} = \frac{|(\boldsymbol{w}_m, \boldsymbol{r})|^2}{\|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2} = c_m^2 \quad \text{and} \quad \widehat{\boldsymbol{w}}_m = \boldsymbol{w}_m^{\mathrm{MR}}.$$

The analogous assertion for quasi-minimal residual smoothing follows, with (2.32) and (2.33), immediately from the orthogonality of the error vectors $\boldsymbol{d}_m^{\mathrm{OR}}$. □

It should not come as a surprise that, in our setting, a one-dimensional minimization procedure like minimal residual smoothing produces the best approximation $\boldsymbol{w}_m^{\mathrm{MR}}$, which is the global optimum on $\mathscr{W}_m$. Recall that $\boldsymbol{w}_m^{\mathrm{OR}}$ is already "nearly optimal" and needs to be corrected only in the direction of $\boldsymbol{w}_m$.

Notes and Remarks

The relations given in Theorem 2.2.5, Theorem 2.2.6 and Corollary 2.2.7 have been derived, rederived and generalized many times in the literature on Krylov subspace methods for solving linear systems of equations, see for example Paige & Saunders (1975), Weiss (1990), Brown (1991), Freund & Nachtigal (1991), Freund (1993), Gutknecht (1993*a*), and Cullum & Greenbaum (1996). We emphasize here that these results are a direct consequence of the orthogonal and oblique projection methods with the test space defined by (2.20), and therefore hold for any approximation method based on these projections. In particular, they are neither restricted to Krylov subspace methods, nor even to solving linear operator equations.

The characterization of the parameters c_m and s_m occurring in these results as angles between the spaces $\mathscr{V}_m$ and $\mathscr{W}_m$ as well as the expression of the oblique projection as the singular value expansion (2.22) are due to Eiermann & Ernst (1998).

The smoothing transformations of Section 2.2.4 were introduced by Schönauer (1987) to obtain a smoothly converging sequence of approximations and residuals associated with a linear system (1.1) when certain Lanczos-based iterative solvers are used, which are known to often produce approximations with wildly oscillating curves of residual norms. As the approximations are linear functions of the residuals, these are obtained as the linear combinations of previous approximations using the same coefficients as for the residuals. It was shown by Weiss (1994) that the iterates/residuals of the QMR iterative method of Freund & Nachtigal (1991) can be obtained by applying the MR smoothing

procedure to those of the biconjugate gradient method. QMR smoothing was introduced by Walker (1995). See also Gutknecht (1993*a*) and Gutknecht (1997, Section 17). Again, these two smoothing transforms work in the very general setting of Chapter 2, and the fact that the orthogonal projection associated with the MR approximation is a rank-one modification of the oblique projection associated with the OR approximation is the key to the simple proof of Proposition 2.2.9.

2.3 Working with Coordinates

The implementation of the MR and OR approximations on nested subspaces described in the previous sections only become applicable for computations once suitable bases are chosen for the subspaces involved. For the MR approximation the most obvious choice is to use an ascending basis of the approximation spaces $\mathscr{W}_m$ and, in view of (2.22), this basis is also the most practical for OR approximation, and this is briefly mentioned in Section 2.3.1. When these methods are used for the approximate solution of operator equations, however, many important methods—among these the class of Krylov subspace methods—become more efficient and stable when formulated in terms of a (nested) basis of the error spaces $\mathscr{V}_m$, and this is treated in Section 2.3.2 Another distinction is whether or not the bases are kept orthogonal. There are important cases where it is worthwhile to give up the stability that comes with orthogonal bases for the economy afforded by certain formulations which use non-orthogonal bases, and these are the subject of Section 2.3.3.

We shall see that the resulting coordinate formulations of Sections 2.3.2 and 2.3.3 are identical to those of familiar Krylov subspace methods, which shows that, not only are the latter specializations of these general approximation schemes to Krylov spaces, but that even the coordinate calculations are the same for the general case.

2.3.1 Using an Orthonormal Basis of $\mathscr{W}_m$

Let the vectors $\{\boldsymbol{w}_m\}_{m\geq 1}$ form a nested sequence of orthonormal bases for the sequence of approximation spaces $\{\mathscr{W}_m\}_{m\geq 0}$, i.e., let the vectors $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ form an orthonormal basis of $\mathscr{W}_m$ for each $m \geq 1$. In this case each $\boldsymbol{w} \in \mathscr{W}_m$ possesses the unique representation $\boldsymbol{w} = W_m \boldsymbol{y}$, in which W_m denotes the row vector $W_m := [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m]$ and $\boldsymbol{y} \in \mathbb{C}^m$ is the coordinate vector of $\boldsymbol{w}$ with respect to this basis. The characterization $\boldsymbol{d}_m^{\mathrm{MR}} = \boldsymbol{r} - W_m \boldsymbol{y}_m^{\mathrm{MR}} \perp \mathscr{W}_m$ then immediately determines the coordinate vector $\boldsymbol{y}_m^{\mathrm{MR}}$ of $\boldsymbol{w}_m^{\mathrm{MR}}$ as

$$\boldsymbol{y}_m^{\mathrm{MR}} = [(\boldsymbol{r}, \boldsymbol{w}_1), \ldots, (\boldsymbol{r}, \boldsymbol{w}_m)]^\top. \tag{2.37}$$

The coordinate vector $\boldsymbol{y}_m^{\mathrm{OR}}$ of the corresponding OR approximation $\boldsymbol{w}_m^{\mathrm{OR}}$ is given by

$$\boldsymbol{y}_m^{\mathrm{OR}} = [(\boldsymbol{r}, \boldsymbol{w}_1), \ldots, (\boldsymbol{r}, \boldsymbol{w}_{m-1}), \|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2/(\boldsymbol{w}_m, \boldsymbol{r})]^\top \tag{2.38}$$

since $\boldsymbol{w}_m^{\mathrm{OR}} = \boldsymbol{w}_{m-1}^{\mathrm{MR}} + \|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|^2/(\boldsymbol{w}_m, \boldsymbol{r})\boldsymbol{w}_m$ (cf. (2.29)).

2.3.2 Using an Orthonormal Basis of $\mathscr{V}_m$

We now drop the orthogonality requirement on the vectors $\boldsymbol{w}_j$, assuming only that, for each m, $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ forms a basis of $\mathscr{W}_m$. In the same manner, let $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m+1}\}$ form

a basis of the corresponding error space $\mathscr{V}_{m+1}$. Since $\mathscr{W}_m \subseteq \operatorname{span}\{\boldsymbol{r}\} + \mathscr{W}_m = \mathscr{V}_{m+1}$, we may represent each $\boldsymbol{w}_k$ as a linear combination of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k+1}$,

$$\boldsymbol{w}_k = \sum_{j=1}^{k+1} \eta_{j,k} \boldsymbol{v}_j, \qquad k = 1, \ldots, m,$$

or, employing the more compact row vector notation $W_m = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m]$ and $V_m := [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m]$, $m = 1, 2, \ldots$,

$$W_m = V_{m+1} \widetilde{H}_m = V_m H_m + [\boldsymbol{0}, \ldots, \boldsymbol{0}, \eta_{m+1,m} \boldsymbol{v}_{m+1}], \tag{2.39}$$

where $\widetilde{H}_m =: [\eta_{j,k}] \in \mathbb{C}^{(m+1)\times m}$ is an upper Hessenberg matrix and $H_m := [I_m \; \boldsymbol{0}]\widetilde{H}_m$ is the square matrix obtained by omitting the last row of $\widetilde{H}_m$. Note that, as long as $\boldsymbol{r} \notin \mathscr{W}_m$, i.e., $\boldsymbol{w}_m \notin \mathscr{V}_m$, we have $\eta_{m+1,m} \neq 0$, which implies $\operatorname{rank}(\widetilde{H}_m) = m$. If $\boldsymbol{r} \in \mathscr{W}_m$ for some index m, we set

$$L := \min\{m : \boldsymbol{r} \in \mathscr{W}_m\} \tag{2.40}$$

to be the smallest such index and observe that $\mathscr{V}_{L+1} = \mathscr{V}_L = \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_L\}$, i.e., $W_L = V_L H_L$, implying $\operatorname{rank}(H_L) = \operatorname{rank}(\widetilde{H}_L) = L$. If such an index does not exist, we set $L = \infty$. To avoid cumbersome notation we shall restrict our attention to the case $L < \infty$, which is the most relevant for practical applications. We note, however, that all our conclusions equally apply in the general case.

In the equation-solving algorithms to be discussed in the remaining chapters, the Hessenberg matrix introduced in (2.39) becomes the representation of the operator A in (1.1) on subspaces of $\mathscr{H}$, usually of low dimension. As we see here, this matrix also occurs naturally in this abstract setting as the link between the approximation space $\mathscr{W}_m$ and the error space $\mathscr{V}_{m+1}$.

The Orthogonalization Process

For a given sequence $\{\boldsymbol{w}_j\}_{j\geq 1}$, an *orthonormal* sequence of vectors $\{\boldsymbol{v}_j\}_{j\geq 1}$ may be constructed recursively starting with $\boldsymbol{v}_1 := \boldsymbol{r}/\|\boldsymbol{r}\|$ and, in view of $\mathscr{V}_{m+1} = \mathscr{V}_m + \operatorname{span}\{\boldsymbol{w}_m\}$, successively orthogonalizing each $\boldsymbol{w}_m$ against the previously generated $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$:

$$\begin{aligned} \boldsymbol{v}_1 &:= \boldsymbol{r}/\beta, \quad \beta := \|\boldsymbol{r}\|, \\ \boldsymbol{v}_{m+1} &:= \frac{(I - P_{\mathscr{V}_m})\boldsymbol{w}_m}{\|(I - P_{\mathscr{V}_m})\boldsymbol{w}_m\|}, \qquad m = 1, 2, \ldots, L-1. \end{aligned} \tag{2.41}$$

Of course, this is nothing but the Gram-Schmidt orthogonalization procedure applied to the basis $\{\boldsymbol{r}, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{m-1}\}$ of $\mathscr{V}_m$, and hence in this case the entries in the resulting Hessenberg matrix (2.39) are given by

$$\eta_{j,m} = (\boldsymbol{w}_m, \boldsymbol{v}_j), \qquad j = 1, \ldots, m+1, \quad m \geq 1. \tag{2.42}$$

We also note that

$$\eta_{m+1,m} = (\boldsymbol{w}_m, \boldsymbol{v}_{m+1}) = \|(I - P_{\mathscr{V}_m})\boldsymbol{w}_m\| \geq 0 \tag{2.43}$$

with equality holding if and only if $\boldsymbol{w}_m \in \mathscr{V}_m$ or, equivalently, $\boldsymbol{r} \in \mathscr{W}_m$, i.e., $m = L$.

We now turn to the determination of the coordinate vectors of the MR and OR approximations with respect to the basis $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$. In the sequel we will use $\boldsymbol{u}_j$ to denote the j-th unit coordinate vector in $\mathbb{C}^m$, and, whenever this adds clarity, write $\boldsymbol{u}_j^{(m)}$ to specify the dimension of the underlying space.

Lemma 2.3.1. *The coordinate vector $\boldsymbol{y}_m^{MR} \in \mathbb{C}^m$ of the MR approximation $\boldsymbol{w}_m^{MR}$ with respect to the basis $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ is the solution of the least-squares problem*

$$\|\beta \boldsymbol{u}_1^{(m+1)} - \widetilde{H}_m \boldsymbol{y}\|_2 \to \min_{\boldsymbol{y} \in \mathbb{C}^m}, \tag{2.44}$$

whereas the coordinate vector $\boldsymbol{y}_m^{OR}$ of the OR approximation solves the linear system of equations

$$H_m \boldsymbol{y} = \beta \boldsymbol{u}_1^{(m)}. \tag{2.45}$$

In short,

$$\boldsymbol{y}_m^{MR} = \beta \widetilde{H}_m^{+} \boldsymbol{u}_1^{(m+1)} \quad \text{and} \quad \boldsymbol{y}_m^{OR} = \beta H_m^{-1} \boldsymbol{u}_1^{(m)},$$

where $\widetilde{H}_m^{+} = (\widetilde{H}_m^H \widetilde{H}_m)^{-1} \widetilde{H}_m^H$ is the Moore-Penrose pseudoinverse of $\widetilde{H}_m$.

Proof. The assertions of the lemma become obvious once the relevant quantities are represented in terms of the orthonormal basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m+1}\}$ of $\mathscr{V}_{m+1}$. The vector $\boldsymbol{r}$ to be approximated possesses the coordinate vector $\beta \boldsymbol{u}_1^{(m+1)}$ and the approximation space $\mathscr{W}_m = \operatorname{span}\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ is represented by the span of the columns of $\widetilde{H}_m$. In other words, if $\boldsymbol{w} \in \mathscr{W}_m$ has the coordinate vector $\boldsymbol{y}$ with respect to $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$, then $\boldsymbol{d} = \boldsymbol{r} - \boldsymbol{w} \in \mathscr{V}_{m+1}$ has the coordinate vector $\beta \boldsymbol{u}_1^{(m+1)} - \widetilde{H}_m \boldsymbol{y}$ with respect to $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m+1}\}$. More formally, for any $\boldsymbol{w} = W_m \boldsymbol{y} \in \mathscr{W}_m$ ($\boldsymbol{y} \in \mathbb{C}^m$), there holds

$$\boldsymbol{r} - \boldsymbol{w} = \boldsymbol{r} - W_m \boldsymbol{y} = \beta \boldsymbol{v}_1 - V_{m+1} \widetilde{H}_m \boldsymbol{y} = V_{m+1}(\beta \boldsymbol{u}_1^{(m+1)} - \widetilde{H}_m \boldsymbol{y}).$$

As the vectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m+1}\}$ are orthonormal, it follows that

$$\|\boldsymbol{r} - \boldsymbol{w}\| = \|\beta \boldsymbol{u}_1^{(m+1)} - \widetilde{H}_m \boldsymbol{y}\|_2$$

($\|\cdot\|_2$ denoting the Euclidean norm in $\mathbb{C}^{m+1}$). Similarly, $\boldsymbol{r} - \boldsymbol{w} \perp \mathscr{V}_m$ if and only if the first m coefficients of $\beta \boldsymbol{u}_1^{(m+1)} - \widetilde{H}_m \boldsymbol{y}$ vanish, i.e., if $\beta \boldsymbol{u}_1^{(m)} - H_m \boldsymbol{y} = \boldsymbol{0}$. □

Remark 2.3.2. To determine $\boldsymbol{y}_m^{\mathrm{OR}}$ using Lemma 2.3.1 we must of course assume that the linear system $H_m \boldsymbol{y} = \beta \boldsymbol{u}_1$ is solvable. But this is equivalent to our previous characterization of the existence of $\boldsymbol{w}_m^{\mathrm{OR}}$, namely with $c_m = \cos \measuredangle(\boldsymbol{d}_{m-1}^{\mathrm{MR}}, \mathscr{W}_m) \neq 0$ (cf. (2.18) and Corollary 2.2.2), which can be seen as follows: First, note that $\boldsymbol{u}_1^{(m)}$ together with the first $m-1$ column vectors of H_m form a basis of $\mathbb{C}^m$ as long as $m \leq L$ (since $\eta_{j+1,j} \neq 0$ for $j = 1, 2, \ldots, L-1$). This implies that $H_m \boldsymbol{y} = \beta \boldsymbol{u}_1$ is consistent, i.e., $\boldsymbol{u}_1 \in \operatorname{range}(H_m)$, if and only if H_m is nonsingular.

Next, recall from Remark 2.1.6 that c_m equals the smallest singular value of the matrix $[(\boldsymbol{v}_j, \widehat{\boldsymbol{w}}_k)]_{j,k=1,2,\ldots,m}$, where $\{\widehat{\boldsymbol{w}}_1, \widehat{\boldsymbol{w}}_2, \ldots, \widehat{\boldsymbol{w}}_m\}$ is any orthonormal basis of $\mathscr{W}_m$. We select

such an orthonormal basis and represent its elements as linear combinations in the original basis $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$. In our row vector notation, this leads to a nonsingular matrix $T \in \mathbb{C}^{m\times m}$ with $[\widehat{\boldsymbol{w}}_1, \widehat{\boldsymbol{w}}_2, \ldots, \widehat{\boldsymbol{w}}_m] = [\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m]T$. Now,

$$[(\boldsymbol{v}_j, \widehat{\boldsymbol{w}}_k)] = [(\boldsymbol{v}_j, \boldsymbol{w}_k)]\, T = H_m^H T$$

and, consequently, the smallest singular value of $[(\boldsymbol{v}_j, \widehat{\boldsymbol{w}}_k)]$ is positive if and only if H_m is nonsingular.

Remark 2.3.3. In view of (2.39) and the result of Lemma 2.3.1, the approximations $\boldsymbol{w}_m^{\mathrm{MR}}$ and $\boldsymbol{w}_m^{\mathrm{OR}}$ and their associated errors possess the following representations in terms of the basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m+1}\}$:

$$\begin{aligned} \boldsymbol{w}_m^{\mathrm{MR}} &= V_{m+1}\widetilde{H}_m\widetilde{H}_m^{+}\beta\boldsymbol{u}_1, \qquad & \boldsymbol{d}_m^{\mathrm{MR}} &= V_{m+1}\left(I_{m+1} - \widetilde{H}_m\widetilde{H}_m^{+}\right)\beta\boldsymbol{u}_1, \\ \boldsymbol{w}_m^{\mathrm{OR}} &= V_{m+1}\widetilde{H}_m H_m^{-1}\beta\boldsymbol{u}_1, \qquad & \boldsymbol{d}_m^{\mathrm{OR}} &= V_{m+1}\left(\begin{bmatrix} I_m \\ \mathbf{0} \end{bmatrix} - \widetilde{H}_m H_m^{-1}\right)\beta\boldsymbol{u}_1. \end{aligned} \tag{2.46}$$

The last identity shows that the coordinate vector of $\boldsymbol{d}_m^{\mathrm{OR}}$ has a particularly simple form: Introducing the notation $H_m^{-1} = \left[\eta_{j,k}^{[-1]}\right]$, we obtain for $m < L$,

$$\begin{aligned} \boldsymbol{d}_m^{\mathrm{OR}} &= \beta V_{m+1}\left(\boldsymbol{u}_1^{(m+1)} - \widetilde{H}_m H_m^{-1}\boldsymbol{u}_1^{(m)}\right) \\ &= \beta V_{m+1}\left(I_{m+1} - \widetilde{H}_m[H_m^{-1}\ \mathbf{0}\,]\right)\boldsymbol{u}_1^{(m+1)} \\ &= \beta V_{m+1}\begin{bmatrix} \mathbf{0} \\ -\eta_{m+1,m}\,\eta_{m,1}^{[-1]} \end{bmatrix} = -\beta\eta_{m+1,m}\,\eta_{m,1}^{[-1]}\boldsymbol{v}_{m+1}. \end{aligned}$$

The matrix $I_{m+1} - \widetilde{H}_m[H_m^{-1}\ \mathbf{0}\,]$ represents $I - P_{\mathscr{W}_m}^{\mathscr{V}_m}$ restricted to $\mathscr{V}_{m+1}$ with respect to the orthonormal basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m+1}\}$. The following lemma, which was recently obtained by Hochbruck and Lubich (Hochbruck & Lubich 1998), provides a simpler expression for this projection.

Lemma 2.3.4. *If $\widehat{\boldsymbol{w}}_{m+1} \in \mathscr{V}_{m+1} \cap \mathscr{W}_m^{\perp}$ is defined by the condition $(\boldsymbol{v}_{m+1}, \widehat{\boldsymbol{w}}_{m+1}) = 1$, then for all $\boldsymbol{v} \in \mathscr{V}_{m+1}$ there holds*

$$(I - P_{\mathscr{W}_m}^{\mathscr{V}_m})\boldsymbol{v} = (\boldsymbol{v}, \widehat{\boldsymbol{w}}_{m+1})\boldsymbol{v}_{m+1}. \tag{2.47}$$

Moreover, the coordinate vector $\widehat{\boldsymbol{y}}_{m+1}$ of $\widehat{\boldsymbol{w}}_{m+1}$ with respect to $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m+1}\}$ has the form

$$\widehat{\boldsymbol{y}}_{m+1} = \begin{bmatrix} \boldsymbol{g}_m \\ 1 \end{bmatrix}, \qquad \text{where } \boldsymbol{g}_m \text{ solves} \qquad H_m^H\boldsymbol{g}_m = -\eta_{m+1,m}\boldsymbol{u}_m.$$

Proof. In $\mathscr{V}_{m+1}$, the projection $I - P_{\mathscr{W}_m}^{\mathscr{V}_m}$ is characterized by the two properties

$$\begin{aligned} (I - P_{\mathscr{W}_m}^{\mathscr{V}_m})\boldsymbol{v} &= \boldsymbol{v} && \forall \boldsymbol{v} \in \mathscr{V}_{m+1} \cap \mathscr{V}_m^{\perp} = \operatorname{span}\{\boldsymbol{v}_{m+1}\}, \\ (I - P_{\mathscr{W}_m}^{\mathscr{V}_m})\boldsymbol{v} &= \mathbf{0} && \forall \boldsymbol{v} \in \mathscr{V}_{m+1} \cap \mathscr{W}_m, \end{aligned}$$

i.e., it is the oblique projection onto $\mathscr{V}_{m+1} \cap \mathscr{V}_m^{\perp}$ orthogonal to $\mathscr{V}_{m+1} \cap \mathscr{W}_m^{\perp}$, both of which are one-dimensional spaces of which $\{\boldsymbol{v}_{m+1}\}$ and $\{\widehat{\boldsymbol{w}}_{m+1}\}$ are biorthonormal bases. Thus, (2.47) is the singular value expansion of $I - P_{\mathscr{W}_m}^{\mathscr{V}_m}$ restricted to $\mathscr{V}_{m+1}$.

To obtain the coordinate vector $\widehat{\boldsymbol{y}}_{m+1}$, note first that the condition $(\boldsymbol{v}_{m+1}, \widehat{\boldsymbol{w}}_{m+1}) = 1$ implies that its last component is equal to one. Next, $\widehat{\boldsymbol{w}}_{m+1} \perp \mathscr{W}_m$ translates to $\widehat{\boldsymbol{y}}_{m+1} \in \mathscr{N}(\widetilde{H}_m^H)$, since the columns of $\widetilde{H}_m$ span the coordinate space of $\mathscr{W}_m$. Denoting by $\boldsymbol{g}_m \in \mathbb{C}^m$ the first m components of $\widehat{\boldsymbol{y}}_{m+1}$ and recalling that $\eta_{m+1,m} > 0$, we obtain

$$\mathbf{0} = \widetilde{H}_m^H \widehat{\boldsymbol{y}}_{m+1} = \begin{bmatrix} H_m^H & \eta_{m+1,m}\boldsymbol{u}_m \end{bmatrix} \begin{bmatrix} \boldsymbol{g}_m \\ 1 \end{bmatrix} = H_m^H \boldsymbol{g}_m + \eta_{m+1,m}\boldsymbol{u}_m.$$

□

The representation (2.47) can be used to obtain another expression for the OR approximation error as follows: by virtue of the inclusion $\mathscr{W}_{m-1} \subset \mathscr{W}_m$, an arbitrary vector $\boldsymbol{w} \in \mathscr{W}_{m-1}$ must lie in the nullspace of $I - P_{\mathscr{W}_m}^{\mathscr{V}_m}$. Furthermore, the difference $\boldsymbol{r} - \boldsymbol{w} \in \operatorname{span}\{\boldsymbol{r}\} + \mathscr{W}_{m-1} = \mathscr{V}_m$ has a representation $\boldsymbol{r} - \boldsymbol{w} = V_m \boldsymbol{z}$ with $\boldsymbol{z} \in \mathbb{C}^m$. It follows that

$$\begin{aligned} \boldsymbol{d}_m^{\mathrm{OR}} &= (I - P_{\mathscr{W}_m}^{\mathscr{V}_m})\boldsymbol{r} = (I - P_{\mathscr{W}_m}^{\mathscr{V}_m})(\boldsymbol{r} - \boldsymbol{w}) \\ &= (\boldsymbol{r} - \boldsymbol{w}, \widehat{\boldsymbol{w}}_m)\boldsymbol{v}_{m+1} = \left(V_{m+1} \begin{bmatrix} \boldsymbol{z} \\ 0 \end{bmatrix}, V_{m+1} \begin{bmatrix} \boldsymbol{g}_m \\ 1 \end{bmatrix} \right) \boldsymbol{v}_{m+1} \\ &= (\boldsymbol{g}_m^H \boldsymbol{z})\boldsymbol{v}_{m+1}, \end{aligned} \tag{2.48}$$

and therefore $\|\boldsymbol{d}_m^{\mathrm{OR}}\| = |\boldsymbol{g}_m^H \boldsymbol{z}| \leq \|\boldsymbol{g}_m\|_2 \|\boldsymbol{z}\|_2$ with equality holding if and only if $\boldsymbol{g}$ and $\boldsymbol{z}$ are collinear. At the same time, as $\boldsymbol{g}_m$ is fixed, equality must occur when $\|\boldsymbol{z}\|_2 = \|\boldsymbol{r} - \boldsymbol{w}\|$ is minimized among all $\boldsymbol{w} \in \mathscr{W}_{m-1}$, which is the case for $\boldsymbol{w} = \boldsymbol{w}_{m-1}^{\mathrm{MR}}$. As a result, $\|\boldsymbol{d}_m^{\mathrm{OR}}\| = \|\boldsymbol{g}_m\|_2 \|\boldsymbol{d}_{m-1}^{\mathrm{MR}}\|$ which, in view of (2.24) and (2.26), implies $\|\boldsymbol{g}_m\|_2 = s_m/c_m$, an identity which could also have been derived directly from the definition of $\boldsymbol{g}_m$.

Angles and the QR-factorization of $\widetilde{H}_m$

The least-squares problem (2.44) can be solved with the help of a QR decomposition of the Hessenberg matrix $\widetilde{H}_m$,

$$\widetilde{H}_m = Q_m \begin{bmatrix} R_m \\ \mathbf{0} \end{bmatrix}, \tag{2.49}$$

where $Q_m \in \mathbb{C}^{(m+1)\times(m+1)}$ is unitary ($Q_m^H Q_m = I_{m+1}$) and $R_m \in \mathbb{C}^{m\times m}$ is upper triangular. Substituting this QR factorization in (2.44) yields

$$\begin{aligned} \min_{\boldsymbol{y}\in\mathbb{C}^m} \|\beta\boldsymbol{u}_1 - \widetilde{H}_m \boldsymbol{y}\|_2 &= \min_{\boldsymbol{y}\in\mathbb{C}^m} \left\| Q_m \left(\beta Q_m^H \boldsymbol{u}_1 - \begin{bmatrix} R_m \\ \mathbf{0} \end{bmatrix} \boldsymbol{y} \right) \right\|_2 \\ &= \min_{\boldsymbol{y}\in\mathbb{C}^m} \left\| \beta Q_m^H \boldsymbol{u}_1 - \begin{bmatrix} R_m \\ \mathbf{0} \end{bmatrix} \boldsymbol{y} \right\|_2 = \min_{\boldsymbol{y}\in\mathbb{C}^m} \left\| \begin{bmatrix} \beta\widetilde{Q}_m^H \boldsymbol{u}_1 - R_m \boldsymbol{y} \\ \beta\overline{q}_{1,m+1}^{(m)} \end{bmatrix} \right\|_2, \end{aligned}$$

where

$$\widetilde{Q}_m := Q_m \begin{bmatrix} I_m \\ \mathbf{0} \end{bmatrix} \in \mathbb{C}^{(m+1)\times m} \tag{2.50}$$

denotes the first m columns of Q_m. Since $\widetilde{H}_m$ has full rank, R_m is nonsingular and the solution of the above least-squares problem is $\boldsymbol{y}_m^{\mathrm{MR}} = \beta R_m^{-1} \widetilde{Q}_m^H \boldsymbol{u}_1$. The associated least-squares error is given by $\|\boldsymbol{d}_m^{\mathrm{MR}}\| = \beta |q_{1,m+1}^{(m)}|$.

The following theorem links the angles between $\boldsymbol{r}$ and $\mathscr{W}_m$ as well as those between the spaces $\mathscr{V}_m$ and $\mathscr{W}_m$ with quantities appearing in the first row of the matrix Q_m:

Theorem 2.3.5. *If, for $m = 1, \ldots, L$, $Q_m = [q_{j,k}^{(m)}]_{j,k=1}^{m+1} \in \mathbb{C}^{(m+1)\times(m+1)}$ is the unitary matrix in the QR decomposition* (2.49) *of the Hessenberg matrix $\widetilde{H}_m$ in* (2.39)*, then there holds*

$$\sin \measuredangle(\boldsymbol{r}, \mathscr{W}_m) = \left| q_{1.m+1}^{(m)} \right|, \tag{2.51}$$

$$\sin \measuredangle(\boldsymbol{d}_{m-1}^{MR}, \mathscr{W}_m) = \sin \measuredangle(\mathscr{V}_m, \mathscr{W}_m) = \left| q_{1,m+1}^{(m)} / q_{1,m}^{(m-1)} \right|. \tag{2.52}$$

Proof. As mentioned earlier (cf. the proof of Lemma 2.3.1) the vector $\boldsymbol{r}$ possesses the coordinates $\beta \boldsymbol{u}_1^{(m+1)}$ with respect to the orthonormal basis $\{\boldsymbol{v}_1, \ldots \boldsymbol{v}_{m+1}\}$ of $\mathscr{V}_{m+1}$, whereas $\mathscr{W}_m$ is represented by $\mathscr{R}(\widetilde{H}_m) \subset \mathbb{C}^{m+1}$. This implies

$$\measuredangle(\boldsymbol{r}, \mathscr{W}_m) = \measuredangle_2(\beta \boldsymbol{u}_1, \mathscr{R}(\widetilde{H}_m)) = \measuredangle_2(\boldsymbol{u}_1, \mathscr{R}(\widetilde{H}_m)),$$

where the index 2 indicates that the last two angles are defined with respect to the Euclidean inner product on $\mathbb{C}^{m+1}$.

The vectors $[\boldsymbol{v}_1, \ldots \boldsymbol{v}_{m+1}] Q_m$ form an alternate orthonormal basis of $\mathscr{V}_{m+1}$, with respect to which $\boldsymbol{r}$ possesses the coordinate vector $\beta Q_m^H \boldsymbol{u}_1 \in \mathbb{C}^{m+1}$, a multiple of the first column of Q_m^H. The vectors in $\mathscr{W}_m$ are represented by $Q_m^H \widetilde{H}_m \boldsymbol{y} = \left[\begin{smallmatrix} R_m \boldsymbol{y} \\ \mathbf{0} \end{smallmatrix} \right]$ $(\boldsymbol{y} \in \mathbb{C}^m)$, a subspace of $\mathbb{C}^{m+1}$ which we identify with $\mathbb{C}^m$ because it consists of those vectors of $\mathbb{C}^{m+1}$ whose last component equals zero. Consequently, there holds

$$\measuredangle(\boldsymbol{r}, \mathscr{W}_m) = \measuredangle_2(\beta Q_m^H \boldsymbol{u}_1, \mathbb{C}^m) = \measuredangle_2(Q_m^H \boldsymbol{u}_1, \mathbb{C}^m)$$

which, in view of (2.4), proves assertion (2.51). Formula (2.52) follows directly from (2.16) and (2.21). □

The matrix Q_m is usually constructed as a product of Givens rotations such that

$$Q_m^H = G_m \begin{bmatrix} G_{m-1} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} G_{m-2} & \mathbf{0} \\ \mathbf{0} & I_2 \end{bmatrix} \cdots \begin{bmatrix} G_1 & \mathbf{0} \\ \mathbf{0} & I_{m-1} \end{bmatrix} \tag{2.53}$$

where, for $k = 1, 2, \ldots, m$,

$$G_k := \begin{bmatrix} I_{k-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \widetilde{c}_k & \widetilde{s}_k e^{-i\varphi_k} \\ \mathbf{0} & -\widetilde{s}_k e^{i\varphi_k} & \widetilde{c}_k \end{bmatrix}, \qquad \widetilde{c}_k, \widetilde{s}_k \geq 0, \quad \widetilde{c}_k^2 + \widetilde{s}_k^2 = 1, \varphi_k \in \mathbb{R}, \tag{2.54}$$

is chosen to introduce a zero in the k-th subdiagonal position in the process of unitarily transforming $\widetilde{H}_m$ to upper triangular form. The details of the m-th rotation are as follows: Suppose we have constructed $G_1, \ldots, G_{m-2}, G_{m-1}$ such that

$$\begin{bmatrix} G_{m-1} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} G_{m-2} & \mathbf{0} \\ \mathbf{0} & I_2 \end{bmatrix} \cdots \begin{bmatrix} G_1 & \mathbf{0} \\ \mathbf{0} & I_{m-1} \end{bmatrix} \widetilde{H}_m = \begin{bmatrix} R_{m-1} & \boldsymbol{r} \\ \mathbf{0} & \tau \\ \mathbf{0} & \eta_{m+1,m} \end{bmatrix}. \tag{2.55}$$

For later use we rewrite this identity as

$$\begin{bmatrix} Q_{m-1}^H & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} H_m \\ 0 \cdots 0 \; \eta_{m+1,m} \end{bmatrix} = \begin{bmatrix} R_{m-1} & \boldsymbol{r} \\ \mathbf{0} & \tau \\ \mathbf{0} & \eta_{m+1,m} \end{bmatrix}. \tag{2.56}$$

We now set

$$\begin{aligned} \widetilde{c}_m &:= \frac{|\tau|}{\sqrt{|\tau|^2 + \eta_{m+1,m}^2}}, \quad \widetilde{s}_m := \frac{\eta_{m+1,m}}{\sqrt{|\tau|^2 + \eta_{m+1,m}^2}}, \\ \varphi_m &:= \arg(\eta_{m+1,m}) - \arg(\tau) = -\arg(\tau) \end{aligned} \tag{2.57}$$

(recall $\eta_{m+1,m} \geq 0$) and verify by a simple calculation that

$$\begin{bmatrix} I_{m-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \widetilde{c}_m & \widetilde{s}_m e^{-i\varphi_m} \\ \mathbf{0} & -\widetilde{s}_m e^{i\varphi_m} & \widetilde{c}_m \end{bmatrix} \begin{bmatrix} R_{m-1} & \boldsymbol{r} \\ \mathbf{0} & \tau \\ \mathbf{0} & \eta_{m+1,m} \end{bmatrix} = \begin{bmatrix} R_{m-1} & \boldsymbol{r} \\ \mathbf{0} & \rho \\ \mathbf{0} & 0 \end{bmatrix}$$

with $\rho = \sqrt{|\tau|^2 + \eta_{m+1,m}^2}\, e^{-i\varphi_m}$.

To see that the quantities $\widetilde{s}_m$ and $\widetilde{c}_m$ are indeed the sines and cosines of the angles $\measuredangle(\boldsymbol{d}_{m-1}^{\mathrm{MR}}, \mathscr{W}_m) = \measuredangle(\mathscr{V}_m, \mathscr{W}_m)$, note that $q_{1,m+1}^{(m)} = -\widetilde{s}_m e^{-i\varphi_m} q_{1,m}^{(m-1)}$, which, with Theorem 2.3.5, yields

$$s_m = \measuredangle(\boldsymbol{d}_{m-1}^{\mathrm{MR}}, \mathscr{W}_m) = |q_{1,m+1}^{(m)} / q_{1,m}^{(m-1)}| = \widetilde{s}_m.$$

The Paige-Saunders Basis

The alternate orthonormal basis of $\mathscr{V}_{m+1}$ which occurred in the proof of Theorem 2.3.5 turns out to be extremely useful when describing the MR and OR approximations. We therefore introduce the notation

$$\widetilde{\widehat{V}}_{m+1} := [\widehat{\boldsymbol{v}}_1^{(m+1)}, \ldots, \widehat{\boldsymbol{v}}_{m+1}^{(m+1)}] := V_{m+1} Q_m$$

for this basis and, since it was first employed by Paige & Saunders (1975), refer to it as the *Paige-Saunders basis*. This notation for the Paige-Saunders basis vectors is not entirely appropriate, since all but the last do not change with the index m, as shown in the following proposition.

Proposition 2.3.6. *There holds*

$$\left[\widehat{\boldsymbol{v}}_1^{(m+1)}, \ldots, \widehat{\boldsymbol{v}}_{m+1}^{(m+1)}\right] \equiv [\widehat{\boldsymbol{v}}_1, \ldots, \widehat{\boldsymbol{v}}_m, \widetilde{\boldsymbol{v}}_{m+1}],$$

where $\widetilde{\boldsymbol{v}}_1 = \boldsymbol{v}_1$, *and, for* $m = 1, \ldots, L-1$,

$$\widehat{\boldsymbol{v}}_m = \quad c_m \widetilde{\boldsymbol{v}}_m + s_m e^{i\varphi_m} \boldsymbol{v}_{m+1}, \tag{2.58}$$

$$\widetilde{\boldsymbol{v}}_{m+1} = -s_m e^{-i\varphi_m} \widetilde{\boldsymbol{v}}_m + \quad c_m \boldsymbol{v}_{m+1}. \tag{2.59}$$

The vectors $\{\widehat{\boldsymbol{v}}_1, \ldots, \widehat{\boldsymbol{v}}_m, \widetilde{\boldsymbol{v}}_{m+1}\}$ *form an orthonormal basis of* $\mathscr{V}_{m+1}$ *whose first* m *vectors* $\{\widehat{\boldsymbol{v}}_1, \ldots, \widehat{\boldsymbol{v}}_m\}$ *form an orthonormal basis of* $\mathscr{W}_m$. *In addition, there holds*

$$\boldsymbol{w}_m^{MR} = \beta \widetilde{\widehat{V}}_{m+1} \begin{bmatrix} \widetilde{Q}_m^H \boldsymbol{u}_1^{(m+1)} \\ 0 \end{bmatrix} \quad \textit{and} \quad \boldsymbol{d}_m^{MR} = \beta \widetilde{\widehat{V}}_{m+1} \begin{bmatrix} \mathbf{0} \\ \overline{q}_{1,m+1}^{(m)} \end{bmatrix}. \tag{2.60}$$

Proof. To prove the first assertion we observe

$$\begin{aligned}\widetilde{\widehat{V}}_{m+1} = V_{m+1}Q_m &= [V_m, \boldsymbol{v}_{m+1}] \begin{bmatrix} Q_{m-1} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} G_m^H \\ &= \left[\widetilde{\widehat{V}}_m, \boldsymbol{v}_{m+1}\right] \begin{bmatrix} I_{m-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & c_m & -s_m e^{-i\varphi_m} \\ \mathbf{0} & s_m e^{i\varphi_m} & c_m \end{bmatrix}.\end{aligned}$$

That the first m elements of $\widetilde{\widehat{V}}_{m+1}$ form a basis of the approximation space $\mathscr{W}_m$ follows from

$$W_m = V_{m+1}\widetilde{H}_m = V_{m+1}Q_m \begin{bmatrix} R_m \\ \mathbf{0} \end{bmatrix} = [\widehat{\boldsymbol{v}}_1, \dots, \widehat{\boldsymbol{v}}_m]\, R_m. \tag{2.61}$$

The discussion following (2.49) revealed that the coordinate vector of $\boldsymbol{w}_m^{\mathrm{MR}}$ with respect to $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$ is $\boldsymbol{y}_m^{\mathrm{MR}} = \beta R_m^{-1}\widetilde{Q}_m^H \boldsymbol{u}_1^{(m+1)}$, and this, together with (2.61) implies (2.60). □

Remark 2.3.7. We conclude from (2.59) that the vector $\widehat{\boldsymbol{w}}_{m+1}$ introduced in Lemma 2.3.4 is given by $\widetilde{\boldsymbol{v}}_{m+1}/c_m$. An equivalent formulation of (2.47) therefore reads

$$(I - P_{\mathscr{W}_m}^{\mathscr{V}_m})\boldsymbol{v} = \frac{(\boldsymbol{v}, \widetilde{\boldsymbol{v}}_{m+1})}{c_m}\boldsymbol{v}_{m+1} \quad \text{for all } \boldsymbol{v} \in \mathscr{V}_{m+1}.$$

The following proposition summarizes the coordinate representations of the MR and OR errors with respect to the two orthonormal bases of $\mathscr{V}_{m+1}$.

Proposition 2.3.8. *The MR and OR approximation errors satisfy:*

$$\begin{aligned}\boldsymbol{d}_m^{MR} &= \beta\, \overline{q}_{1,m+1}^{(m)} \widetilde{\boldsymbol{v}}_{m+1} = \beta \prod_{j=1}^{m} \left(-s_j e^{i\varphi_j}\right) \widetilde{\boldsymbol{v}}_{m+1}, \\ \boldsymbol{d}_m^{OR} &= -\beta\, \frac{s_m}{c_m} e^{i\varphi_m} \overline{q}_{1,m}^{(m-1)} \boldsymbol{v}_{m+1} = \frac{\beta}{c_m} \prod_{j=1}^{m} \left(-s_j e^{i\varphi_j}\right) \boldsymbol{v}_{m+1}, \\ \boldsymbol{d}_{m-1}^{MR} - \boldsymbol{d}_m^{OR} &= \frac{\beta}{c_m} \overline{q}_{1,m}^{(m-1)} \widehat{\boldsymbol{v}}_m = \frac{\beta}{c_m} \prod_{j=1}^{m-1} \left(-s_j e^{i\varphi_j}\right) \widehat{\boldsymbol{v}}_m.\end{aligned}$$

Proof. The recursive definition (2.53), (2.54) of Q_m allows us to express its entries $q_{1,k}^{(m)}$ explicitly in terms of the Givens parameters as

$$q_{1,k}^{(m)} = c_k \prod_{j=1}^{k-1} \left(-s_j e^{-i\varphi_j}\right), \quad 1 \le k \le m, \qquad\qquad q_{1,m+1}^{(m)} = \prod_{j=1}^{m} \left(-s_j e^{-i\varphi_j}\right). \tag{2.62}$$

This, together with (2.60), proves the first identity.

Next, we recall from Remark 2.3.3 that $\boldsymbol{d}_m^{\mathrm{OR}} = -\beta\eta_{m+1,m}\, \eta_{m,1}^{[-1]} \boldsymbol{v}_{m+1}$. To eliminate $\eta_{m,1}^{[-1]}$ from this relation, we note that the matrix H_m possesses the QR decomposition (cf. (2.56))

$$Q_{m-1}^H H_m = \begin{bmatrix} R_{m-1} & \boldsymbol{r} \\ \mathbf{0} & \tau \end{bmatrix}, \quad \text{i.e.,} \quad H_m^{-1} = \begin{bmatrix} R_{m-1}^{-1} & \widetilde{\boldsymbol{r}} \\ \mathbf{0} & 1/\tau \end{bmatrix} Q_{m-1}^H,$$

which implies $\eta_{m,1}^{[-1]} = \overline{q}_{1,m}^{(m-1)}/\tau$. Since $\eta_{m+1,m}/\tau = e^{i\varphi_m} s_m/c_m$ (cf. (2.57)) we conclude $\eta_{m,1}^{[-1]}\eta_{m+1,m} = q_{m,1}^{(m-1)} e^{i\phi_m} s_m/c_m$. This proves the second identity.

The desired representation of $\boldsymbol{d}_{m-1}^{\mathrm{MR}} - \boldsymbol{d}_m^{\mathrm{OR}}$ now follows from (2.58). □

We note that all relations contained in Theorems 2.2.5 and 2.2.6 which link the MR and OR approaches can just as well be obtained by manipulating the error representations of Proposition 2.3.8. Indeed, this is essentially how these relations have previously been proven in the literature. The main difference to the approach taken in Chapter 2.2 is that the sines and cosines which occur in these relations are usually taken to be those from the Givens rotations needed to construct the QR decomposition of $\widetilde{H}_m$. By identifying these parameters as the sines and cosines of $\measuredangle(\mathscr{V}_m, \mathscr{W}_m)$, these are seen to express intrinsic relations between these spaces rather than being mere artifacts of the algorithm.

The utility of the Paige-Saunders basis $\widetilde{\widehat{V}}_{m+1}$ lies in the fact that its first m elements $V_{m+1}\widetilde{Q}_m$, which we denote for future reference by

$$\widehat{V}_m := [\widehat{\boldsymbol{v}}_1, \ldots, \widehat{\boldsymbol{v}}_m], \tag{2.63}$$

constitute an orthonormal basis of the approximation space $\mathscr{W}_m$, in terms of which the best approximation $\boldsymbol{w}_m^{\mathrm{MR}}$ of $\boldsymbol{r}$ from $\mathscr{W}_m$ is given by the truncated Fourier expansion $\boldsymbol{w}_m^{\mathrm{MR}} = \sum_{j=1}^m (\boldsymbol{r}, \widehat{\boldsymbol{v}}_j)\widehat{\boldsymbol{v}}_j$. On the other hand, introducing the notation $V_{m+1}^*\boldsymbol{r}$ to denote the column vector $[(\boldsymbol{r}, \boldsymbol{v}_1), \ldots, (\boldsymbol{r}, \boldsymbol{v}_{m+1})]^\top \in \mathbb{C}^{m+1}$, we also have

$$\boldsymbol{w}_m^{\mathrm{MR}} = V_{m+1}\widetilde{Q}_m\widetilde{Q}_m^H V_{m+1}^*\boldsymbol{r} = \beta\widehat{V}_m\widetilde{Q}_m^H\boldsymbol{u}_1^{(m+1)} = \beta\sum_{j=1}^m \overline{q}_{1,j}^{(m)}\widehat{\boldsymbol{v}}_j. \tag{2.64}$$

When the termination index $m = L$ is reached, we have $W_L = V_L H_L = V_L Q_{L-1} R_L$, and we thus conclude that the Fourier coefficients of $\boldsymbol{r}$ with respect to the Paige-Saunders basis at step L are given by

$$(\boldsymbol{r}, \widehat{\boldsymbol{v}}_j) = \overline{q}_{1,j}^{(L-1)} = \overline{q}_{1,j}^{(j)}, \qquad j = 1, \ldots, L,$$

where we have set $q_{1,L}^{(L)} := q_{1,L}^{(L-1)}$. For future reference, we observe that, in view of (2.62), the vector of Fourier coefficients $\boldsymbol{z}_m := \widehat{V}_m^*\boldsymbol{r} = \widetilde{Q}_m^H\boldsymbol{u}_1^{(m+1)}$ of $\boldsymbol{r}$ with respect to $\{\widehat{\boldsymbol{v}}_1, \ldots, \widehat{\boldsymbol{v}}_m\}$ satisfies the recurrence

$$\boldsymbol{z}_m = \begin{bmatrix} \boldsymbol{z}_{m-1} \\ c_m\zeta_m \end{bmatrix}, \quad \zeta_m = \overline{q}_{1,m}^{(m-1)} = -s_{m-1}e^{i\varphi_{m-1}}\zeta_{m-1}, \quad m = 2, \ldots, L, \tag{2.65}$$

with $\boldsymbol{z}_1 = [c_1]$, $\zeta_1 = -s_1 e^{i\varphi_1}$.

The transformation which takes the vectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ to $\{\widehat{\boldsymbol{v}}_1, \ldots, \widehat{\boldsymbol{v}}_m\}$ has a simple geometric description: by Proposition 2.3.6, it consists of two plane rotations for each basis vector except $\boldsymbol{v}_1$ and $\boldsymbol{v}_L$, which each undergo only one rotation. Beginning with the orthonormal basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ of $\mathscr{V}_2 = \mathrm{span}\{\boldsymbol{r}\} + \mathscr{W}_1$, $\mathscr{W}_1 = \mathrm{span}\{\boldsymbol{w}_1\}$, the first rotation occurs in the plane spanned by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ and is such that $\boldsymbol{v}_1$ is rotated so as to be collinear with $\boldsymbol{w}_1$. The intermediate vector $\widetilde{\boldsymbol{v}}_2$ results from applying this same rotation to $\boldsymbol{v}_2$. In the next step, applying the first rotation to the orthonormal basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ of $\mathscr{V}_3 = \mathrm{span}\{\boldsymbol{r}\} + \mathscr{W}_2$ yields the orthonormal vectors $\{\widehat{\boldsymbol{v}}_1, \widetilde{\boldsymbol{v}}_2, \boldsymbol{v}_3\}$. Since $\widehat{\boldsymbol{v}}_1$ spans $\mathscr{W}_1$,

there must exist a vector $\widehat{\boldsymbol{v}}_2 \in \operatorname{span}\{\widetilde{\boldsymbol{v}}_2, \boldsymbol{v}_3\}$ such that $\mathscr{W}_2 = \operatorname{span}\{\widehat{\boldsymbol{v}}_1, \widehat{\boldsymbol{v}}_2\}$, and the second rotation is chosen as that in the plane spanned by $\widetilde{\boldsymbol{v}}_2$ and $\boldsymbol{v}_3$ which takes $\widetilde{\boldsymbol{v}}_2$ into $\widehat{\boldsymbol{v}}_2$, and $\widetilde{\boldsymbol{v}}_3$ is the result of its action on $\boldsymbol{v}_3$. This process continues in this manner up to the termination index $m = L$, for which $\boldsymbol{r} \in \mathscr{W}_L = \mathscr{V}_L$, and we see that $\{\widehat{\boldsymbol{v}}_1 \ldots, \widehat{\boldsymbol{v}}_{L-1}, \widetilde{\boldsymbol{v}}_L\}$ must span $\mathscr{W}_L$, and thus we set $\widehat{\boldsymbol{v}}_L := \widetilde{\boldsymbol{v}}_L$.

2.3.3 Using Arbitrary Bases of $\mathscr{W}_m$ and $\mathscr{V}_m$

For practical computations it is desirable that the matrices $\widetilde{H}_m$ have small bandwidth. If $\widetilde{H}_m$ has only k nonvanishing diagonals, namely the ones with indices $-1, 0, 1, \ldots, k-2$ (we follow the standard notation according to which a diagonal has index k if its entries $\eta_{j,\ell}$ are characterized by $\ell - j = k$), then only k diagonals of the upper triangular matrices R_m are nonzero, namely those with indices $k = 0, 1, \ldots, k-1$. This follows easily from the fact that Q_m has upper Hessenberg form. The banded structure of the matrices R_m can then be used to derive k-term recurrence formulas for the coordinate vectors $\boldsymbol{y}_m$ in terms of $\boldsymbol{y}_{m-1}, \boldsymbol{y}_{m-2}, \ldots, \boldsymbol{y}_{m-k+1}$ and for the approximations $\boldsymbol{w}_m^{\mathrm{MR/OR}}$ in terms of $\boldsymbol{w}_{m-1}^{\mathrm{MR/OR}}, \boldsymbol{w}_{m-2}^{\mathrm{MR/OR}}, \ldots, \boldsymbol{w}_{m-k+1}^{\mathrm{MR/OR}}$. (This statement applies to both the MR and the OR approach.) The most important consequence of this observation is that, at each step, only the k previous approximations (or, in other implementations, the last k basis vectors) need to be stored, which means that storage requirements do not increase with m. If we insist on choosing $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_L$ to be orthogonal vectors, then the Hessenberg matrices will generally not have banded form. The main motivation for doing without an orthonormal basis of $\mathscr{V}_m$ is therefore to constrain the bandwidth of H_m in order to keep storage requirements low.

A Basis-dependent Inner Product

As explained at the beginning of Section 2.3.2, no orthogonality conditions are required to derive the fundamental relationship (2.39)

$$W_m = V_{m+1}\widetilde{H}_m = V_m H_m + [\boldsymbol{0}, \ldots, \boldsymbol{0}, \eta_{m+1,m}\boldsymbol{v}_{m+1}].$$

In this section we require only that $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots$ be linearly independent, so that $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ constitutes a basis of $\mathscr{V}_m$ for each $m = 1, 2, \ldots, L$. Just as before, the j-th column of the upper Hessenberg matrix $\widetilde{H}_m \in \mathbb{C}^{(m+1)\times m}$ contains the coefficients of $\boldsymbol{w}_j \in \mathscr{W}_j \subset \mathscr{V}_{m+1}$ with respect to the basis vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m+1}$. The difference is that, since now the vectors $\boldsymbol{v}_j$ need not be orthogonal, these coefficients can no longer be expressed in terms of the inner product with which $\mathscr{H}$ was originally endowed. We shall see below, however, that the familiar inner product representation does indeed still hold, but with respect to a different inner product. As in the proof of Lemma 2.3.1 we see that, for each $\boldsymbol{w} = W_m\boldsymbol{y} \in \mathscr{W}_m$ ($\boldsymbol{y} \in \mathbb{C}^m$), the associated error is represented by

$$\boldsymbol{d} = \boldsymbol{r} - \boldsymbol{w} = \boldsymbol{r} - W_m\boldsymbol{y} = \beta\boldsymbol{v}_1 - V_{m+1}\widetilde{H}_m\boldsymbol{y} = V_{m+1}(\beta\boldsymbol{u}_1 - \widetilde{H}_m\boldsymbol{y}).$$

Minimizing the norm of $\boldsymbol{d}$ among all $\boldsymbol{w} \in \mathscr{W}_m$ leads, as before, to the least squares problem

$$\min_{\boldsymbol{y}\in\mathbb{C}^m} \left\| V_{m+1}\left(\beta\boldsymbol{u}_1 - \widetilde{H}_m\boldsymbol{y}\right)\right\| = \min_{\boldsymbol{y}\in\mathbb{C}^m} \|\beta\boldsymbol{u}_1 - \widetilde{H}_m\boldsymbol{y}\|_{\boldsymbol{v}}, \tag{2.66}$$

in which $\|\cdot\|_{\boldsymbol{v}}$ denotes the norm induced on the coordinate space with respect to the basis V_L by the inner product $(\cdot,\cdot)$ given on $\mathscr{H}$. More precisely, if we set for $\boldsymbol{y},\boldsymbol{z}\in\mathbb{C}^L$,

$$(\boldsymbol{y},\boldsymbol{z})_{\boldsymbol{v}} := (V_L\boldsymbol{y},V_L\boldsymbol{z}) = \boldsymbol{z}^H M\boldsymbol{y}, \qquad \text{where } M := [(\boldsymbol{v}_j,\boldsymbol{v}_k)]\in\mathbb{C}^{L\times L}, \tag{2.67}$$

then $(\cdot,\cdot)_{\boldsymbol{v}}$ is an inner product on $\mathbb{C}^L$ and induces a norm, namely $\|\cdot\|_{\boldsymbol{v}} := (\cdot,\cdot)_{\boldsymbol{v}}^{1/2}$, on $\mathbb{C}^L$.

At this point one could proceed as in the algorithms which use an orthogonal basis, the only difference being that all inner products in the coordinate space now require knowledge of the Gram matrix M. In particular, the Givens rotations and the matrices Q_m in the QR factorizations (2.49) must now be unitary with respect to the inner product $(\cdot,\cdot)_{\boldsymbol{v}}$, i.e., they must satisfy $Q_m^H M_m Q_m = I_m$, where $M_m\in\mathbb{C}^{(m+1)\times(m+1)}$ is the $(m+1)$-st leading principal submatrix of M. The submatrices M_m, however, *cannot* be computed unless all basis vectors $\boldsymbol{v}_1,\boldsymbol{v}_2,\ldots,\boldsymbol{v}_{m+1}$ are available, regardless of whether these may obey short recurrences. Unfortunately, this is precisely what we sought to avoid by giving up the orthogonality of the vectors $\boldsymbol{v}_j$.

Quasi-Minimal Approximations

An alternative was proposed by Freund (1992*b*): Rather than solving the minimization problem (2.66), we instead solve

$$\min_{\boldsymbol{y}\in\mathbb{C}^m}\|\beta\boldsymbol{u}_1-\widetilde{H}_m\boldsymbol{y}\|_2,$$

and, if $\boldsymbol{y}_m^{\mathrm{QMR}}\in\mathbb{C}^m$ denotes the unique solution to this least-squares problem, regard

$$\boldsymbol{w}_m^{\mathrm{QMR}} := W_m\boldsymbol{y}_m^{\mathrm{QMR}}$$

as an approximation of $\boldsymbol{r}$. Using the terminology introduced by Freund, we refer to this approach as the *quasi-minimal residual* (QMR) approximation. In the context of Krylov subspace methods the vector $\boldsymbol{s}_m^{\mathrm{QMR}} := \beta\boldsymbol{u}_1-\widetilde{H}_m\boldsymbol{y}_m^{\mathrm{QMR}}\in\mathbb{C}^{m+1}$ is usually called the *quasi-residual* of $\boldsymbol{w}_m^{\mathrm{QMR}}$.

We note that, instead of changing the inner product in the coordinate space from $(\cdot,\cdot)_{\boldsymbol{v}}$ to the Euclidean inner product, one could equivalently have replaced the given inner product $(\cdot,\cdot)$ on $\mathscr{V}_L\subseteq\mathscr{H}$ by

$$(\boldsymbol{u},\boldsymbol{v})_V = (V_L\boldsymbol{x},V_L\boldsymbol{y})_V =: \boldsymbol{y}^H\boldsymbol{x} \qquad \text{for all } \boldsymbol{u}=V_L\boldsymbol{x},\boldsymbol{v}=V_L\boldsymbol{y}\in\mathscr{V}_L \tag{2.68}$$

and proceeded as in the MR algorithm of Section 2.3.2. The basis vectors $\{\boldsymbol{v}_1,\ldots,\boldsymbol{v}_L\}$ are orthonormal with respect to the new inner product $(\cdot,\cdot)_V$ thus defined, and this is its essential feature.

The assertions in the remainder of this section rest on the following basic identities for the new inner product which are direct consequences of its definition: for all coordinate vectors $\boldsymbol{x},\boldsymbol{y}\in\mathbb{C}^L$, there holds

$$\begin{aligned}(V_L\boldsymbol{x},V_L\boldsymbol{y})_V &= (V_L\boldsymbol{x},V_LM^{-1}\boldsymbol{y}) = (V_LM^{-1}\boldsymbol{x},V_L\boldsymbol{y}) = (V_LM^{-1/2}\boldsymbol{x},V_LM^{-1/2}\boldsymbol{y}),\\ (V_L\boldsymbol{x},V_L\boldsymbol{y}) &= (V_L\boldsymbol{x},V_LM\boldsymbol{y})_V = (V_LM\boldsymbol{x},V_L\boldsymbol{y})_V = (V_LM^{1/2}\boldsymbol{x},V_LM^{1/2}\boldsymbol{y})_V.\end{aligned} \tag{2.69}$$

As a consequence, we obtain for instance

Theorem 2.3.9. *The QMR iterates are the MR iterates with respect to the inner product* $(\cdot,\cdot)_V$*:*

$$\|\boldsymbol{d}_m^{QMR}\|_V = \|\boldsymbol{r} - \boldsymbol{w}_m^{QMR}\|_V = \min_{\boldsymbol{w}\in\mathscr{W}_m} \|\boldsymbol{r} - \boldsymbol{w}\|_V.$$

With regard to the original norm on $\mathscr{H}$*, the QMR errors may be bounded in terms of the MR errors as*

$$\|\boldsymbol{d}_m^{MR}\| \le \|\boldsymbol{d}_m^{QMR}\| \le \sqrt{\kappa_2(M_m)}\,\|\boldsymbol{d}_m^{MR}\|, \tag{2.70}$$

in which $\kappa_2(M_m)$ *denotes the (Euclidean) condition number of the (Hermitian positive definite) matrix* M_m*. Moreover, we have*

$$\|\boldsymbol{d}_m^{QMR}\| \le \sqrt{\lambda_{\max}(M_m)}\,\|\boldsymbol{s}_m^{QMR}\|.$$

Proof. The minimization property in the new norm is by construction, and the first inequality in (2.70) is the minimization property of MR in the original norm. The second inequality follows when we recall that, by (2.69), for $\boldsymbol{u} = V_{m+1}\boldsymbol{x}$, $\boldsymbol{x}\in\mathbb{C}^{m+1}$, we have

$$\|\boldsymbol{u}\|^2 = (V_{m+1}\boldsymbol{x}, V_{m+1}\boldsymbol{x}) = \boldsymbol{x}^H M_m \boldsymbol{x} \quad \text{and} \quad \|\boldsymbol{u}\|_V^2 = (V_{m+1}\boldsymbol{x}, V_{m+1}\boldsymbol{x})_V = \boldsymbol{x}^H\boldsymbol{x},$$

and therefore $\lambda_{\min}(M_m) \le \|\boldsymbol{u}\|^2/\|\boldsymbol{u}\|_V^2 \le \lambda_{\max}(M_m)$. This, along with the optimality properties of MR and QMR in their respective norms now yields the chain of inequalities

$$\begin{aligned}\|\boldsymbol{d}_m^{\mathrm{QMR}}\|^2 &\le \lambda_{\max}(M_m)\|\boldsymbol{d}_m^{\mathrm{QMR}}\|_V^2 \le \lambda_{\max}(M_m)\|\boldsymbol{d}_m^{\mathrm{MR}}\|_V^2\\ &\le \frac{\lambda_{\max}(M_m)}{\lambda_{\min}(M_m)}\|\boldsymbol{d}_m^{\mathrm{MR}}\|^2 \le \frac{\lambda_{\max}(M_m)}{\lambda_{\min}(M_m)}\|\boldsymbol{d}_m^{\mathrm{QMR}}\|^2.\end{aligned}$$

Similarly, for the last assertion,

$$\|\boldsymbol{d}_m^{\mathrm{QMR}}\|^2 = (V_{m+1}\boldsymbol{s}_m^{\mathrm{QMR}}, V_{m+1}\boldsymbol{s}_m^{\mathrm{QMR}}) = [\boldsymbol{s}_m^{\mathrm{QMR}}]^H M_m \boldsymbol{s}_m^{\mathrm{QMR}} \le \lambda_{\max}(M_m)\|\boldsymbol{s}_m^{\mathrm{QMR}}\|^2.$$

□

In view of (2.70) the deviation of the QMR approach from the MR approach is bounded by the condition numbers $\kappa_2(M_m)$, i.e., by the ratio of the extreme eigenvalues of M_m. The largest eigenvalue $\lambda_{\max}(M_m)$ is easily controlled: this merely requires choosing the basis vectors $\boldsymbol{v}_m$ to have unit length—i.e., $\|\boldsymbol{v}_m\| = 1$ for all m—to guarantee $\lambda_{\max}(M_m) \le m+1$. (Note that $\lambda_{\max}(M_m) \le \|M_m\|_F := [\sum_{j,k=1}^{m+1}(\boldsymbol{v}_j,\boldsymbol{v}_k)^2]^{1/2} \le [\sum_{j,k=1}^{m+1}\|\boldsymbol{v}_j\|\,\|\boldsymbol{v}_k)\|]^{1/2}$.) The crucial point is to construct the basis V_m in such a way that $\lambda_{\min}(M_m)$ does not approach zero (or does so only slowly).

Another immediate consequence of (2.69) is the following characterization of the QMR approximation as an *oblique* projection with respect to the original inner product.

Proposition 2.3.10. *The error vectors of the QMR approach satisfy*

$$\boldsymbol{d}_m^{QMR} \perp \mathscr{U}_m,$$

where $\mathscr{U}_m := \{\boldsymbol{v} = V_{m+1}\boldsymbol{y} \,:\, \boldsymbol{y} = M_m^{-1}\widetilde{H}_m\boldsymbol{z} \text{ for some } \boldsymbol{z}\in\mathbb{C}^m\}$ *is an* m*-dimensional subspace of* $\mathscr{V}_{m+1}$ *(for* $m = L$ *there holds* $\mathscr{U}_L = \mathscr{V}_L$*) and orthogonality is understood with respect to the original inner product* $(\cdot,\cdot)$ *on* $\mathscr{H}$*. Consequently,*

$$\boldsymbol{d}_m^{QMR} = \left(I - P_{\mathscr{W}_m}^{\mathscr{U}_m}\right)\boldsymbol{r},$$

where $P_{\mathscr{W}_m}^{\mathscr{U}_m}$ *denotes the oblique projection onto* $\mathscr{W}_m$ *orthogonal to* $\mathscr{U}_m$*.*

Proof. Since the QMR approximations are merely the MR approximations with respect to $(\cdot,\cdot)_V$, their errors $\boldsymbol{d}_m^{\mathrm{QMR}} \in \mathscr{V}_{m+1}$ are characterized by

$$\boldsymbol{d}_m^{\mathrm{QMR}} \perp_V \mathscr{W}_m.$$

From this observation and (2.69) the proof easily follows.

An equally simple alternative proof notes that $\boldsymbol{d}_m^{\mathrm{QMR}} = V_{m+1}\big(\beta\boldsymbol{u}_1 - \widetilde{H}_m \boldsymbol{y}_m^{\mathrm{QMR}}\big)$, where $\boldsymbol{y}_m^{\mathrm{QMR}}$ solves the least-squares problem $\|\beta\boldsymbol{u}_1 - \widetilde{H}_m\boldsymbol{y}\|_2 \to \min$. In other words, $\beta\boldsymbol{u}_1 - \widetilde{H}_m\boldsymbol{y}_m^{\mathrm{QMR}} \perp \mathscr{R}(\widetilde{H}_m)$, or equivalently, $\beta\boldsymbol{u}_1 - \widetilde{H}_m\boldsymbol{y}_m^{\mathrm{QMR}} \perp_{\boldsymbol{v}} M^{-1}\mathscr{R}(\widetilde{H}_m)$. □

Note that the orthogonal complement of $\mathscr{U}_m$ is given by $\mathscr{U}_m^{\perp} = \mathrm{span}\{\boldsymbol{d}_m^{\mathrm{QMR}}\} + \mathscr{V}_{m+1}^{\perp}$, i.e., $\mathscr{U}_m^{\perp} \oplus \mathscr{V}_m = \mathscr{H}$, which ensures that the oblique projection $P_{\mathscr{W}_m}^{\mathscr{U}_m}$ exists.

Quasi-Orthogonal Approximations

Next, we briefly describe the analogue of the OR approach for the case of a non-orthogonal basis. Instead of seeking $\boldsymbol{y} \in \mathbb{C}^m$ such that

$$0 = (V_{m+1}[\beta\boldsymbol{u}_1 - \widetilde{H}_m\boldsymbol{y}], \boldsymbol{v}_j) = (\beta\boldsymbol{u}_1 - \widetilde{H}_m\boldsymbol{y}, \boldsymbol{u}_j)_{\boldsymbol{v}}. \qquad j = 1, 2, \ldots, m,$$

which would lead to $\boldsymbol{r} - W_m\boldsymbol{y} \perp \mathscr{V}_m$, i.e., to a proper OR approximation, we determine $\boldsymbol{y}_m^{\mathrm{QOR}} \in \mathbb{C}^m$ such that

$$0 = (\beta\boldsymbol{u}_1 - \widetilde{H}_m\boldsymbol{y}_m^{\mathrm{QOR}}, \boldsymbol{u}_j)_2, \quad j = 1, 2, \ldots, m, \qquad \text{i.e.,} \quad H_m\boldsymbol{y}_m^{\mathrm{QOR}} = \beta\boldsymbol{u}_1, \tag{2.71}$$

provided H_m is nonsingular. The corresponding approximants to $\boldsymbol{r}$ are then defined by $\boldsymbol{w}_m^{\mathrm{QOR}} = W_m\boldsymbol{y}_m^{\mathrm{QOR}}$. In terms of the inner product $(\cdot,\cdot)_V$ on $\mathscr{V}_L$ they are characterized by

$$\boldsymbol{d}_m^{\mathrm{QOR}} = \boldsymbol{r} - \boldsymbol{w}_m^{\mathrm{QOR}} \perp_V \mathscr{V}_m,$$

i.e., as the OR iterates with respect to $(\cdot,\cdot)_V$. As the analogue to Proposition 2.3.10 we obtain

Proposition 2.3.11. *The errors of the QOR approximants satisfy*

$$\boldsymbol{d}_m^{QOR} \perp \mathscr{T}_m,$$

where $\mathscr{T}_m := \{\boldsymbol{v} = V_{m+1}\boldsymbol{y} \in \mathscr{V}_{m+1} : \boldsymbol{y} = M_m^{-1}[\boldsymbol{z}^T\ 0]^T$ *with* $\boldsymbol{z} \in \mathbb{C}^m\}$ *is an* m*-dimensional subspace of* $\mathscr{V}_{m+1}$ *(for* $m = L$ *there holds* $\mathscr{T}_L = \mathscr{V}_L$*) and orthogonality is understood with respect to the original inner product* $(\cdot,\cdot)$ *on* $\mathscr{H}$*. Consequently,*

$$\boldsymbol{d}_m^{QOR} = \big(I - P_{\mathscr{W}_m}^{\mathscr{T}_m}\big)\,\boldsymbol{r},$$

where $P_{\mathscr{W}_m}^{\mathscr{T}_m}$ *denotes the oblique projection onto* $\mathscr{W}_m$ *orthogonal to* $\mathscr{T}_m$*.*

We know from Remark 2.3.2 that H_m being nonsingular is equivalent $\mathscr{W}_m \oplus \mathscr{V}_m^{\perp_V} = \mathscr{H}$. But since, for every $\boldsymbol{u} \in \mathscr{H}$, $\boldsymbol{u} \perp_V \mathscr{V}_m$ if and only if $\boldsymbol{u} \perp \mathscr{T}_m$, the oblique projection $P_{\mathscr{W}_m}^{\mathscr{T}_m}$ exists if H_m is nonsingular.

Recall that the QMR and QOR approximations are the MR and OR approximations, respectively, obtained when replacing the original inner product $(\cdot,\cdot)$ by the basis-dependent inner product $(\cdot,\cdot)_V$. This simple observation implies that the assertions of the preceding sections, particularly those of Theorem 2.2.5 and Propositions 2.2.9, 2.3.8, are valid for any pair of QMR/QOR methods. Note, however, that when formulating these results for QMR/QOR methods, each occurrence of the original norm must be replaced by the $\|\cdot\|_V$-norm and that angles are understood to be defined with respect to $(\cdot,\cdot)_V$.

As an example, we mention that

$$\boldsymbol{w}_m^{\mathrm{QMR}} = \widehat{s}_m^2 \boldsymbol{w}_{m-1}^{\mathrm{QMR}} + \widehat{c}_m^2 \boldsymbol{w}_m^{\mathrm{QOR}},$$

where

$$\widehat{s}_m := \sin \measuredangle_V(\boldsymbol{d}_{m-1}^{\mathrm{QMR}}, \mathscr{W}_m), \quad \widehat{c}_m := \cos \measuredangle_V(\boldsymbol{d}_{m-1}^{\mathrm{QMR}}, \mathscr{W}_m)$$

and $\measuredangle_V(\boldsymbol{d}_{m-1}^{\mathrm{QMR}}, \mathscr{W}_m)$ denotes the angle between $\boldsymbol{d}_{m-1}^{\mathrm{QMR}}$ and $\mathscr{W}_m$ with respect to $(\cdot,\cdot)_V$.

The preceding analysis might lead one to believe that the QMR approximations will move steadily farther away from the MR approximation at each step. The following observation due to Stewart (1998) shows that this is not necessarily the case, but that the QMR approximation may under certain conditions recover, regardless of how far it may have deviated from the (optimal) MR approximation in earlier steps.

Proposition 2.3.12. *There holds $\boldsymbol{w}_m^{QMR} = \boldsymbol{w}_m^{MR}$ if and only if $\widetilde{\boldsymbol{v}}_{m+1} \perp \mathscr{W}_m$, and there holds $\boldsymbol{w}_m^{QOR} = \boldsymbol{w}_m^{OR}$ if and only if $\boldsymbol{v}_{m+1} \perp \mathscr{V}_m$.*

Proof. By Proposition 2.3.6, we have $\mathscr{W}_m = \mathrm{span}\{\widehat{\boldsymbol{v}}_1, \ldots, \widehat{\boldsymbol{v}}_m\}$ and $\boldsymbol{d}_m^{\mathrm{QMR}} \in \mathrm{span}\{\widetilde{\boldsymbol{v}}_{m+1}\}$. Since $\boldsymbol{d}_m^{\mathrm{QMR}} = \boldsymbol{d}_m^{\mathrm{MR}}$ if and only if $\boldsymbol{d}_m^{\mathrm{QMR}} \perp \mathscr{W}_m$, the first assertion is proven. The analogous assertion for the QOR approximation follows from $\boldsymbol{d}_m^{\mathrm{QOR}} \in \mathrm{span}\{\boldsymbol{v}_{m+1}\}$, hence $\boldsymbol{w}^{\mathrm{QOR}} \perp \mathscr{V}_m$ if and only if $\boldsymbol{v}_{m+1} \perp \mathscr{V}_m$. □

Smoothing Procedures Revisited

Next, we comment on the effect of applying the smoothing procedures introduced in Section 3, namely minimal and quasi-minimal residual smoothing, to the QOR approximations. We first note that they are no longer equivalent: More precisely, if we define

$$\alpha_m^{\mathrm{MR}} := \frac{(\boldsymbol{d}_{m-1}^{\mathrm{QMR}}, \boldsymbol{d}_{m-1}^{\mathrm{QMR}} - \boldsymbol{d}_m^{\mathrm{QOR}})}{\|\boldsymbol{d}_{m-1}^{\mathrm{QMR}} - \boldsymbol{d}_m^{\mathrm{QOR}}\|^2} \tag{2.72}$$

then, in general, $\boldsymbol{w}_m^{\mathrm{QMR}} \neq (1-\alpha_m^{\mathrm{MR}})\boldsymbol{w}_{m-1}^{\mathrm{QMR}} + \alpha_m^{\mathrm{MR}} \boldsymbol{w}_m^{\mathrm{QOR}}$ because the formula (2.72) for the smoothing parameter α_m^{MR} was derived in order that the errors of the smoothed approximations solve a local approximation problem with respect to the inner product $(\cdot,\cdot)$, which is different from the inner product $(\cdot,\cdot)_V$ which characterizes the QMR and QOR approximations. Minimal residual smoothing therefore does not lead from QOR to QMR.

It is an easy consequence of Remark 2.3.3 that the situation is different if we apply QMR smoothing: If we set

$$\alpha_m^{\mathrm{QMR}} = \frac{1/\|\boldsymbol{d}_m^{\mathrm{QOR}}\|_V^2}{\sum_{j=0}^m 1/\|\boldsymbol{d}_j^{\mathrm{QOR}}\|_V^2},$$

then $\boldsymbol{w}_m^{\mathrm{QMR}} = (1-\alpha_m^{\mathrm{QMR}})\boldsymbol{w}_{m-1}^{\mathrm{QMR}} + \alpha_m^{\mathrm{QMR}}\boldsymbol{w}_m^{\mathrm{QOR}}$ does indeed hold (cf. Proposition 2.2.9). But since $\boldsymbol{d}_m^{\mathrm{QOR}} = \gamma\boldsymbol{v}_{m+1}$ for some $\gamma \in \mathbb{C}$, we have $\|\boldsymbol{d}_m^{\mathrm{QOR}}\|_V = |\gamma| = \|\boldsymbol{d}_m^{\mathrm{QOR}}\|/\|\boldsymbol{v}_{m+1}\|$, and consequently

$$\alpha_m^{\mathrm{QMR}} = \frac{\|\boldsymbol{v}_{m+1}\|^2/\|\boldsymbol{d}_m^{\mathrm{QOR}}\|^2}{\sum_{j=0}^{m}\|\boldsymbol{v}_{j+1}\|^2/\|\boldsymbol{d}_j^{\mathrm{QOR}}\|^2}.$$

If we again make the common assumption that the basis vectors $\boldsymbol{v}_j$ have unit length ($\|\boldsymbol{v}_j\| = 1$ for all j), then

$$\alpha_m^{\mathrm{QMR}} = \frac{1/\|\boldsymbol{d}_m^{\mathrm{QOR}}\|^2}{\sum_{j=0}^{m}1/\|\boldsymbol{d}_j^{\mathrm{QOR}}\|^2},$$

and the QMR and QOR approximants are related by exactly the formulas which hold for a proper MR/OR pair, namely:

$$\boldsymbol{w}_m^{\mathrm{QMR}} = \frac{\sum_{j=0}^{m}\boldsymbol{w}_j^{\mathrm{QOR}}/\|\boldsymbol{d}_j^{\mathrm{QOR}}\|^2}{\sum_{j=0}^{m}1/\|\boldsymbol{d}_j^{\mathrm{QOR}}\|^2} \quad \text{and} \quad \boldsymbol{d}_m^{\mathrm{QMR}} = \frac{\sum_{j=0}^{m}\boldsymbol{d}_j^{\mathrm{QOR}}/\|\boldsymbol{d}_j^{\mathrm{QOR}}\|^2}{\sum_{j=0}^{m}1/\|\boldsymbol{d}_j^{\mathrm{QOR}}\|^2}.$$

QMR smoothing applied to CGS and Bi-CGSTAB is discussed in Walker (1995). For smoothing techniques applied to the general class of Lanczos-type product methods, see Ressel & Gutknecht (1996).

2.4 Every Method is an MR and an OR Method

In Section 2.3.3 we saw how the QMR and QOR approximations can be reinterpreted as MR and OR approximations with respect to the basis-dependent inner product $(\cdot,\cdot)_V$. It turns out that an analogous interpretation is possible for *any* reasonable sequence of approximations $\{\boldsymbol{w}_m\}$ to a given $\boldsymbol{r} \in \mathscr{H}$, namely both as an MR *and* as an OR approximation.

In this section $\{\mathscr{W}_m\}_{m=0}^{L}$ denotes any sequence of nested spaces with $\dim\mathscr{W}_m = m$ (in particular, $\mathscr{W}_0 = \{\boldsymbol{0}\}$), and $\{\mathscr{V}_m\}_{m=1}^{L}$ denotes the associated sequence of error spaces $\mathscr{V}_m = \operatorname{span}\{\boldsymbol{r}\} + \mathscr{W}_{m-1}$ with respect to $\boldsymbol{r} \in \mathscr{H}$.

Theorem 2.4.1. *Assume $\{\boldsymbol{h}_m\}_{m=0}^{L}$ is a sequence of approximations to $\boldsymbol{r} \in \mathscr{H}$ such that $\boldsymbol{h}_m \in \mathscr{W}_m$ and $\boldsymbol{h}_L = \boldsymbol{r}$. Then there exists an inner product $(\cdot,\cdot)_V$ on $\mathscr{V}_L = \mathscr{W}_L$ such that*

$$\|\boldsymbol{r}-\boldsymbol{h}_m\|_V = \min_{\boldsymbol{w}\in\mathscr{W}_m}\|\boldsymbol{r}-\boldsymbol{w}\|_V, \qquad m = 1,2,\ldots,L-1,$$

if and only if $\boldsymbol{h}_m \in \mathscr{W}_{m-1}$ implies $\boldsymbol{h}_m = \boldsymbol{h}_{m-1}$ for $m = 1,2,\ldots,L-1$ or, in other words, if and only if

$$\textit{either} \quad \boldsymbol{h}_m \in \mathscr{W}_m \setminus \mathscr{W}_{m-1} \textit{ or } \boldsymbol{h}_m = \boldsymbol{h}_{m-1}, \qquad m = 1,2,\ldots,L-1. \tag{2.73}$$

Proof. If the vectors $\{\boldsymbol{h}_m\}_{m=1}^{L-1}$ are the best approximations to $\boldsymbol{r}$ from $\mathscr{W}_m$ with respect to some inner product $(\cdot,\cdot)_V$, then whenever $\boldsymbol{h}_m$ happens to lie also in $\mathscr{W}_{m-1}$, $\boldsymbol{h}_m$ must

also be the best approximation to $\boldsymbol{r}$ from $\mathscr{W}_{m-1}$, whereby $\boldsymbol{h}_m = \boldsymbol{h}_{m-1}$, which proves the necessity of (2.73).

Conversely, assuming that (2.73) is satisfied, we write

$$\boldsymbol{r} = (\boldsymbol{h}_1 - \boldsymbol{h}_0) + (\boldsymbol{h}_2 - \boldsymbol{h}_1) + \cdots + (\boldsymbol{h}_L - \boldsymbol{h}_{L-1})$$

and construct a basis $\{\boldsymbol{w}_1, \ldots \boldsymbol{w}_L\}$ of $\mathscr{W}_L$ by setting

$$\boldsymbol{w}_m := \begin{cases} \boldsymbol{h}_m - \boldsymbol{h}_{m-1}, & \text{if } \boldsymbol{h}_m \in \mathscr{W}_m \setminus \mathscr{W}_{m-1}, \\ \text{an arbitrary vector from } \mathscr{W}_m \setminus \mathscr{W}_{m-1}, & \text{if } \boldsymbol{h}_m = \boldsymbol{h}_{m-1}. \end{cases}$$

Note that, for each m, $\{\boldsymbol{w}_1, \ldots \boldsymbol{w}_m\}$ is a basis of $\mathscr{W}_m$. We further define the "Fourier coefficients" α_m by

$$\alpha_m := \begin{cases} 1, & \text{if } \boldsymbol{h}_m \in \mathscr{W}_m \setminus \mathscr{W}_{m-1}, \\ 0, & \text{if } \boldsymbol{h}_m = \boldsymbol{h}_{m-1}, \end{cases} \qquad m = 1, \ldots, L,$$

so that $\boldsymbol{r} = \alpha_1 \boldsymbol{w}_1 + \alpha_2 \boldsymbol{w}_2 + \cdots \alpha_L \boldsymbol{w}_L$ and

$$\boldsymbol{h}_m = \alpha_1 \boldsymbol{w}_1 + \alpha_2 \boldsymbol{w}_2 + \cdots + \alpha_m \boldsymbol{w}_m, \qquad m = 1, \ldots, L,$$

i.e., $\boldsymbol{h}_m$ is nothing but the truncated "Fourier expansion" of $\boldsymbol{r}$. Defining the inner product $(\cdot, \cdot)_V$ such that $\{\boldsymbol{w}_1, \ldots \boldsymbol{w}_L\}$ are orthonormal then leads to the desired conclusion. □

The next theorem establishes the analogous result for the OR (or, more precisely, the QOR) approximation.

Theorem 2.4.2. *If $\{\boldsymbol{h}_m\}_{m=1}^L$ is a sequence of approximations to $\boldsymbol{r} \in \mathscr{H}$ such that $\boldsymbol{h}_m \in \mathscr{W}_m$ and $\boldsymbol{h}_L = \boldsymbol{r}$, then an inner product $(\cdot, \cdot)_{\widetilde{V}}$ on $\mathscr{V}_L = \mathscr{W}_L$ such that*

$$\boldsymbol{r} - \boldsymbol{h}_m \perp_{\widetilde{V}} \mathscr{V}_m, \qquad m = 1, 2, \ldots, L-1,$$

exists if and only if $\boldsymbol{h}_m \in \mathscr{W}_m \setminus \mathscr{W}_{m-1}$ for $m = 1, 2, \ldots, L-1$.

Proof. Assume that, for all $m = 1, 2, \ldots, L-1$, the vectors $\boldsymbol{h}_m$ are the OR approximations to $\boldsymbol{r}$ from $\mathscr{W}_m$ with respect to some inner product $(\cdot, \cdot)_{\widetilde{V}}$. If now, for some m, $\boldsymbol{h}_m \in \mathscr{W}_{m-1}$, then $\boldsymbol{d}_m := \boldsymbol{r} - \boldsymbol{h}_m \in \operatorname{span}\{\boldsymbol{r}\} + \mathscr{W}_{m-1} = \mathscr{V}_m$, i.e., $\boldsymbol{d}_m \in \mathscr{V}_m \cap \mathscr{V}_m^{\perp} = \{\boldsymbol{0}\}$, which implies $\boldsymbol{h}_m = \boldsymbol{r}$. But this is impossible unless $\boldsymbol{r} \in \mathscr{W}_m$, i.e., $m = L$, and we have thus established that $\boldsymbol{h}_m \in \mathscr{W}_m \setminus \mathscr{W}_{m-1}$ for $m = 1, \ldots, L-1$.

Conversely, since $\boldsymbol{h}_m \in \mathscr{W}_m \setminus \mathscr{W}_{m-1}$ implies $\boldsymbol{d}_m \in \mathscr{V}_{m+1} \setminus \mathscr{V}_m$ for $m = 1, \ldots, L-1$, we see that $\{\boldsymbol{d}_0, \boldsymbol{d}_1, \ldots \boldsymbol{d}_{L-1}\}$ $(\boldsymbol{d}_0 = \boldsymbol{r})$ is a basis of $\mathscr{V}_L$ such that, for every $m = 1, \ldots, L-1$, $\{\boldsymbol{d}_0, \boldsymbol{d}_1, \ldots \boldsymbol{d}_{m-1}\}$ is a basis of $\mathscr{V}_m$. Defining the inner product $(\cdot, \cdot)_{\widetilde{V}}$ such that $\{\boldsymbol{d}_0, \ldots \boldsymbol{d}_{L-1}\}$ is an orthogonal basis of $\mathscr{V}_L$ leads to the desired conclusion. □

We have formulated these two theorems for the case of a sequence terminating with $\boldsymbol{h}_L = \boldsymbol{r}$, as this is the situation when solving linear equations in finite dimensions by MR and OR approximation based on Krylov spaces, at least in the absence of rounding errors. When the sequence of approximations does not terminate, we may proceed analogously as in Theorem 2.4.1 with the difference that the inner product is then defined on the union

of all error spaces $\mathscr{V}_m$ and we need not have convergence of the approximations to $\boldsymbol{r}$ in the norm induced by this inner product. Similar considerations apply for a formulation of Theorem 2.4.2 for a nonterminating sequence of approximations.

We conclude that, by allowing the inner product to vary, the concept of MR and OR approximations becomes sufficiently general to include any reasonable sequence of approximations.

Of course, these results are of a rather academic nature since an application in which the linear system arises often comes with a natural norm to be minimized in the iteration. However, these results show that the MR/OR framework really does include all reasonable approximation schemes and that methods should not be compared on the grounds of whether they minimize a norm, but whether this norm is appropriate for the problem.

Chapter 3

Solving Equations with Projection Methods

In this chapter we apply the results of Chapter 2 to the solution of linear equations (1.1). The three basic types or orthonormal bases used in the MR/OR approximations lead to equation-solving methods for general correction spaces. Sections 3.3 and 3.4 discuss MR approximation with respect to the direct sum of two correction spaces, which orthogonality relations this involves, and how these orthogonality relations may be relaxed in an optimal manner.

3.1 Basic Setting

We now return to our original problem of approximating the solution of the operator equation (1.1). The methods we shall consider for this purpose belong to the class of *subspace correction methods*, which, starting from an initial approximation $\boldsymbol{x}_0$ of the solution of (1.1), determine corrections $\boldsymbol{c}$ to $\boldsymbol{x}_0$ which are chosen from a sequence of nested *correction spaces*

$$\{\boldsymbol{0}\} = \mathscr{C}_0 \subset \mathscr{C}_1 \subset \mathscr{C}_2 \subset \cdots \subset \mathscr{C}_m \subset \mathscr{C}_{m+1} \subset \cdots \subset \mathscr{H},$$

leading to approximate solutions of the form $\boldsymbol{x}_m = \boldsymbol{x}_0 + \boldsymbol{c}_m$, $\boldsymbol{c}_m \in \mathscr{C}_m$. Methods of this kind abound in applied mathematics; as a familiar example we mention the class of *additive multilevel methods* for approximating the solution of an elliptic boundary value problem, in which the correction spaces $\mathscr{C}_m$ are nested finite element subspaces of a function space $\mathscr{H}$ known to contain the solution (see, e.g., Xu (1992)). Unless stated otherwise, we shall only consider correction spaces for which $\dim \mathscr{C}_m = m$.

In the methods to be considered, the determination of the correction is based on conditions imposed on the *residual* $\boldsymbol{r}_m = \boldsymbol{b} - A\boldsymbol{x}_m$ of an approximation $\boldsymbol{x}_m$ to the solution of (1.1). In particular, and this is the natural connection to the abstract formulation in Chapter 2, these methods can be viewed as methods which approximate the *initial residual* $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$ from the sequence of subspaces

$$\mathscr{W}_m = A\mathscr{C}_m, \qquad m = 0, 1, \ldots, \tag{3.1}$$

i.e., from the images of the correction spaces under the operator A.

Since A is assumed invertible, the identification (3.1) in case of the MR approximation leads to

$$\|\boldsymbol{r}_0 - \boldsymbol{w}_m^{\mathrm{MR}}\| = \min_{\boldsymbol{w}\in\mathscr{W}_m = A\mathscr{C}_m} \|\boldsymbol{r}_0 - \boldsymbol{w}\| = \min_{\boldsymbol{c}\in\mathscr{C}_m} \|\boldsymbol{r}_0 - A\boldsymbol{c}\| = \min_{\boldsymbol{c}\in\mathscr{C}_m} \|\boldsymbol{b} - A(\boldsymbol{x}_0 + \boldsymbol{c})\|,$$

so that the MR approximation selects that correction in $\mathscr{C}_m$ which minimizes the residual. Analogously, for the OR approximation, with $\boldsymbol{c}_m^{\mathrm{OR}}$ defined by $A\boldsymbol{c}_m^{\mathrm{OR}} = \boldsymbol{w}_m^{\mathrm{OR}}$,

$$\mathscr{V}_m \perp \boldsymbol{r}_0 - \boldsymbol{w}_m^{\mathrm{OR}} = \boldsymbol{r}_0 - A\boldsymbol{c}_m^{\mathrm{OR}} = \boldsymbol{b} - A(\boldsymbol{x}_0 + \boldsymbol{c}_m^{\mathrm{OR}}),$$

where now, in view of (3.1),

$$\mathscr{V}_m = \mathrm{span}\{\boldsymbol{r}_0\} + A\mathscr{C}_m. \tag{3.2}$$

The OR approximation thus selects corrections whose residuals are orthogonal to the space $\mathscr{V}_m$. The identification $\mathscr{W}_m = A\mathscr{C}_m$ also identifies approximation errors for $\boldsymbol{r}_0$ with the corresponding residuals with respect to (1.1). For this reason, it is more appropriate in this equation-solving context to refer to the error spaces $\mathscr{V}_m$ as *residual spaces*.

Of course, upon setting (3.1), all results of Chapter 2 carry over to the equation-solving context. Reformulating each result would be rather tedious, so we restrict ourselves to a few crucial issues. First, the termination index L defined in (2.40) can now be characterized as

$$\begin{aligned} L &= \min\{m : \boldsymbol{r}_0 \in A\mathscr{C}_m\} = \min\{m : \boldsymbol{b} = A(\boldsymbol{x}_0 + \boldsymbol{c}), \boldsymbol{c} \in \mathscr{C}_m\} \\ &= \min\{m : A^{-1}\boldsymbol{b} \in \boldsymbol{x}_0 + \mathscr{C}_m\}. \end{aligned} \tag{3.3}$$

Second, the relations between MR and OR errors given in Theorem 2.2.5, Theorem 2.2.6 and Corollary 2.2.7 translate to relations between MR and OR residuals by identifying $\boldsymbol{d}_m^{\mathrm{MR}} = \boldsymbol{r}_m^{\mathrm{MR}}$ and $\boldsymbol{d}_m^{\mathrm{OR}} = \boldsymbol{r}_m^{\mathrm{OR}}$. The analogous expression to (2.30) follows for the approximations $\boldsymbol{x}_m^{\mathrm{MR}}, \boldsymbol{x}_m^{\mathrm{OR}} \in \mathscr{C}_m$ as a result of $\boldsymbol{r}_m = \boldsymbol{b} - A\boldsymbol{x}_m$ and since $s_m^2 + c_m^2 = 1$:

$$\boldsymbol{x}_m^{\mathrm{MR}} = s_m^2\boldsymbol{x}_{m-1}^{\mathrm{MR}} + c_m^2\boldsymbol{x}_m^{\mathrm{OR}}. \tag{3.4}$$

In the same manner, the MR and QMR smoothing transformations given in Section 2.2.4 also carry over from residuals to approximations. For later use, we note that setting $\mathscr{W} = A\mathscr{C}_m$ in (2.5) yields

$$\|\boldsymbol{r}_m^{\mathrm{MR}}\| = \sin\measuredangle(\boldsymbol{r}_0, A\mathscr{C}_m)\|\boldsymbol{r}_0\|. \tag{3.5}$$

The angle-free expressions of (3.4) from Corollary 2.2.7 become

$$\begin{aligned} \|\boldsymbol{r}_m^{\mathrm{MR}}\| &= \sqrt{1 - \frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2}{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2}}\,\|\boldsymbol{r}_m^{\mathrm{OR}}\|, \\ \boldsymbol{x}_m^{\mathrm{OR}} &= \frac{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2}{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2 - \|\boldsymbol{r}_m^{\mathrm{MR}}\|^2}\boldsymbol{x}_m^{\mathrm{MR}} - \frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2}{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2 - \|\boldsymbol{r}_m^{\mathrm{MR}}\|^2}\boldsymbol{x}_{m-1}^{\mathrm{MR}}. \end{aligned}$$

As can be seen from the last two equations, the OR approximation and residual can easily be computed from the corresponding MR quantities. Moreover, since the MR quantities can always be computed in a stable fashion, this is the preferable way to

compute these quantities. (An exception is, of course, the Hermitian positive definite case, in which the OR quantities may be computed stably and at lower expense than the MR counterparts by the classical method of conjugate gradients, cf. Section 4.4.)

Finally, we reformulate the characterizations of a Galerkin breakdown as discussed for the abstract setting in Remarks 2.1.5 and 2.2.8.

Proposition 3.1.1. *For every $m = 1, 2, \ldots, L-1$, the following four statements are equivalent:*

(a) The OR approximation $\boldsymbol{x}_m^{OR}$ (i.e., $P_{\mathscr{W}_m}^{\mathscr{V}_m}$) does not exist.
(b) The MR approximation makes no progress, i.e., $\boldsymbol{r}_m^{MR} = \boldsymbol{r}_{m-1}^{MR}$.
(c) The Hessenberg matrix H_m of (2.39) is singular.
(d) $\varphi_m = \measuredangle(\boldsymbol{r}_{m-1}^{MR}, A\mathscr{C}_m) = \pi/2$ (equivalently, $\sin\varphi_m = 1$ or $\cos\varphi_m = 0$).

3.2 Algorithms for General Correction Spaces

In this section we formulate the basic coordinate representations presented in Section 2.3 as algorithms for solving (1.1) and thereby arrive at the most important algorithms based on projection methods. Since the remainder of this thesis is concerned primarily with MR methods, we shall restrict ourselves to MR algorithms and only indicate whenever an algorithm can also be used to compute OR approximations with minor modifications.

The methods we discuss below were originally presented in the literature in the guise of Krylov subspace methods, which are MR and OR methods in which the correction space $\mathscr{C}_m$ is the mth Krylov space with respect to A and $\boldsymbol{r}_0$, and we discuss this most investigated case in Chapter 4. We choose to first reformulate these algorithms for general correction spaces not only because this is the canonical formulation of the abstract MR and OR methods of Chapter 2 for solving operator equations, entailing no added complication versus the Krylov subspace case, but also because in recent work attention has shifted to correction spaces other than Krylov spaces, and it is therefore of practical relevance to investigate this situation as well.

As the MR approximation computes the best approximation of $\boldsymbol{r}_0$ from $\mathscr{W}_m = A\mathscr{C}_m$, its implementation requires the (explicit or implicit) construction of orthonormal bases of the approximation spaces $\mathscr{W}_m$. The most direct approach is, given a basis $C_m = [\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_m]$ of $\mathscr{C}_m$, to construct an orthonormal basis of $\mathscr{W}_m$ by orthonormalizing the images $A\boldsymbol{c}_j$ and then proceed as in Section 2.3.1. Because the spaces are nested, the orthogonalization results in a QR decomposition

$$AC_m = W_m R_m \tag{3.6}$$

with a triangular matrix $R_m \in \mathbb{C}^{m\times m}$ and a set of orthonormal vectors $W_m = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m]$ which form a basis of $\mathscr{W}_m$. The MR approximation $\boldsymbol{x}_m^{\mathrm{MR}}$ of the solution of (1.1) with respect to $\mathscr{C}_m$ is given by $\boldsymbol{x}_m^{\mathrm{MR}} = \boldsymbol{x}_0 + C_m \boldsymbol{y}_m^{\mathrm{MR}}$ with a coefficient vector $\boldsymbol{y}_m^{\mathrm{MR}} \in \mathbb{C}^m$. By (2.37), the coefficient vector of $P_{\mathscr{W}_m}\boldsymbol{r}_0$ with respect to $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ is $W_m^*\boldsymbol{r}_0$, hence we must have $AC_m\boldsymbol{y}_m^{\mathrm{MR}} = W_m W_m^*\boldsymbol{r}_0$, which, in view of (3.6), leads to $\boldsymbol{y}_m^{\mathrm{MR}} = R_m^{-1}W_m^*\boldsymbol{r}_0$. If no Galerkin breakdown occurs at this step, i.e., if $(\boldsymbol{r}_0, \boldsymbol{w}_m) \neq 0$, then the OR approximation

$\boldsymbol{x}_m^{\mathrm{OR}} = \boldsymbol{x}_0 + C_m \boldsymbol{y}_m^{\mathrm{OR}}$ may be computed following (2.38) by solving

$$R_m \boldsymbol{y}_m^{\mathrm{OR}} = \begin{bmatrix} (\boldsymbol{r}_0, \boldsymbol{w}_1) \\ \vdots \\ (\boldsymbol{r}_0, \boldsymbol{w}_{m-1}) \\ \|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2/(\boldsymbol{w}_m, \boldsymbol{r}_0) \end{bmatrix}, \tag{3.7}$$

or by using (3.4) and noting $c_m = |(\boldsymbol{w}_m, \boldsymbol{r}_0)|/\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|$, $s_m = \sqrt{1 - c_m^2}$.

The main computational expense of this algorithm lies in the orthogonalization process and the solution of a triangular system whenever the approximations are desired. Although it appears to be the most straightforward implementation, this method was only recently proposed for computing MR approximations by Walker & Zhou (1994). We summarize m steps of the Walker-Zhou algorithm in Algorithm 3.2.1. Of course, the

Algorithm 3.2.1 The algorithm of Walker and Zhou for general correction spaces.

$\boldsymbol{r}_0 := \boldsymbol{b} - A\boldsymbol{x}_0$
for $j := 1$ **to** m
  Select $\boldsymbol{c}_j \in \mathscr{C}_j \setminus \mathscr{C}_{j-1}$, $\widehat{\boldsymbol{w}} := A\boldsymbol{c}_j$
  for $i = 1$ **to** $j - 1$
    $r_{i,j} := (\widehat{\boldsymbol{w}}, \boldsymbol{w}_i)$
    $\widehat{\boldsymbol{w}} := \widehat{\boldsymbol{w}} - r_{i,j}\boldsymbol{w}_i$
  end
  $r_{j,j} := \|\widehat{\boldsymbol{w}}\|$
  $\boldsymbol{w}_j := \widehat{\boldsymbol{w}}/r_{j,j}$
end
$\boldsymbol{y}_m^{\mathrm{MR}} := R_m^{-1} W_m^* \boldsymbol{r}_0$
$\boldsymbol{x}_m^{\mathrm{MR}} := \boldsymbol{x}_0 + C_m \boldsymbol{y}_m^{\mathrm{MR}}$
$\boldsymbol{r}_m^{\mathrm{MR}} = W_m W_m^* \boldsymbol{r}_0$
If desired, compute $\boldsymbol{y}_m^{\mathrm{OR}}$, $\boldsymbol{x}_m^{\mathrm{OR}}$ and $\boldsymbol{r}_m^{\mathrm{OR}}$ using (3.7) or (3.4)

residual and its norm may also be updated as

$$\boldsymbol{r}_m = \boldsymbol{r}_{m-1} - (\boldsymbol{r}_0, \boldsymbol{w}_m)\boldsymbol{w}_m, \quad \|\boldsymbol{r}_m\|^2 = \|\boldsymbol{r}_{m-1}\|^2 - |(\boldsymbol{r}_0, \boldsymbol{w}_m)|^2, \qquad m = 1, 2, \ldots,$$

in which $(\boldsymbol{r}_0, \boldsymbol{w}_m)$ may be replaced by $(\boldsymbol{r}_{m-1}, \boldsymbol{w}_m)$ by orthogonality.

Remark 3.2.1. This is the first occasion where we have given an algorithm for an orthogonalization process (as opposed to the abstract formulation (2.41)), in this case orthogonalizing the image $A\boldsymbol{c}_j$ against $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{j-1}$. Different mathematically equivalent formulations of the orthogonalization process, such as the Gram-Schmidt, modified Gram-Schmidt or Householder orthogonalization algorithms, each possibly combined with various reorthogonalization schemes (see e.g. Golub & van Loan (1996)) can differ markedly in the presence of roundoff error. Investigations into the effect of the orthogonalization on the finite-precision behavior of MR methods have been carried out by Greenbaum and Strakos and their co-workers (see Rozložník & Strakoš (1996) and Greenbaum, Rozložník & Strakoš (1997)). Their results show that the modified Gram-Schmidt process seems to

be the best compromise between efficiency and accuracy, in the sense that loss of orthogonality due to roundoff does not affect convergence until the residual norm has reached its final accuracy. For this reason, we will formulate orthogonalizations in the modified Gram-Schmidt form when formulating algorithms.

Since the image of the MR correction $\boldsymbol{c}_m^{\mathrm{MR}}$ under A is the best approximation $W_m W_m^* \boldsymbol{r}_0$ of $\boldsymbol{r}_0$ from $\mathscr{W}_m$, we conclude that

$$\boldsymbol{c}_m^{\mathrm{MR}} = A^{-1} W_m W_m^* \boldsymbol{r}_0 = C_m W_m^* \boldsymbol{r}_0$$

if the basis C_m consists of the pre-images under A of the orthonormal basis vectors W_m of $\mathscr{W}_m$. Such a pair of bases is generated by an algorithm known as the *generalized conjugate residual (GCR) method*, which was introduced as a Krylov subspace method by Eisenstat, Elman & Schultz (1983). In this case the coefficient vector $\boldsymbol{y}_m^{\mathrm{MR}}$ of the MR correction with respect to C_m consists simply of the Fourier coefficients $W_m^* \boldsymbol{r}_0$, i.e., no triangular system needs to be solved. The corresponding OR coefficient vector is obtained, again in view of (2.38), by

$$\boldsymbol{y}_m^{\mathrm{OR}} = [(\boldsymbol{r}_0, \boldsymbol{w}_1), \ldots, (\boldsymbol{r}_0, \boldsymbol{w}_{m-1}), \|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2 / (\boldsymbol{w}_m, \boldsymbol{r}_0)]^\top. \tag{3.8}$$

The associated residual vectors may be formed by using the same coefficient vectors, but with respect to the orthonormal basis W_m. The GCR algorithm, which is listed as Algorithm 3.2.2, belongs to the vast lineage of generalizations of the *conjugate gradient* and *conjugate residual methods* of Hestenes & Stiefel (1952), and we refer to Section 4.4 for a brief survey of this family of Krylov subspace methods.

Algorithm 3.2.2 GCR for general correction spaces.

```
1  r_0 := b − A x_0
2  for j := 1 to m
3      Select ĉ ∈ C_j \ C_{j−1}, ŵ := A ĉ
4      for i := 1 to j − 1
5          ĉ := ĉ − (ŵ, w_i) c_i
6          ŵ := ŵ − (ŵ, w_i) w_i
7      end
8      w_j := ŵ/‖ŵ‖
9      c_j := ĉ/‖ŵ‖
10 end
11 y_m^MR = W_m^* r_0
12 x_m^MR := x_0 + C_m y_m^MR = x_{m−1}^MR + (r_0, w_m) c_m
13 r_m^MR := r_0 − W_m y_m^MR = r_{m−1}^MR − (r_0, w_m) w_m
14 If desired, compute y_m^OR, x_m^OR and r_m^OR using (3.8)
```

A third fundamental algorithmic approach, which has become the most popular for computing MR approximations, relies on an orthonormal basis of the residual space, and thus allows us to apply the results of Section 2.3.2. Its Krylov subspace counterparts were introduced by Paige & Saunders (1975) for selfadjoint A and by Saad & Schultz (1986) for the general, nonselfadjoint case. The latter, general form is known as the *generalized*

minimum residual method (GMRES). Subsequently, Saad (1993) introduced a variant of GMRES for general correction spaces in order to accommodate variable preconditioners, and introduced the name *flexible GMRES (FGMRES)* for the resulting algorithm, which is listed as Algorithm 3.2.3. A typical application of FGMRES is the use of GMRES in conjunction with a preconditioner which can vary from step to step. Such an example can be found in Elman, Ernst & O'Leary (1999).

The point of departure in FGMRES is the (in general not orthogonal) basis C_m of $\mathscr{C}_m$, the image $W_m = AC_m$ of which is a basis of $\mathscr{W}_m$. FGMRES then proceeds to compute an orthonormal basis V_{m+1} of the residual space $\mathscr{V}_{m+1} = \text{span}\{\boldsymbol{r}_0\} + \mathscr{W}_m$, resulting in the decomposition

$$AC_m = V_{m+1}\widetilde{H}_m, \tag{3.9}$$

which is relation (2.39) with $W_m = AC_m$ as the basis of the approximation space $\mathscr{W}_m$. Due to its similarity with an *Arnoldi decomposition* (Arnoldi 1951), which results when $C_m = V_m$, we shall refer to (3.9) as an *Arnoldi-type decomposition*. The orthogonalization process is the same as (2.41), and, in view of (2.42) and (2.43), the entries $\eta_{j,m}$ of $\widetilde{H}_m$ are

$$\eta_{j,m} = (A\boldsymbol{c}_m, \boldsymbol{v}_j), \quad j = 1, \ldots, m+1, \quad m \geq 1, \tag{3.10}$$

where $\eta_{m+1,m} \geq 0$. With the QR-decomposition (2.49) and $\widetilde{Q}_m$ defined in (2.50), we also have (cf. (2.61) and (2.63))

$$AC_m = V_{m+1}\widetilde{Q}_m R_m = \widehat{V}_m R_m. \tag{3.11}$$

Since the coefficient vectors of the MR and OR corrections with respect to C_m are those of the associated residual approximations with respect to the basis $W_m = AC_m$, their computation proceeds exactly as in Section 2.3.2. Therefore, we obtain $\boldsymbol{x}_m^{\text{MR}} = \boldsymbol{x}_0 + C_m\boldsymbol{y}_m^{\text{MR}}$ and $\boldsymbol{x}_m^{\text{OR}} = \boldsymbol{x}_0 + C_m\boldsymbol{y}_m^{\text{OR}}$, where $\boldsymbol{y}_m^{\text{MR}}$ and $\boldsymbol{y}_m^{\text{OR}}$ solve (2.44) and (2.45), respectively.

To relate FGMRES to the Walker-Zhou variant of MR, note that $\widehat{V}_m R_m$ is a QR-factorization of the form (3.6). Similarly, the orthonormal basis $\widehat{V}_m$ of $\mathscr{W}_m$ plays the role of W_m in GCR, and the second basis occurring in GCR—i.e., the pre-image of W_m under A—is given by $A^{-1}\widehat{V}_m = C_m R_m^{-1}$.

We thus conclude that GCR and the Walker-Zhou variant of MR both generate a subset of the quantities produced in FGMRES. As we will see in Chapter 4, when $\mathscr{C}_m$ is a sequence of Krylov spaces FGMRES possesses the additional advantage that only one basis needs to be generated and stored. Finally, we note that these three algorithms compute identical approximations only under the assumption of exact arithmetic, but will generally differ with regard to their stability in the presence of roundoff error. A thorough comparison along these lines is carried out by Rozložník & Strakoš (1996), where theoretical and numerical comparisons of variants of these three algorithms in the Krylov subspace case favor the GMRES algorithm. Aside from stability issues, the choice between these algorithms should be based on efficiency considerations, since, especially when combined with inner iterations, one or the other may be more convenient.

We conclude this section with an algorithmic formulation for the case treated in Section 2.3.3 in which none of the bases used is orthogonal with regard to the given inner product $(\cdot, \cdot)$ on $\mathscr{H}$: the generic QMR/QOR algorithm for solving operator equations, listed as Algorithm 3.2.4.

Algorithm 3.2.3 GMRES for a general correction space (FGMRES).

$\boldsymbol{r}_0 := \boldsymbol{b} - A\boldsymbol{x}_0, \beta := \|\boldsymbol{r}_0\|, \boldsymbol{v}_1 := \boldsymbol{r}_0/\beta$
for $j := 1$ **to** m
  Select $\boldsymbol{c}_j \in \mathscr{C}_j \setminus \mathscr{C}_{j-1}, \boldsymbol{w} := A\boldsymbol{c}_j$
  for $i = 1$ **to** j
    $h_{i,j} := (\boldsymbol{w}, \boldsymbol{v}_i)$
    $\boldsymbol{w} := \boldsymbol{w} - h_{i,j}\boldsymbol{v}_i$
  end
  $h_{j+1,j} := \|\boldsymbol{w}\|$
  $\boldsymbol{v}_{j+1} := \boldsymbol{w}/h_{j+1,j}$
end
Compute $\boldsymbol{y}_m^{\mathrm{MR}}$to minimize $\|\beta\boldsymbol{u}_1^{(m+1)} - \widetilde{H}_m\boldsymbol{y}\|$
Compute $\boldsymbol{y}_m^{\mathrm{OR}}$ to solve $H_m\boldsymbol{y} = \beta\boldsymbol{u}_1^{(m)}$
$\boldsymbol{x}_m^{\mathrm{MR}} := \boldsymbol{x}_0 + C_m\boldsymbol{y}_m^{\mathrm{MR}}$
$\boldsymbol{x}_m^{\mathrm{OR}} := \boldsymbol{x}_0 + C_m\boldsymbol{y}_m^{\mathrm{OR}}$

We observe that the only difference to the FGMRES algorithm is that the corresponding construction of the Arnoldi-type decomposition (3.9) no longer proceeds by successive orthogonalization of the vectors $\boldsymbol{r}_0, A\boldsymbol{c}_1, \ldots, A\boldsymbol{c}_m$ with respect to $(\cdot,\cdot)$, but that any set of vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m+1}$ and unreduced Hessenberg matrix $\widetilde{H}_m$ satisfying (3.9) is permitted. In particular, as observed in Section 2.3.3, the calculations in the coordinate space, i.e., the updated QR factorization of $\widetilde{H}_m$ by Givens rotations as well as the solution of the least-squares problem and Galerkin equation are the same as in the FGMRES algorithm. Of course, it should be mentioned that all the difficulty now lies in the generation of a suitable Arnoldi-type factorization (3.9), i.e., one for which the matrices $\widetilde{H}_m$ have a small bandwidth and for which the resulting basis-dependent inner product $(\cdot,\cdot)_V$ (cf. (2.68)) is close to the original inner product $(\cdot,\cdot)$ or some other inner product appropriate for the underlying operator equation. Currently the only general purpose method of this type is the original method to use this approach, the Krylov subspace QMR method of Freund & Nachtigal (1991), which employs the look-ahead Lanczos process (see Gutknecht (1992), Gutknecht (1994)) for this purpose.

Algorithm 3.2.4 Generic QMR algorithm for a general correction space.

$\boldsymbol{r}_0 := \boldsymbol{b} - A\boldsymbol{x}_0, \beta := \|\boldsymbol{r}_0\|, \boldsymbol{v}_1 := \boldsymbol{r}_0/\beta$
for $j := 1$ **to** m
  Select $\boldsymbol{c}_j \in \mathscr{C}_j \setminus \mathscr{C}_{j-1}, \boldsymbol{w} := A\boldsymbol{c}_j$
  Determine $\boldsymbol{v}_{j+1}, \{h_{i,j}\}_{i=1}^{j+1}$ such that $A\boldsymbol{c}_j = \sum_{i=1}^{j+1} h_{i,j}\boldsymbol{v}_i$
end
Compute $\boldsymbol{y}_m^{\mathrm{QMR}}$to minimize $\|\beta\boldsymbol{u}_1^{(m+1)} - \widetilde{H}_m\boldsymbol{y}\|_2$
Compute $\boldsymbol{y}_m^{\mathrm{QOR}}$ to solve $H_m\boldsymbol{y} = \beta\boldsymbol{u}_1^{(m)}$
$\boldsymbol{x}_m^{\mathrm{QMR}} := \boldsymbol{x}_0 + C_m\boldsymbol{y}_m^{\mathrm{QMR}}$
$\boldsymbol{x}_m^{\mathrm{QOR}} := \boldsymbol{x}_0 + C_m\boldsymbol{y}_m^{\mathrm{QOR}}$

3.3 Multiple Subspace Correction

Various recently developed enhancements of the basic MR and OR schemes presented above are based on introducing additional subspace corrections besides those associated with the stepwise increasing correction spaces. Existing approaches include generating such auxiliary projections from spectral information on the operator A gained during the iteration process or from additional inner iteration or restart cycles. In addition, time and storage constraints often make it necessary to form these projections only approximately, while at the same time keeping this approximation as effective as possible. In order to better describe and compare these new developments, we first formulate the basic projection steps required to combine *two* subspace corrections. Again, such a combination of two or more subspace corrections is a commonly used device in other areas of numerical analysis, e.g. in the field of additive Schwarz methods for solving boundary value problems, in which the different correction spaces correspond to finite-dimensional spaces of functions with support on subdomains of the original domain (which can also be interpreted as a block-Jacobi method). In the following Section 3.4, we then discuss how subspace information may be quantified in order to construct effective approximate projections.

Consider an initial approximation $\boldsymbol{x}_0$ to the solution of (1.1) for which we seek the MR approximation $\boldsymbol{x}_0 + \boldsymbol{c}$ with $\boldsymbol{c}$ selected from the correction space $\mathscr{C}$. We assume $\mathscr{C}$ to be the direct sum $\mathscr{C} = \mathscr{C}_1 \oplus \mathscr{C}_2$ of two spaces $\mathscr{C}_1$ and $\mathscr{C}_2$, and our goal is to obtain the MR approximation as the result of two separate projection steps involving $\mathscr{C}_1$ and $\mathscr{C}_2$, respectively. This task is equivalent to finding the best approximation $\boldsymbol{w} = A\boldsymbol{c} \in \mathscr{W} = A\mathscr{C} = \mathscr{W}_1 \oplus \mathscr{W}_2$ to $\boldsymbol{r}_0$, where $\mathscr{W}_j := A\mathscr{C}_j$, $j = 1, 2$.

If, in a first step, we obtain the best approximation $\boldsymbol{w}_1 = P_{\mathscr{W}_1}\boldsymbol{r}_0$ in $\mathscr{W}_1$, then the best approximation in $\mathscr{W}$ is obtained by introducing the orthogonal complement $\mathscr{Z} := \mathscr{W} \cap \mathscr{W}_1^{\perp}$ of $\mathscr{W}_1$ in $\mathscr{W}$, in terms of which $\mathscr{W}$ has the direct *and orthogonal* decomposition $\mathscr{W} = \mathscr{W}_1 \oplus \mathscr{Z}$. The global best approximation is now given by

$$\boldsymbol{w} := P_{\mathscr{W}}\boldsymbol{r}_0 = (P_{\mathscr{W}_1} + P_{\mathscr{Z}})\boldsymbol{r}_0 = P_{\mathscr{W}_1}\boldsymbol{r}_0 + P_{\mathscr{Z}}(I - P_{\mathscr{W}_1})\boldsymbol{r}_0. \tag{3.12}$$

The last expression shows that the contribution from the second projection consists of the orthogonal projection onto $\mathscr{Z}$ of the error $(I - P_{\mathscr{W}_1})\boldsymbol{r}_0$ of the first approximation.

Expressing all spaces in terms of $\mathscr{C}_1$ and $\mathscr{C}_2$ and noting that $\mathscr{Z} = (I - P_{A\mathscr{C}_1})A\mathscr{C}_2$, we conclude that the correction $\boldsymbol{c}$ associated with the residual approximation $\boldsymbol{w}$ satisfies

$$A\boldsymbol{c} = \boldsymbol{w} = P_{A\mathscr{C}_1}\boldsymbol{r}_0 + P_{(I-P_{A\mathscr{C}_1})A\mathscr{C}_2}(I - P_{A\mathscr{C}_1})\boldsymbol{r}_0.$$

The global correction is thus of the form $\boldsymbol{c} = \boldsymbol{c}_1 + \boldsymbol{d}$, where

$$A\boldsymbol{c}_1 = P_{A\mathscr{C}_1}\boldsymbol{r}_0 \tag{3.13}$$

$$A\boldsymbol{d} = P_{(I-P_{A\mathscr{C}_1})A\mathscr{C}_2}(I - P_{A\mathscr{C}_1})\boldsymbol{r}_0. \tag{3.14}$$

The solution $\boldsymbol{c}_1$ of (3.13) is simply the MR solution of the equation $A\boldsymbol{c} = \boldsymbol{r}_0$ with respect to the correction space $\mathscr{C}_1$. To obtain a useful representation of $\boldsymbol{d}$, we note that the right hand side of (3.14) may be viewed as the MR approximation with respect to $\mathscr{C}_2$ of the equation

$$(I - P_{A\mathscr{C}_1})A\boldsymbol{c} = (I - P_{A\mathscr{C}_1})\boldsymbol{r}_0. \tag{3.15}$$

Lemma 3.3.1. *The operator $(I - P_{A\mathscr{C}_1})A$ restricted to $\mathscr{C}_2$ is a bijection from $\mathscr{C}_2$ to $\mathscr{Z}$.*

Proof. The assertion follows by showing that the operator in question is one-to-one: $(I - P_{A\mathscr{C}_1})A\widetilde{\boldsymbol{c}} = \boldsymbol{0}$ for $\widetilde{\boldsymbol{c}} \in \mathscr{C}_2$ implies $A\widetilde{\boldsymbol{c}} \in A\mathscr{C}_1 \cap A\mathscr{C}_2 = \{\boldsymbol{0}\}$. □

The solution $\boldsymbol{d}$ of (3.14) yielding the second component of the combined correction $\boldsymbol{c}$ may thus be obtained by first determining the MR solution $\boldsymbol{c}_2$ of (3.15) and then evaluating

$$\boldsymbol{d} = A^{-1}(I - P_{A\mathscr{C}_1})A\boldsymbol{c}_2 = \boldsymbol{c}_2 - A^{-1}P_{A\mathscr{C}_1}A\boldsymbol{c}_2. \tag{3.16}$$

Lemma 3.3.2. *The operator $P := A^{-1}(I - P_{A\mathscr{C}_1})A$ restricted to $\mathscr{C}_2$ is the oblique projection onto $A^{-1}\mathscr{Z}$ along $\mathscr{C}_1$.*

Proof. The projection property follows immediately upon squaring P. Since A is nonsingular, $\mathscr{N}(P) = A^{-1}\mathscr{W}_1 = \mathscr{C}_1$ and $\mathscr{R}(P) = A^{-1}(A\mathscr{C}_1)^{\perp}$. Restricted to $\mathscr{C}_2$, the range reduces to the preimage under A of the orthogonal complement of $A\mathscr{C}_1$ with respect to $A\mathscr{C}_2$, i.e., $A^{-1}\mathscr{Z}$. □

At first glance, the evaluation of $\boldsymbol{d}$ as given in (3.16) appears to require a multiplication by A as well as the solution of another equation involving A with a right hand side from $A\mathscr{C}_1$, in addition to the computation of the two projections. In fact, we show how $\boldsymbol{d}$ can be calculated inexpensively using quantities generated in the course of the two MR approximation steps.

Assume $\mathscr{C}_1$ has dimension m and that Algorithm 3.2.3 (FGMRES) has been employed to obtain the MR approximation to the solution of $A\boldsymbol{c} = \boldsymbol{r}_0$ with respect to $\mathscr{C}_1$. If $C_m^{(1)} = [\boldsymbol{c}_1^{(1)}, \ldots, \boldsymbol{c}_m^{(1)}]$ denotes a basis of $\mathscr{C}_1$, then, besides the MR approximation $\boldsymbol{c}_1$, which has the coordinate representation $\boldsymbol{c}_1 = C_m^{(1)}\boldsymbol{y}_1$ with $\boldsymbol{y}_1 \in \mathbb{C}^m$, FGMRES also constructs the Arnoldi-type decomposition (3.9), which we write here as $AC_m^{(1)} = V_{m+1}^{(1)}\widetilde{H}_m^{(1)}$. The QR decomposition $\widetilde{H}_m^{(1)} = \widetilde{Q}_m^{(1)}R_m^{(1)}$ (cf. (2.49),(2.50)) makes available the Paige-Saunders basis $\widehat{V}_m^{(1)}$ (cf. (2.63)), which forms an orthonormal basis of $A\mathscr{C}_1$. Note also that, in view of relation (3.11), there holds

$$A^{-1}\widehat{V}_m^{(1)} = C_m^{(1)}R_m^{-1}. \tag{3.17}$$

The orthogonal projection $P_{A\mathscr{C}_1}$ may be expressed in terms of $\widehat{V}_m^{(1)}$ as $\widehat{V}_m^{(1)}\big[\widehat{V}_m^{(1)}\big]^*$, and, denoting the residual of the first MR approximation by $\boldsymbol{r}_1 := \boldsymbol{r}_0 - A\boldsymbol{c}_1$, equation (3.15) may be written

$$\Big(I - \widehat{V}_m^{(1)}\big[\widehat{V}_m^{(1)}\big]^*\Big)A\boldsymbol{c} = \boldsymbol{r}_1.$$

Applying Algorithm 3.2.3 to equation (3.15) using the basis $C_k^{(2)} = [\boldsymbol{c}_1^{(2)}, \ldots, \boldsymbol{c}_k^{(2)}]$ of the k-dimensional correction space $\mathscr{C}_2$ thus produces the decomposition

$$\Big(I - \widehat{V}_m^{(1)}\big[\widehat{V}_m^{(1)}\big]^*\Big)AC_k^{(2)} = V_{k+1}^{(2)}\widetilde{H}_k^{(2)} \tag{3.18}$$

as well as the MR approximation $\boldsymbol{c}_2 = C_k^{(2)}\boldsymbol{y}_2$, $\boldsymbol{y}_2 \in \mathbb{C}^k$. The solution $\boldsymbol{d}$ of (3.14) as given in (3.16) can now be expressed as

$$\begin{aligned}\boldsymbol{d} &= \boldsymbol{c}_2 - A^{-1}P_{A\mathscr{C}_1}A\boldsymbol{c}_2 = C_k^{(2)}\boldsymbol{y}_2 - A^{-1}\widehat{V}_m^{(1)}\big[\widehat{V}_m^{(1)}\big]^* AC_k^{(2)}\boldsymbol{y}_2 \\ &= C_k^{(2)}\boldsymbol{y}_2 - C_m^{(1)}\big[R_m^{(1)}\big]^{-1}\big(\big[\widehat{V}_m^{(1)}\big]^* AC_k^{(2)}\big)\boldsymbol{y}_2,\end{aligned}$$

which shows that the action of A^{-1} in (3.16) is effected by the inverse of the (small) triangular matrix $R_m^{(1)}$. We further observe that the evaluation of $A\boldsymbol{c}_2$ in (3.16) is accomplished through the $m \times k$ matrix $\big[\widehat{V}_m^{(1)}\big]^* AC_k^{(2)}$, which is available at no extra cost as a by-product of the orthogonalization process carried out in the second MR step to obtain (3.18). In fact, (3.17) and (3.18) can be combined to yield the global decomposition

$$A\begin{bmatrix}C_m^{(1)} & C_k^{(2)}\end{bmatrix} = \begin{bmatrix}\widehat{V}_m^{(1)} & V_{k+1}^{(2)}\end{bmatrix}\begin{bmatrix}R_m^{(1)} & \big[\widehat{V}_m^{(1)}\big]^* AC_k^{(2)} \\ O & \widetilde{H}_k^{(2)}\end{bmatrix} \tag{3.19}$$

with respect to $\mathscr{C}$. We summarize the coordinate representation of these two successive projections in

Theorem 3.3.3. *The MR approximation of the solution of $A\boldsymbol{c} = \boldsymbol{r}_0$ with respect to the correction space $\mathscr{C} = \mathscr{C}_1 \oplus \mathscr{C}_2$ is given by*

$$\boldsymbol{c} = C_m^{(1)}\boldsymbol{y}_1 + C_k^{(2)}\boldsymbol{y}_2 + C_m^{(1)}\big[R_m^{(1)}\big]^{-1}\big(\big[\widehat{V}_m^{(1)}\big]^* AC_k^{(2)}\big)\boldsymbol{y}_2,$$

where the coefficient vectors $\boldsymbol{y}_1 \in \mathbb{C}^m$ and $\boldsymbol{y}_2 \in \mathbb{C}^k$ satisfy the least-squares problems

$$\Big\|\|\boldsymbol{r}_0\|\boldsymbol{u}_1^{(m+1)} - \widetilde{H}_m^{(1)}\boldsymbol{y}\Big\| \to \min_{\boldsymbol{y}\in\mathbb{C}_m}, \qquad \Big\|\|\boldsymbol{r}_1\|\boldsymbol{u}_1^{(k+1)} - \widetilde{H}_k^{(2)}\boldsymbol{y}\Big\| \to \min_{\boldsymbol{y}\in\mathbb{C}_k}$$

and the matrices $C_m^{(1)}$, $C_k^{(2)}$, $\widehat{V}_m^{(1)}$, $R_m^{(1)}$, $\widetilde{H}_m^{(1)}$, and $\widetilde{H}_k^{(2)}$ as well as the vector $\boldsymbol{r}_1$ are defined above.

Remark 3.3.4. The decomposition (3.19) is a slight modification of the standard Arnoldi-type decomposition (3.9), which, translated to the present context, would have the form

$$A\begin{bmatrix}C_m^{(1)} & C_k^{(2)}\end{bmatrix} = \begin{bmatrix}V_{m+1}^{(1)} & V_k^{(2)}\end{bmatrix}\widetilde{H}_{m+k}, \qquad \widetilde{H}_{m+k} = \begin{bmatrix}\widetilde{H}_m^{(1)} & H_{m+1,k} \\ O & \widetilde{R}_k\end{bmatrix} \tag{3.20}$$

with an upper Hessenberg matrix $\widetilde{H}_{m+k} \in \mathbb{C}^{(m+k+1)\times(m+k)}$ composed of the submatrices $H_{m+1,k} \in \mathbb{C}^{(m+1)\times k}$, the upper triangular matrix $\widetilde{R}_k \in \mathbb{C}^{k\times k}$ and the upper Hessenberg matrix $\widetilde{H}_m^{(1)} \in \mathbb{C}^{(m+1)\times m}$ associated with the Arnoldi decomposition of A with respect to $C_m^{(1)}$. The modified decomposition (3.19) is obtained from (3.20) by substituting the QR decomposition (2.49) and introducing the Paige and Saunders basis (2.63), which also reveals the relations

$$V_{k+1}^{(2)} = \begin{bmatrix}\widetilde{\boldsymbol{v}}_{m+1} & V_k^{(2)}\end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix}H_{m+1,k} \\ \widetilde{R}_k\end{bmatrix} = \begin{bmatrix}\big[\widehat{V}_m^{(1)}\big]^* AC_k^{(2)} \\ \widetilde{H}_k^{(2)}\end{bmatrix}.$$

(Note that the last equation contains two different block partitionings of the $(m+k+1)\times k$ matrices: that on the left is split into an $(m+1)\times k$ and a $k\times k$ block, while the blocks on the right are of dimension $m\times k$ and $(k+1)\times k$, respectively.)

3.4 Incomplete Orthogonalization

The MR approximation applied to equation (3.15) in effect maintains orthogonality of the basis vectors of the residual space $\mathscr{V}_2$ against $\mathscr{W}_1 = A\mathscr{C}_1$. Computationally, this is manifested in the generation of the $m \times k$ matrix $\big[\widehat{V}_m^{(1)}\big]^* AC_k^{(2)}$ in the course of the orthogonalization process (cf. (3.19)). In order to reduce the cost of both the storage of $\widehat{V}_m^{(1)}$ and the work involved in the orthogonalization, we now consider performing the MR approximation to the solution of (3.15) only approximately in the sense that orthogonality is maintained only against a subspace of $\mathscr{W}_1$ of fixed dimension. When faced with the choice of such a subspace against which one can afford to maintain orthogonality, one possible criterion is to select that space which results in the greatest reduction of the residual norm after the second MR approximation. Such an approach was proposed by de Sturler (1996), and will be further described in Section 5.4.4.

As in Section 3.3, consider the MR approximation with respect to the correction space $\mathscr{C} = \mathscr{C}_1 \oplus \mathscr{C}_2$. The global MR approximation (3.12) consists of an MR approximation with respect to $\mathscr{C}_1$ followed by a second projection involving the orthogonal complement $\mathscr{Z} := (I - P_{\mathscr{W}_1})\mathscr{W}_2$ of $\mathscr{W}_1 = A\mathscr{C}_1$ with respect to $\mathscr{W}_2 = A\mathscr{C}_2$. The simplest approach of completely omitting the orthogonalization involved in constructing $P_{\mathscr{Z}}$ results in the combined approximation

$$\widetilde{\boldsymbol{w}} := P_{\mathscr{W}_1}\boldsymbol{r}_0 + P_{\mathscr{W}_2}(I - P_{\mathscr{W}_1})\boldsymbol{r}_0,$$

in place of (3.12). This is the standard way of restarting an MR algorithm. Besides the two extremes of complete orthogonalization against $\mathscr{W}_1$ or none at all, it is also possible to orthogonalize against only a subspace $\widetilde{\mathscr{W}_1} \subset \mathscr{W}_1$ of dimension $\ell < m$, which brings up the problem of determining $\widetilde{\mathscr{W}_1}$ such that, if orthogonality of the residual space $\mathscr{V}_2$ of the second MR approximation is maintained against $\widetilde{\mathscr{W}_1}$, this results in the smallest residual norm over all ℓ-dimensional subspaces of $\mathscr{W}_1$.

The solution of this problem is greatly facilitated by a judicious choice of bases: Let $W_m^{(1)} = [\boldsymbol{w}_1^{(1)}, \ldots, \boldsymbol{w}_m^{(1)}]$ and $W_k^{(2)} = [\boldsymbol{w}_1^{(2)}, \ldots, \boldsymbol{w}_k^{(2)}]$ denote biorthogonal orthonormal bases of $\mathscr{W}_1$ and $\mathscr{W}_2$ ordered such that the (diagonal) $m \times k$ matrix $\Gamma := \big[W_m^{(1)}\big]^* W_k^{(2)}$ has nonincreasing nonnegative entries $\gamma_1, \ldots, \gamma_{\min\{m,k\}}$. The numbers γ_j are the cosines of the canonical angles between the spaces $\mathscr{W}_1$ and $\mathscr{W}_2$ (cf. (Stewart 1998, Chapter 4.5)) and therefore lie between zero and one. In addition, the assumption $\mathscr{C}_1 \cap \mathscr{C}_2 = \{\boldsymbol{0}\}$ along with the nonsingularity of A implies $\mathscr{W}_1 \cap \mathscr{W}_2 = \{\boldsymbol{0}\}$ and therefore each γ_j is strictly less than one.

An orthogonal basis of $\mathscr{Z}$ is given by $\widehat{Z}_k := \big(I - W_m^{(1)}\big[W_m^{(1)}\big]^*\big)W_k^{(2)}$, and we set $\widehat{Z}_k^*\widehat{Z}_k = I - \Gamma^*\Gamma =: \Sigma^2 \in \mathbb{C}^{k\times k}$, where $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_k)$ with

$$\sigma_j = \begin{cases} \sqrt{1-\gamma_j^2}, & 1 \le j \le \min\{k,m\} \\ 1, & \text{otherwise,} \end{cases}$$

in view of which $Z_k := \widehat{Z}_k\Sigma^{-1}$ is an orthonormal basis of $\mathscr{Z}$. Denoting $Z_k = [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k]$, the following theorem expresses the effect of complete orthogonalization versus none at all:

Theorem 3.4.1. *In the notation introduced above and with* $\boldsymbol{r}_1 := (I - P_{\mathscr{W}_1})\boldsymbol{r}_0$, *there holds*

$$(P_{\mathscr{W}_2} - P_{\mathscr{Z}})\boldsymbol{r}_1 = \sum_{j=1}^{\min\{k,m\}} (\boldsymbol{r}_1, \boldsymbol{z}_j)\gamma_j(\sigma_j \boldsymbol{w}_j^{(1)} - \gamma_j \boldsymbol{z}_j), \tag{3.21}$$

$$\|(P_{\mathscr{W}_2} - P_{\mathscr{Z}})\boldsymbol{r}_1\|^2 = \sum_{j=1}^{\min\{k,m\}} \gamma_j^2 |(\boldsymbol{r}_1, \boldsymbol{z}_j)|^2. \tag{3.22}$$

Proof. Taking note of $\boldsymbol{r}_1 \perp \mathscr{W}_1$ and $W_k^{(2)} = Z_k\Sigma + W_m^{(1)}\Gamma$, we obtain

$$\begin{aligned}(P_{\mathscr{W}_2} - P_{\mathscr{Z}})\boldsymbol{r}_1 &= \left(W_k^{(2)}\left[W_k^{(2)}\right]^* - Z_k Z_k^*\right)\boldsymbol{r}_1 \\ &= \left((Z_k\Sigma + W_m^{(1)}\Gamma)(Z_k\Sigma + W_m^{(1)}\Gamma)^* - Z_k Z_k^*\right)\boldsymbol{r}_1 \\ &= (W_m^{(1)}\Gamma\Sigma - Z_k\Gamma^H\Gamma)Z_k^*\boldsymbol{r}_1,\end{aligned}$$

which is a reformulation of (3.21). Taking norms and noting $\mathscr{W}_1 \perp \mathscr{Z}$ as well as $\gamma_j^2 + \sigma_j^2 = 1$ yields (3.22):

$$\|(P_{\mathscr{W}_2} - P_{\mathscr{Z}})\boldsymbol{r}_1\|^2 = (Z^*\boldsymbol{r}_1)^*(\Sigma\Gamma^H\Gamma\Sigma + (\Gamma^H\Gamma)^2)(Z^*\boldsymbol{r}_1) = \|\Gamma Z^*\boldsymbol{r}_1\|_2^2.$$

□

We see that the difference between the two projection depends on the $\mathscr{Z}$-components of the approximation error $\boldsymbol{r}_1$ remaining after the first projection weighted by the corresponding cosines γ_j of the canonical angles between $\mathscr{W}_1$ and $\mathscr{W}_2$. Whenever $\gamma_j = 0$, the projection onto $\mathscr{W}_2$ would already have produced the correct component in the direction $\boldsymbol{z}_j$, whereas in case $\gamma_j = 1$ the associated basis vectors $\boldsymbol{w}_j^{(1)}$ and $\boldsymbol{w}_j^{(2)}$ are collinear and $P_{\mathscr{W}_2}$ would have yielded no component in direction $\boldsymbol{z}_j$.

To consider the case of incomplete orthogonalization, let $\widetilde{\mathscr{W}_1} \subset \mathscr{W}_1$ with $\dim \widetilde{\mathscr{W}_1} = \ell < m$. By orthogonalizing the basis of $\mathscr{W}_2$ against $\widetilde{\mathscr{W}_1}$, we construct the orthogonal projection onto $\widetilde{\mathscr{Z}} := (\widetilde{\mathscr{W}_1} \oplus \mathscr{W}_2) \cap \widetilde{\mathscr{W}_1}^{\perp}$, which, applied to $\boldsymbol{r}_1$, yields the difference between $P_{\mathscr{W}_1}\boldsymbol{r}_0$ and best approximation of $\boldsymbol{r}_0$ in $\widetilde{\mathscr{W}_1} \oplus \mathscr{W}_2$.

Theorem 3.4.2. *Of all* ℓ*-dimensional subspaces* $\widetilde{\mathscr{W}_1} \subset \mathscr{W}_1$, *that which minimizes* $\|(P_{\widetilde{\mathscr{Z}}} - P_{\mathscr{Z}})(I - P_{\mathscr{W}_1})\boldsymbol{r}_0\|$ *over all* $\boldsymbol{r}_0 \in \mathscr{H}$ *is given by* $\widetilde{\mathscr{W}_1} = \mathrm{span}\{\boldsymbol{w}_1^{(1)}, \ldots, \boldsymbol{w}_\ell^{(1)}\}$, *and results in*

$$\|(P_{\widetilde{\mathscr{Z}}} - P_{\mathscr{Z}})(I - P_{\mathscr{W}_1})\boldsymbol{r}_0\| = \|(P_{\widetilde{\mathscr{Z}}} - P_{\mathscr{Z}})\boldsymbol{r}_1\| = \sum_{j=\ell+1}^{\min\{k,m\}} \gamma_j^2 |(\boldsymbol{r}_1, \boldsymbol{z}_j)|^2. \tag{3.23}$$

Proof. Any orthonormal basis $\widetilde{W}_\ell^{(1)} = [\widetilde{\boldsymbol{w}}_1, \ldots, \widetilde{\boldsymbol{w}}_\ell]$ of $\widetilde{\mathscr{W}_1}$ has the form $\widetilde{W}_\ell^{(1)} = W_m^{(1)}Q_1$ with a matrix $Q_1 \in \mathbb{C}^{m\times\ell}$ consisting of the first ℓ columns of a unitary $m \times m$ matrix $Q = [Q_1\, Q_2]$. We obtain a basis of $\widetilde{\mathscr{Z}}$ by orthogonalizing $W_k^{(2)}$ against $\widetilde{W}_\ell^{(1)}$:

$$\begin{aligned}\widehat{Z}_\ell :&= \left(I - \widetilde{W}_\ell^{(1)}\left[\widetilde{W}_\ell^{(1)}\right]^*\right)W_k^{(2)} = W_k^{(2)} - W_m^{(1)}Q_1Q_1^H\Gamma \\ &= (Z_k\Sigma + W_k^{(2)}\Gamma) - W_m^{(1)}Q_1Q_1^H\Gamma = Z_k\Sigma + W_m^{(1)}(I - Q_1Q_1^H)\Gamma \\ &= Z_k\Sigma + W_m^{(1)}Q_2Q_2^H\Gamma.\end{aligned}$$

Because of $0 \le \gamma_j < 1$ the Hermitian matrix

$$\widehat{Z}_\ell^* \widehat{Z}_\ell = \Sigma^2 + \Gamma^* Q_2 Q_2^H \Gamma = I - \Gamma^* Q_1 Q_1^H \Gamma =: S^2$$

is positive definite and therefore possesses a square root S, by means of which we obtain an orthonormal basis of $\widetilde{\mathscr{Z}}$ as $Z_\ell := \widehat{Z}_\ell S^{-1}$. Again recalling $\boldsymbol{r}_1 \perp \mathscr{W}_1$, we obtain for the difference of the two projections

$$\begin{aligned}(P_{\widetilde{\mathscr{Z}}} - P_{\mathscr{Z}})\boldsymbol{r}_1 &= (Z_\ell Z_\ell^* - Z_k Z_k^*)\boldsymbol{r}_1 \\ &= \left(Z_k(\Sigma S^{-2}\Sigma - I) + W_m^{(1)}(Q_2 Q_2^H \Gamma S^{-2}\Sigma)\right) Z_k^* \boldsymbol{r}_1.\end{aligned} \tag{3.24}$$

From the definition of S^2, we have

$$\Sigma S^{-2}\Sigma = (\Sigma^{-1} S^2 \Sigma^{-1})^{-1} = (I + \Sigma^{-1}\Gamma^H Q_2 Q_2^H \Gamma \Sigma^{-1})^{-1} =: (I + MM^H)^{-1},$$

with $M = \Sigma^{-1}\Gamma^H Q_2$. We thus obtain

$$\Sigma S^{-2}\Sigma - I = (I + MM^H)^{-1} - I = -MM^H(I + MM^H)^{-1}$$

as well as $Q_2 Q_2^H \Gamma S^{-2}\Sigma = Q_2 M^H (I + MM^H)^{-1}$, which we insert in (3.24) to obtain

$$\|(P_{\widetilde{\mathscr{Z}}} - P_{\mathscr{Z}})\boldsymbol{r}_1\|^2 = (Z_k^* \boldsymbol{r}_1)\left[(I + MM^H)^{-1} MM^H\right](Z_k^* \boldsymbol{r}_1). \tag{3.25}$$

This expression is minimized for all $\boldsymbol{r}_1$—hence also for all $\boldsymbol{r}_0$—by choosing Q_1 to minimize the largest eigenvalue of the Hermitian matrix $(I + MM^H)^{-1}MM^H$ or, equivalently, that of $MM^H = \Sigma^{-1}\Gamma^H(I - Q_1 Q_1^H)\Gamma\Sigma^{-1}$. The entries γ_j/σ_j of the $m \times k$ diagonal matrix $\Gamma\Sigma^{-1}$ are nonincreasing, hence the minimum occurs for

$$Q_1 = \begin{bmatrix} I_\ell \\ O \end{bmatrix}$$

and the assertion follows by inserting the resulting choice of M in (3.25). □

Chapter 4

Krylov Subspace Methods

The overwhelming majority of subspace correction methods for solving linear systems of equations employ correction spaces of a particularly simple structure. These spaces are known as *Krylov spaces* (or *Krylov subspaces*) and are defined by

$$\mathscr{K}_m := \mathscr{K}_m(A, \boldsymbol{v}) := \operatorname{span}\{\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^{m-1}\boldsymbol{v}\}, \qquad \boldsymbol{v} \in \mathscr{H}. \tag{4.1}$$

In other words, $\mathscr{K}_m$ consist of all polynomials in A of a degree at most $m-1$ applied to a fixed vector $\boldsymbol{v}$, which in the sequel will always be the initial residual $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$. The name refers to a method introduced by Krylov (1931) for determining divisors of the minimal polynomial of an operator for the purpose of computing eigenvalues, in which such spaces were used (see also Householder (1964, Section 6.1)). In this chapter we survey some of the ramifications of this choice. Section 4.1 collects some advantages of using Krylov spaces, and Section 4.2 specializes the results of Chapter 2 to Krylov spaces. Section 4.3 discusses the close link between a Krylov space and an associated ring of polynomials and derives the polynomials which correspond to the MR and OR residual vectors, the zeros of which are harmonic Ritz values and Ritz values, respectively, of A with respect to the Krylov space. This section closes with a brief derivation of Sorensen's *implicitly restarted Arnoldi method* (Sorensen 1992) as a device for manipulating Krylov spaces and their Arnoldi factorizations. Section 4.4 attempts to sketch the development of Krylov subspace MR and OR methods since the introduction of the conjugate gradient method in Hestenes & Stiefel (1952). Finally, Section 4.5 contains a brief discussion of short recurrence formulas for Krylov subspace methods.

4.1 Why Krylov Subspaces?

Using (shifted) Krylov spaces to construct approximate solutions to linear equations is as old as stationary iterative methods (cf. Varga (1999)): given a splitting $A = M - N$ with M nonsingular, the induced stationary iteration

$$\boldsymbol{x}_m = T\boldsymbol{x}_{m-1} + \boldsymbol{c}, \qquad m = 1, 2, \ldots \tag{4.2}$$

with $T = M^{-1}N$ and $\boldsymbol{c} = M^{-1}\boldsymbol{b}$ generates the approximations

$$\boldsymbol{x}_m = \boldsymbol{x}_0 + (I + T + \cdots + T^{m-1})\boldsymbol{r}_0 \in \boldsymbol{x}_0 + \mathscr{K}_m(T, \boldsymbol{r}_0).$$

One way of viewing Krylov subspace MR and OR methods is as a more clever strategy for choosing the approximations in $\boldsymbol{x}_0 + \mathscr{K}_m$ or, which is the same, as techniques which accelerate the basic stationary iterative method (4.2). This was the motivation in Varga (1999), where the term *semi-iterative methods* is used for this approach. Even earlier, Chebyshev polynomials were used to accelerate stationary iterations in the case when the splitting leads to a Hermitian positive definite matrix T (Shortley 1953). As we shall see, a subtle difference between such stationary iterations and Krylov subspace MR and OR methods is that the latter terminate with the exact solution whenever the Krylov space becomes A-invariant, which is not the case for stationary methods or those based on Chebyshev polynomials.

One regard in which (4.1) is a reasonable choice for a correction space is that it enables the successive generation of the sequence of correction spaces $\{\mathscr{C}_m\}$ using only multiplication of vectors by A, an operation which is inexpensive whenever A is represented by a sparse or structured matrix or when the action of A can be implemented efficiently without reference to a matrix representation. Moreover, note that $\mathscr{C}_m = \mathscr{K}_m(A, \boldsymbol{r}_0)$ results in the residual space (cf. (3.2))

$$\mathscr{V}_{m+1} = \operatorname{span}\{\boldsymbol{r}_0\} + A\mathscr{C}_m = \operatorname{span}\{\boldsymbol{r}_0\} + A\mathscr{K}_m = \mathscr{K}_{m+1},$$

i.e., the residual space $\mathscr{V}_{m+1}$ of index $m+1$ coincides with the correction space $\mathscr{C}_{m+1}$ of the next iteration, obviating the need to store two sets of basis vectors in Algorithm 3.2.1 and Algorithm 3.2.3, as we shall see in Section 4.2.1. This effectively halves the storage requirements of these algorithms, which in the Krylov subspace case are known as the GMRES algorithm of Saad & Schultz (1986) and Walker and Zhou's "simpler GMRES" (Walker & Zhou 1994), versus the GCR variant. In the Krylov subspace case, the Gram-Schmidt orthogonalization procedure (2.41) is known as the *Arnoldi process* (Arnoldi 1951), and the resulting Arnoldi-type decomposition (3.9) becomes the proper Arnoldi decomposition

$$AV_m = V_{m+1}\widetilde{H}_m = V_m H_m + \eta_{m+1,m}\boldsymbol{v}_{m+1}\boldsymbol{u}_m^\top, \tag{4.3}$$

with the entries of the Hessenberg matrix $\widetilde{H}_m$ given by (cf. (2.42), (3.10))

$$\eta_{j,m} = (A\boldsymbol{v}_m, \boldsymbol{v}_j), \qquad j = 1, \ldots, m+1, \quad m \geq 1. \tag{4.4}$$

This identifies H_m as the *orthogonal section* of A onto $\mathscr{K}_m$, i.e., it represents the linear map $A_{\mathscr{K}_m} := P_{\mathscr{K}_m} A|_{\mathscr{K}_m} : \mathscr{K}_m \to \mathscr{K}_m$ with respect to the basis V_m. We also note that the characterization of the termination index L (cf. (2.40) and (3.3)) now becomes

$$L = \min\{m \,:\, \mathscr{K}_m(A, \boldsymbol{r}_0) = \mathscr{K}_{m+1}(A, \boldsymbol{r}_0)\} = \min\{m \,:\, A^{-1}\boldsymbol{r}_0 \in \mathscr{K}_m(A, \boldsymbol{r}_0)\}. \tag{4.5}$$

In particular, L is the index of the smallest A-invariant Krylov subspace with initial vector $\boldsymbol{r}_0$. This fact can sometimes exploited, e.g. when equations with a nearby operator $\widetilde{A}$ can be solved inexpensively and $\widetilde{A} - A$ has low rank, since then $\widetilde{A}$ can be used as a *preconditioner*, resulting in a system with $L = \operatorname{rank}(\widetilde{A} - A)$.

Whether or not Krylov spaces are well suited as correction spaces will, as shown before, depend on the behavior of the angles $\measuredangle(\mathscr{K}_m, A\mathscr{K}_m)$ as m approaches ∞. There are classes of problems for which this behavior is very favorable. An example where the angles

actually tend to zero, which, in view of (2.25), implies superlinear convergence of the MR and OR approximants, is given by second-kind Fredholm equations (cf. Eiermann & Ernst (1998, Theorem 6.12)). On the other hand, there are matrix problems of dimension n for which $\measuredangle(\mathscr{K}_m, A\mathscr{K}_m) = \pi/2$ $(m = 1, 2, \ldots, n-1)$, i.e., no Krylov subspace method is able to improve the initial residual until the very last step. The convergence properties of Krylov subspace methods will be discussed in more detail in Chapter 6.

Finally, the theoretical investigation of Krylov subspace methods is greatly facilitated due to a close connection between a Krylov space and an associated space of polynomials, as discussed in Section 4.3 below. When A belongs to the class of normal operators, this polynomial description reduces the question of convergence of Krylov subspace methods to one of best approximation by polynomials in the complex plane.

4.2 MR and OR Approximations from Krylov Subspaces

In this section we specialize the abstract results of Chapter 2 to Krylov subspace MR and OR methods and point out the simplifications which result.

4.2.1 Krylov Subspace Methods Based on Orthonormal Bases

Following the MR approach, we now seek approximate solutions $\boldsymbol{x}_m^{\mathrm{MR}} = \boldsymbol{x}_0 + \boldsymbol{c}_m^{\mathrm{MR}}$ of (1.1) with $\boldsymbol{c}_m^{\mathrm{MR}} \in \mathscr{K}_m(A, \boldsymbol{r}_0)$ such that

$$\|\boldsymbol{r}_m^{\mathrm{MR}}\| = \|\boldsymbol{b} - A\boldsymbol{x}_m^{\mathrm{MR}}\| = \|\boldsymbol{r}_0 - A\boldsymbol{c}_m^{\mathrm{MR}}\| = \min_{\boldsymbol{c} \in \mathscr{K}_m(A, \boldsymbol{r}_0)} \|\boldsymbol{r}_0 - A\boldsymbol{c}\| \tag{4.6}$$

$(m = 1, 2, \ldots, L)$. Recall that $A\boldsymbol{c}_m^{\mathrm{MR}}$ is the best approximation to $\boldsymbol{r}_0$ from the space $\mathscr{W}_m = A\mathscr{K}_m(A, \boldsymbol{r}_0)$ and that, in the parlance of Chapter 2, $\boldsymbol{r}_m^{\mathrm{MR}} = \boldsymbol{r}_0 - A\boldsymbol{c}_m^{\mathrm{MR}}$ is the associated approximation error.

As for the OR approach, we now seek approximate solutions $\boldsymbol{x}_m^{\mathrm{OR}} = \boldsymbol{x}_0 + \boldsymbol{c}_m^{\mathrm{OR}}$ with $\boldsymbol{c}_m^{\mathrm{OR}} \in \mathscr{K}_m(A, \boldsymbol{r}_0)$ such that

$$\boldsymbol{r}_m^{\mathrm{OR}} = \boldsymbol{b} - A\boldsymbol{x}_m^{\mathrm{OR}} = \boldsymbol{r}_0 - A\boldsymbol{c}_m^{\mathrm{OR}} \perp \mathscr{K}_m(A, \boldsymbol{r}_0). \tag{4.7}$$

In the terminology of Section 2.2, $A\boldsymbol{c}_m^{\mathrm{OR}}$ is the OR approximation to $\boldsymbol{r}_0$ from $\mathscr{W}_m$ with error $\boldsymbol{r}_m^{\mathrm{OR}}$. As mentioned above, the test spaces defining the orthogonality conditions coincide with the Krylov spaces $\mathscr{V}_m = \mathscr{K}_m(A, \boldsymbol{r}_0)$.

We first turn to the question of existence of the two approximate solutions (cf. Corollary 2.2.2) and express the residual norms in terms of the angles

$$\varphi_m := \measuredangle\left(\mathscr{K}_m(A, \boldsymbol{r}_0), A\mathscr{K}_m(A, \boldsymbol{r}_0)\right) \tag{4.8}$$

between the subspaces $\mathscr{K}_m(A, \boldsymbol{r}_0)$ and $A\mathscr{K}_m(A, \boldsymbol{r}_0)$ (cf. Theorem 2.2.5). As in Chapter 2, we denote the associated sines and cosines by $s_m = \sin\varphi_m$ and $c_m = \cos\varphi_m$, and we recall that these coincide with the Givens parameters which occur in the updated QR-factorization of the Hessenberg matrix $\widetilde{H}_m$ (cf. Theorem 2.3.5 and the discussion that follows). Note that $\varphi_m = 0$ if and only if $m = L$.

Corollary 4.2.1. *The MR approximation $\boldsymbol{x}_m^{MR} \in \boldsymbol{x}_0 + \mathscr{K}_m(A, \boldsymbol{r}_0)$ defined by (4.6) exists for every $m = 1, 2, \ldots, L$. The associated residual vector satisfies*

$$\|\boldsymbol{r}_m^{MR}\| = \|\boldsymbol{r}_{m-1}^{MR}\| s_m = \|\boldsymbol{r}_0\| \prod_{j=1}^{m} s_j.$$

In particular, this implies $\|\boldsymbol{r}_0^{MR}\| \geq \|\boldsymbol{r}_1^{MR}\| \geq \cdots \geq \|\boldsymbol{r}_{L-1}^{MR}\| > \|\boldsymbol{r}_L^{MR}\| = 0$. For $m = 1, 2, \ldots, L$, the OR approximation $\boldsymbol{x}_m^{OR} \in \boldsymbol{x}_0 + \mathscr{K}_m(A, \boldsymbol{r}_0)$ defined by (4.7) exists if and only if $\varphi_m \neq \pi/2$. In this case the associated residual vector satisfies

$$\|\boldsymbol{r}_m^{OR}\| = \frac{\|\boldsymbol{r}_m^{MR}\|}{c_m} = \frac{\|\boldsymbol{r}_0\| \prod_{j=1}^{m} s_j}{c_m}.$$

In particular, $\boldsymbol{x}_L^{OR}$ exists and there holds $\boldsymbol{x}_L^{OR} = \boldsymbol{x}_L^{MR} = A^{-1}\boldsymbol{b}$.

Since $\sin\varphi_m = 1$ and $\cos\varphi_m = 0$ are equivalent statements, we see again that $\boldsymbol{x}_m^{\mathrm{OR}}$ exists if and only if $\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\| > \|\boldsymbol{r}_m^{\mathrm{MR}}\|$.

In Chapter 2 we saw that the existence of the OR approximation at step m requires that $\measuredangle(\mathscr{V}_m, \mathscr{W}_m) < \pi/2$ (cf. Remark 2.1.7), or equivalently, that the direction in which $\mathscr{W}_{m-1}$ is extended to $\mathscr{W}_m$ not be orthogonal to $\operatorname{span}\{\boldsymbol{r}_0\}$. As long as arbitrary expanding sequences are allowed as correction spaces as in Chapter 3, no condition on A can exclude the possibility of the OR approximation failing to exist: choosing e.g. as a new basis vector $\boldsymbol{c}_m$ of the correction space the pre-image under A of any vector orthogonal to $\boldsymbol{r}_0$ will lead to this situation in step m. When $\mathscr{C}_m$ is a Krylov space, however, there is a convenient criterion.

Theorem 4.2.2. *The OR approximation for the solution of (1.1) with respect to the sequence of Krylov spaces $\mathscr{K}_m(A, \boldsymbol{r}_0)$ exists for all initial residuals $\boldsymbol{r}_0$ if and only if the field of values*

$$W(A) := \left\{ \frac{(A\boldsymbol{v}, \boldsymbol{v})}{(\boldsymbol{v}, \boldsymbol{v})} : 0 \neq \boldsymbol{v} \in \mathscr{H} \right\} \tag{4.9}$$

of A does not include the origin.

Proof. By Theorem 2.2.3, existence and uniqueness of the OR approximation is equivalent with $c_m \neq 0$, where $c_m = \cos\measuredangle(\mathscr{V}_m, \mathscr{W}_m)$. In the present situation, using (2.21) and (2.1), this becomes

$$\begin{aligned} c_m &= \cos\measuredangle(\mathscr{K}_m, A\mathscr{K}_m) = \cos\measuredangle(\boldsymbol{r}_{m-1}^{\mathrm{MR}}, A\mathscr{K}_m) = \sup_{\boldsymbol{v}\in\mathscr{K}_m} \cos\measuredangle(\boldsymbol{r}_{m-1}^{\mathrm{MR}}, A\boldsymbol{v}) \\ &\geq \cos\measuredangle(\boldsymbol{r}_{m-1}^{\mathrm{MR}}, A\boldsymbol{r}_{m-1}^{\mathrm{MR}}) = \frac{|(\boldsymbol{r}_{m-1}^{\mathrm{MR}}, A\boldsymbol{r}_{m-1}^{\mathrm{MR}})|}{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\| \, \|A\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|}, \end{aligned}$$

where we have made use of the fact that $\boldsymbol{r}_{m-1}^{\mathrm{MR}} \in \mathscr{K}_m$. Consequently, if $0 \notin W(A)$ this implies $c_m \neq 0$ for all m. Conversely, if A is such that the OR approximation exists for all m and for all initial residuals, $(A\boldsymbol{v}, \boldsymbol{v}) = 0$ for $\boldsymbol{v} \neq 0$ would imply non-existence of the first OR approximation if one were to start with $\boldsymbol{r}_0 = \boldsymbol{v}$, a contradiction. □

The next following lemma, which we will make use of in Section 6.2.3, is due to Moret (1997) and gives a bound on $s_m = |(\boldsymbol{v}_{m+1}, \boldsymbol{w}_m)|$ (cf. Lemma 2.2.4) for the case when $\mathscr{V}_m$ is a Krylov space:

Lemma 4.2.3. *Let $\{\boldsymbol{v}_m\}$ be the Arnoldi basis defined by* (4.3) *and* (4.4) *with respect to $\boldsymbol{r}_0 \in \mathscr{H}$ and $A : \mathscr{H} \to \mathscr{H}$, where we assume A is bounded and possesses a bounded inverse. In addition, let $\{\boldsymbol{w}_j\}_{j\geq 1}$ denote an orthonormal system such that $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ is an orthonormal basis of $\mathscr{W}_m$. Then there holds*

$$|(\boldsymbol{v}_{m+1}, \boldsymbol{w}_m)| \leq \|A^{-1}\| \; (\boldsymbol{v}_{m+1}, A\boldsymbol{v}_m).$$

Proof. Since $A^{-1}\boldsymbol{w}_m \in \mathscr{V}_m$, we have $A^{-1}\boldsymbol{w}_m = P_{\mathscr{V}_m} A^{-1}\boldsymbol{w}_m$ and thus $\boldsymbol{w} = AP_{\mathscr{V}_m} A^{-1}\boldsymbol{w}_m$. We can therefore write $(\boldsymbol{v}_{m+1}, \boldsymbol{w}_m) = (\boldsymbol{v}_{m+1}, AP_{\mathscr{V}_m} A^{-1}\boldsymbol{w}_m)$. Moreover, since $A\boldsymbol{v}_j \in \mathscr{V}_m \perp \boldsymbol{v}_{m+1}$ for $1 \leq j \leq m-1$, we have

$$\begin{aligned} |(\boldsymbol{v}_{m+1}, AP_{\mathscr{V}_m} A^{-1}\boldsymbol{w}_m)| &= |(\boldsymbol{v}_{m+1}, \sum_{j=1}^{m} (A^{-1}\boldsymbol{w}_m, \boldsymbol{v}_j) A\boldsymbol{v}_j)| = |(\boldsymbol{v}_{m+1}, (A^{-1}\boldsymbol{w}_m, \boldsymbol{v}_m) A\boldsymbol{v}_m)| \\ &= |(A^{-1}\boldsymbol{w}_m, \boldsymbol{v}_m)(\boldsymbol{v}_{m+1}, A\boldsymbol{v}_m)| \leq \|A^{-1}\| \, (A\boldsymbol{v}_m, \boldsymbol{v}_{m+1}). \end{aligned}$$

Note that the modulus in $|(A\boldsymbol{v}_m, \boldsymbol{v}_{m+1})|$ can be omitted since $(A\boldsymbol{v}_m, \boldsymbol{v}_{m+1}) = \|(I - P_{\mathscr{V}_m})A\boldsymbol{v}_m\| \geq 0$. □

We next turn to the question of how $\boldsymbol{x}_m^{\mathrm{MR}}$ and $\boldsymbol{x}_m^{\mathrm{OR}}$ and their associated residual vectors are related (cf. (2.30), (2.31) and Corollary 2.2.7).

Corollary 4.2.4. *With the angle φ_m as defined by* (4.8) *and associated sine and cosine s_m and c_m there holds*

$$\boldsymbol{x}_m^{MR} = s_m^2 \boldsymbol{x}_{m-1}^{MR} + c_m^2 \boldsymbol{x}_m^{OR}, \qquad \boldsymbol{r}_m^{MR} = s_m^2 \boldsymbol{r}_{m-1}^{MR} + c_m^2 \boldsymbol{r}_m^{OR}$$

for $m = 1, 2, \ldots, L$, provided $\boldsymbol{x}_m^{OR}$ exists, i.e., that $\varphi_m \neq \pi/2$. Equivalently, we have

$$\begin{aligned} \|\boldsymbol{r}_m^{MR}\| &= \sqrt{1 - \frac{\|\boldsymbol{r}_m^{MR}\|}{\|\boldsymbol{r}_{m-1}^{MR}\|}} \, \|\boldsymbol{r}_m^{OR}\|, \\ \boldsymbol{x}_m^{MR} &= \boldsymbol{x}_m^{OR} + \frac{\|\boldsymbol{r}_m^{MR}\|}{\|\boldsymbol{r}_{m-1}^{MR}\|} \left(\boldsymbol{x}_{m-1}^{MR} - \boldsymbol{x}_m^{OR}\right). \end{aligned}$$

Of course, not only (2.30) and (2.31) but every relation presented in Section 2 can be translated to the Krylov subspace setting. This applies in particular to the results of Proposition 2.2.9 on minimal and quasi-minimal residual smoothing.

The identities contained in Corollaries 4.2.1 and 4.2.4 are not new and may be found e.g. in the papers of Brown (1991), Freund (1992*b*), Gutknecht (1993*a*), and Cullum & Greenbaum (1996). In these works, however, the sines and cosines result from the Givens rotations needed to construct a QR decomposition of the Hessenberg matrix analogous to (2.39) in the course of a specific algorithm for computing the MR and OR approximations. Section 2.3, specifically Theorem 2.3.5, reveals the more fundamental significance of these rotation angles, namely as the angles between the subspaces $\mathscr{K}_m(A, \boldsymbol{r}_0)$ and $A\mathscr{K}_m(A, \boldsymbol{r}_0)$.

We now turn to algorithms for computing MR approximations in the Krylov subspace case by specializing the three basic algorithms introduced in Chapter 3. We note that this by no means covers the multitude of Krylov subspace methods which have been proposed for non-Hermitian problems. A brief survey of the historically important methods is given in Section 4.4.

One step of the Walker and Zhou algorithm (Algorithm 3.2.1) consists of enlarging the correction space $\mathscr{C}_m$ by adding a new basis vector and then orthogonalizing its image against the previously generated orthonormal basis of $\mathscr{W}_m = A\mathscr{C}_m$, resulting in the decomposition (3.6). In the Krylov subspace case we have

$$\mathscr{C}_m = \mathscr{K}_m(A, \boldsymbol{r}_0), \qquad \mathscr{W}_m = A\mathscr{K}_m(A, \boldsymbol{r}_0) = \mathscr{K}_m(A, A\boldsymbol{r}_0),$$

and therefore the next correction space is $\mathscr{C}_{m+1} = \operatorname{span}\{\boldsymbol{r}_0\} + \mathscr{W}_m$. An obvious candidate for $\boldsymbol{c}_{m+1}$ is thus $\boldsymbol{w}_m$, resulting in the basis $C_{m+1} = [\boldsymbol{r}_0, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m]$ for the correction space, so that no separate basis of $\mathscr{C}_m$ is necessary. We summarize the Walker-Zhou MR algorithm for Krylov spaces in Algorithm 4.2.1. As in Algorithm 3.2.1, the residuals may

Algorithm 4.2.1 The Walker-Zhou MR algorithm for Krylov spaces.

$\boldsymbol{x}_0$ given, $\boldsymbol{r}_0 := \boldsymbol{b} - A\boldsymbol{x}_0$
$\boldsymbol{w}_1 = A\boldsymbol{r}_0/\|A\boldsymbol{r}_0\|, r_{1,1} = \|A\boldsymbol{r}_0\|$
for $j := 2$ **to** m
  $\widehat{\boldsymbol{w}} = A\boldsymbol{w}_{m-1}$
  for $i = 1$ **to** $j-1$
    $r_{i,j} := (\widehat{\boldsymbol{w}}, \boldsymbol{w}_i)$
    $\widehat{\boldsymbol{w}} := \widehat{\boldsymbol{w}} - r_{i,j}\boldsymbol{w}_i$
  end
  $r_{j,j} := \|\widehat{\boldsymbol{w}}\|$
  $\boldsymbol{w}_j := \widehat{\boldsymbol{w}}/r_{j,j}$
end
$\boldsymbol{y}_m^{\mathrm{MR}} := R_m^{-1} W_m^* \boldsymbol{r}_0$
$\boldsymbol{x}_m^{\mathrm{MR}} := \boldsymbol{x}_0 + [\boldsymbol{r}_0, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{m-1}]\boldsymbol{y}_m^{\mathrm{MR}}$
$\boldsymbol{r}_m^{\mathrm{MR}} := W_m W_m^* \boldsymbol{r}_0$
If desired, compute $\boldsymbol{y}_m^{\mathrm{OR}}$, $\boldsymbol{x}_m^{\mathrm{OR}}$ and $\boldsymbol{r}_m^{\mathrm{OR}}$ using (3.7) or (3.4)

also be updated from step to step.

For the GCR algorithm for Krylov spaces (cf. Algorithm 3.2.2), different ways of extending the bases C_m and W_m have been proposed in the literature. The older variant, introduced in Eisenstat et al. (1983), generates in step m the new basis vector $\boldsymbol{w}_m$ by orthogonalizing $A\boldsymbol{r}_{m-1}$ against the previously generated orthonormal basis of $\mathscr{W}_{m-1}$, while simultaneously updating $\boldsymbol{r}_{m-1}$ to obtain $\boldsymbol{c}_m = A^{-1}\boldsymbol{w}_m$. This approach, which is given in Algorithm 4.2.2, has the drawback that it fails to extend the Krylov space whenever two consecutive MR approximations coincide. For this reason, Eisenstat et al. (1983) state that the algorithm should only be used for linear systems where A is *positive real*, which means that its Hermitian part $(A + A^*)/2$ is positive definite. By Theorem 4.2.2, we see that the algorithm may in fact be used provided $0 \notin W(A)$, a slightly more general criterion.

Algorithm 4.2.2 GCR for Krylov spaces.

1 $\boldsymbol{r}_0 := \boldsymbol{b} - A\boldsymbol{x}_0$
2 **for** $j := 1$ **to** m
3 $\quad \widehat{\boldsymbol{c}} := \boldsymbol{r}_{j-1}, \widehat{\boldsymbol{w}} := A\widehat{\boldsymbol{c}}_j$
4 $\quad$ **for** $i := 1$ **to** $j-1$
5 $\quad\quad \widehat{\boldsymbol{c}} := \widehat{\boldsymbol{c}} - (\widehat{\boldsymbol{w}}, \boldsymbol{w}_i)\boldsymbol{c}_i$
6 $\quad\quad \widehat{\boldsymbol{w}} := \widehat{\boldsymbol{w}} - (\widehat{\boldsymbol{w}}, \boldsymbol{w}_i)\boldsymbol{w}_i$
7 $\quad$ **end**
8 $\quad \boldsymbol{w}_j := \widehat{\boldsymbol{w}}/\|\widehat{\boldsymbol{w}}\|$
9 $\quad \boldsymbol{c}_j := \widehat{\boldsymbol{c}}/\|\widehat{\boldsymbol{w}}\|$
10 **end**
11 $\boldsymbol{y}_m^{\mathrm{MR}} = W_m^* \boldsymbol{r}_0$
12 $\boldsymbol{x}_m^{\mathrm{MR}} := \boldsymbol{x}_0 + C_m \boldsymbol{y}_m^{\mathrm{MR}} = \boldsymbol{x}_{m-1}^{\mathrm{MR}} + (\boldsymbol{r}_0, \boldsymbol{w}_m)\boldsymbol{c}_m$
13 $\boldsymbol{r}_m^{\mathrm{MR}} := \boldsymbol{r}_0 - W_m \boldsymbol{y}_m^{\mathrm{MR}} = \boldsymbol{r}_{m-1}^{\mathrm{MR}} - (\boldsymbol{r}_0, \boldsymbol{w}_m)\boldsymbol{w}_m$
14 If desired, compute $\boldsymbol{y}_m^{\mathrm{OR}}$, $\boldsymbol{x}_m^{\mathrm{OR}}$ and $\boldsymbol{r}_m^{\mathrm{OR}}$ using (3.8)

An easy remedy for this deficiency is to extend the basis W_{m-1} by instead orthogonalizing $A\boldsymbol{w}_{m-1}$ against W_{m-1}, which results in the Arnoldi process for $\mathscr{W}_m = K_m(A, A\boldsymbol{r}_0)$ for generating the $\boldsymbol{w}_m$-sequence, while the $\boldsymbol{c}_m$-sequence is again maintained such that $A\boldsymbol{c}_m = \boldsymbol{w}_m$. The resulting algorithm is identical with Algorithm 4.2.2 except that line 3 reads $\widehat{\boldsymbol{c}}_j = \boldsymbol{w}_{j-1}, \widehat{\boldsymbol{w}}_j = A\widehat{\boldsymbol{c}}_j$. This is pointed out in Rozložník & Strakoš (1996), where many equivalent MR implementations are also compared with regard to their numerical stability. As we shall see in Section 4.4, the former GCR-variant is an ORTHOMIN-type method, while the latter is of the type ORTHODIR.

The GMRES algorithm for Krylov spaces (cf. Algorithm 3.2.3), introduced in Saad & Schultz (1986), generates the Arnoldi factorization (4.3) (cf. (2.39) and (3.9)) and a recursively updated QR-factorization of the Hessenberg matrix $\widetilde{H}_m$. As in the Walker-Zhou algorithm, only one sequence of basis vectors needs to be generated and stored. These form an orthonormal basis of $\mathscr{K}_m(A, \boldsymbol{r}_0)$. An orthonormal basis of $\mathscr{W}_m = A\mathscr{K}_m(A, \boldsymbol{r}_0)$ is implicitly available in the form of the Paige-Saunders basis (cf. Section 2.3.2). The GMRES algorithm is given in Algorithm 4.2.3. GMRES actually refers specifically to the MR approximation; when the OR approximation is computed based on the Arnoldi factorization as in Algorithm 4.2.3 it is known as the *full orthogonalization method (FOM)* (Saad 1981) or as *Arnoldi's method for solving linear systems*.

4.2.2 Krylov Subspace Methods Based on Non-Orthogonal Bases

The previous section showed how several well-known Krylov subspace algorithms for the iterative solution of (1.1) which use an orthogonal basis of $\mathscr{V}_m = \mathscr{K}_m(A, \boldsymbol{r}_0)$ result from specializing the basic algorithms of Chapters 2 and 3 to Krylov subspaces as correction spaces. In the same manner, we show in this section how Krylov subspace methods such as QMR/BCG and TFQMR/CGS, which use a non-orthogonal basis of the Krylov space, naturally fit into the framework of MR/OR approximations based on non-orthogonal bases

Algorithm 4.2.3 GMRES and FOM for Krylov spaces.

```
1  r_0 := b − Ax_0, β := ‖r_0‖, v_1 := r_0/β
2  for j := 1 to m
3      w := Av_m
4      for i = 1 to j
5          h_{i,j} := (w, v_i)
6          w := w − h_{i,j} v_i
7      end
8      h_{j+1,j} := ‖w‖
9      v_{j+1} := w/h_{j+1,j}
10 end
11 Compute y_m^MR to minimize ‖β u_1^{(m+1)} − H̃_m y‖
12 Compute y_m^OR to solve H_m y = β u_1^{(m)}
13 x_m^MR := x_0 + V_m y_m^MR
14 x_m^OR := x_0 + V_m y_m^OR
```

discussed in Section 2.3.3.

To this end, let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_L \in \mathscr{H}$ be *any* set of linearly independent vectors such that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ form a basis of $\mathscr{V}_m = \mathscr{K}_m(A, \boldsymbol{r}_0)$ for $m = 1, \ldots, L$ (in particular, $\boldsymbol{r}_0 = \beta \boldsymbol{v}_1$ for some $0 \neq \beta \in \mathbb{C}$). This basis leads naturally to a (quasi)-minimal residual as well as to a (quasi-) orthogonal residual method. Since, for every $1 \leq m \leq L$, $A\boldsymbol{v}_m \in \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{m+1}\}$, where we set $\boldsymbol{v}_{L+1} = \boldsymbol{0}$, there exists an upper Hessenberg matrix $\widetilde{H}_m \in \mathbb{C}^{(m+1)\times m}$ such that—just as in (2.39) and in (4.3)—

$$AV_m = V_m H_m + [\boldsymbol{0}, \ldots, \boldsymbol{0}, \eta_{m+1,m}\boldsymbol{v}_{m+1}] = V_{m+1}\widetilde{H}_m. \tag{4.10}$$

Just as in the Arnoldi decomposition (4.3) in the case of an orthogonal basis, the m-th column of $\widetilde{H}_m$ contains the coefficients of $A\boldsymbol{v}_m \in \mathscr{K}_{m+1}(A, \boldsymbol{r}_0)$ with respect to the basis vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m+1}$. We are thus in the situation of Section 2.3.3, with $\mathscr{V}_m = \mathscr{K}_m(A, \boldsymbol{r}_0)$ and $\mathscr{W}_m = A\mathscr{V}_m$. Defining the auxiliary inner products $(\cdot,\cdot)_V$ on $\mathscr{V}_L$ and $(\cdot,\cdot)_{\boldsymbol{v}}$ on the coordinate space $\mathbb{C}^L$ as in (2.68) and (2.67), respectively, we obtain the QMR approximation

$$\boldsymbol{x}_m^{\mathrm{QMR}} := \boldsymbol{x}_0 + V_m \boldsymbol{y}_m^{\mathrm{QMR}} = \boldsymbol{x}_0 + \boldsymbol{v}_m^{\mathrm{QMR}}$$

of the solution $A^{-1}\boldsymbol{b}$ by requiring that $A\boldsymbol{v}_m^{\mathrm{QMR}}$ be the MR approximation to $\boldsymbol{r}_0$ with respect to the inner product $(\cdot,\cdot)_V$. Just as in Section 2.3.3, the coefficient vector $\boldsymbol{y}_m^{\mathrm{QMR}} \in \mathbb{C}^m$ is characterized as the unique solution of the least-squares problem (2.44). Merely translating Theorems 2.3.9 and 2.3.10 to Krylov subspace notation, we obtain

Theorem 4.2.5. *The QMR iterates are the MR iterates with respect to the inner product* $(\cdot,\cdot)_V$*:*

$$\|\boldsymbol{b} - A\boldsymbol{x}_m^{QMR}\|_V = \min_{\boldsymbol{x}\in\boldsymbol{x}_0+\mathscr{K}_m(A,\boldsymbol{r}_0)} \|\boldsymbol{b} - A\boldsymbol{x}\|_V. \tag{4.11}$$

In terms of the original norm in $\mathscr{H}$, the MR and QMR residuals may be bounded by

$$\|\boldsymbol{r}_m^{MR}\| \leq \|\boldsymbol{r}_m^{QMR}\| \leq \sqrt{\kappa_2(M_m)}\,\|\boldsymbol{r}_m^{MR}\|, \tag{4.12}$$

in which $\kappa_2(M_m)$ denotes the (Euclidean) condition number of the (Hermitian positive definite) matrix M_m.

Theorem 4.2.6. *The residual vectors of the QMR iterates are characterized by*

$$\boldsymbol{r}_m^{QMR} \perp \mathscr{U}_m,$$

where $\mathscr{U}_m = \{\boldsymbol{v} = V_{m+1}\boldsymbol{y} \in \mathscr{K}_{m+1}(A, \boldsymbol{r}_0) \,:\, \boldsymbol{y} = M_m^{-1}\widetilde{H}_m\boldsymbol{z} \text{ with } \boldsymbol{z} \in \mathbb{C}^m\}$ is an m-dimensional subspace of $\mathscr{K}_{m+1}(A, \boldsymbol{r}_0)$ (for $m = L$, there holds $\mathscr{U}_L = \mathscr{K}_L(A, \boldsymbol{r}_0)$) and orthogonality is understood with respect to the original inner product $(\cdot,\cdot)$ on $\mathscr{H}$. Consequently,

$$\boldsymbol{r}_m^{QMR} = \left(I - P_{\mathscr{W}_m}^{\mathscr{U}_m}\right)\boldsymbol{r}_0,$$

where $P_{\mathscr{W}_m}^{\mathscr{U}_m}$ denotes the oblique projection onto $\mathscr{W}_m = A\mathscr{K}_m(A, \boldsymbol{r}_0)$ orthogonal to $\mathscr{U}_m$.

Again drawing from the abstract results in Section 2.3.3, the corresponding QOR approximations in the Krylov subspace case are defined by

$$\boldsymbol{x}_m^{\mathrm{QOR}} = \boldsymbol{x}_0 + V_m\boldsymbol{y}_m^{\mathrm{QOR}},$$

where the coordinate vector $\boldsymbol{y}_m^{\mathrm{QOR}} \in \mathbb{C}^m$ is determined by

$$0 = (\beta\boldsymbol{u}_1 - \widetilde{H}_m\boldsymbol{y}_m^{\mathrm{QOR}}, \boldsymbol{u}_j)_2 \quad (j = 1, 2, \ldots, m), \text{ i.e., } H_m\boldsymbol{y}_m^{\mathrm{QOR}} = \beta\boldsymbol{u}_1 \tag{4.13}$$

under the usual assumption that H_m is nonsingular. In terms of the inner product $(\cdot,\cdot)_V$ on $\mathscr{K}_L(A, \boldsymbol{r}_0)$ the QOR approximations are characterized by

$$\boldsymbol{r}_m^{\mathrm{QOR}} = \boldsymbol{b} - A\boldsymbol{x}_m^{\mathrm{QOR}} \perp_V \mathscr{K}_m(A, \boldsymbol{r}_0),$$

i.e., the QOR iterates are the OR iterates with respect to the inner product $(\cdot,\cdot)_V$. Similarly to Theorem 2.3.10 there holds

Theorem 4.2.7. *The residual vectors of the QOR iterates satisfy*

$$\boldsymbol{r}_m^{QOR} \perp \mathscr{T}_m,$$

where $\mathscr{T}_m = \{\boldsymbol{v} = V_{m+1}\boldsymbol{y} \in \mathscr{K}_{m+1}(A, \boldsymbol{r}_0) \,:\, \boldsymbol{y} = M_m^{-1}[\boldsymbol{z}^T\ 0]^T \text{ with } \boldsymbol{z} \in \mathbb{C}^m\}$ is an m-dimensional subspace of $\mathscr{K}_{m+1}(A, \boldsymbol{r}_0)$ (for $m = L$, there holds $\mathscr{T}_L = \mathscr{K}_L(A, \boldsymbol{r}_0)$) and orthogonality is understood with respect to the original inner product $(\cdot,\cdot)$ on $\mathscr{H}$. Consequently,

$$\boldsymbol{r}_m^{QOR} = \left(I - P_{\mathscr{W}_m}^{\mathscr{T}_m}\right)\boldsymbol{r}_0,$$

where $P_{\mathscr{W}_m}^{\mathscr{T}_m}$ denotes the oblique projection onto $\mathscr{W}_m = A\mathscr{K}_m(A, \boldsymbol{r}_0)$ orthogonal to $\mathscr{T}_m$.

The QMR and QOR iterates being identified as MR and OR iterates with respect to the inner product $(\cdot,\cdot)_V$ implies that the assertions of Corollaries 4.2.1 and 4.2.4 are valid for any pair of QMR/QOR methods. The crucial angles φ_m of (4.8), however, are now defined with respect to $(\cdot,\cdot)_V$.

The motivation for using non-orthogonal bases comes from the potential savings in storage and computation when using algorithms for generating a basis of $\mathscr{K}_m(A,\boldsymbol{r}_0)$ which lead to a Hessenberg matrix in relation (4.10) with a small number of bands. A tridiagonal Hessenberg matrix can be achieved using the non-Hermitian Lanczos process. The non-Hermitian variant of the Lanczos process is almost as economical as its Hermitian counterpart, requiring in addition the generation of a basis of a Krylov space with respect to the adjoint A^* and thus storage for several additional vectors and, what can be expensive, multiplication by A^* at each step. The result are two biorthogonal, in general not orthogonal, bases. The non-Hermitian Lanczos process may break down before the Krylov space becomes stationary, and in finite arithmetic this leads to numerical instabilities in the case on near-breakdowns. This problem is addressed by so-called look-ahead techniques for the non-Hermitian Lanczos process (cf. Freund, Gutknecht & Nachtigal (1993), Gutknecht (1997) and the references therein), which result in a pair of block-biorthogonal bases and a Hessenberg matrix which is as close to tridiagonal form as possible while maintaining stability.

Remark 4.2.8. We have mentioned that Krylov subspace QMR/QOR methods have the advantage of being able to work with an arbitrary basis of the Krylov space, in particular also with a non-orthogonal basis. As the bounds (4.12) show, how close the QMR iterates come to also minimizing the residual in the original norm depends on the Euclidean condition number of the Gram matrices M_m, i.e., on the largest and smallest singular value of $V_m = [\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m]$ (interpreted as an operator from $\mathbb{C}^m$ to $\mathscr{K}_m(A,\boldsymbol{r}_0)$). The largest singular value of V_m is bounded by $\sqrt{m}$ if the basis vectors are normalized with respect to the original norm. When the look-ahead Lanczos algorithm is used to generate the basis, bounds on the smallest singular value may be obtained depending on the look-ahead strategy being followed.

Example: QMR/BCG and TFQMR/CGS

We now consider two particular examples of Krylov subspace QMR/QOR algorithms, the QMR method of Freund & Nachtigal (1991) and the TFQMR and CGS methods due to Freund (Freund 1993, Freund 1994) and Sonneveld (1989), respectively. As another QMR/QOR pair, we mention the QMRBICGSTAB method of Chan, Gallopoulos, Simoncini, Szeto & Tong (1994) and BICGSTAB developed by van der Vorst (1992) and Gutknecht (1993*b*).

The QMR algorithm of Freund and Nachtigal proceeds exactly as described above, with the basis of the Krylov space generated by the look-ahead Lanczos algorithm. The QOR counterpart of Freund and Nachtigal's QMR is the BCG algorithm, the iterates of which are characterized by

$$\boldsymbol{x}^{\mathrm{BCG}} \in \boldsymbol{x}_0 + \mathscr{K}_m(A,\boldsymbol{r}_0), \qquad \boldsymbol{r}_m^{\mathrm{BCG}} \perp \mathscr{K}_m(A^*,\widetilde{\boldsymbol{r}}_0).$$

The algorithm proceeds by generating a basis of the Krylov space $\mathscr{K}_m(A^*,\widetilde{\boldsymbol{r}}_0)$, where $\widetilde{\boldsymbol{r}}_0$ is an arbitrary starting vector, simultaneously with a basis V_m of $\mathscr{K}_m(A,\boldsymbol{r}_0)$ in such a way

that these two bases are biorthogonal. If $\boldsymbol{y} \in \mathbb{C}^m$ denotes the coefficient vector of the BCG approximation with respect to V_m, then the biorthogonality requirement implies

$$\begin{aligned} \boldsymbol{r}_m^{\mathrm{BCG}} &= V_{m+1}(\beta \boldsymbol{u}_1 - \widetilde{H}_m \boldsymbol{y}) \perp \mathscr{K}_m(A^*, \widetilde{\boldsymbol{r}}_0) \\ &\Leftrightarrow \boldsymbol{r}_m^{\mathrm{BCG}} \perp_V \mathscr{V}_m \\ &\Leftrightarrow H_m \boldsymbol{y} = \beta \boldsymbol{u}_1. \end{aligned}$$

The last equality identifies $\boldsymbol{x}^{\mathrm{BCG}}$ as the m-th QOR iterate.

To treat the TFQMR/CGS pair, recall that the residual of any Krylov subspace approximation can be expressed as $\boldsymbol{r}_m = \phi_m(A)\boldsymbol{r}_0$ in terms of a polynomial ϕ_m of degree m satisfying $\phi_m(0) = 1$. Sonneveld (1989) defined the CGS iterate $\boldsymbol{x}_m^{\mathrm{CGS}} \in \boldsymbol{x}_0 + \mathscr{K}_{2m}(A, \boldsymbol{r}_0)$ such that

$$\boldsymbol{r}_m^{\mathrm{CGS}} = [\phi_m(A)]^2 \boldsymbol{r}_0, \quad \text{where} \quad \boldsymbol{r}_m^{\mathrm{BCG}} = \phi_m(A)\boldsymbol{r}_0.$$

It is shown in Freund (1993) that $\boldsymbol{x}_m^{\mathrm{CGS}} = \boldsymbol{x}_0 + Y_{2m}\boldsymbol{z}_{2m}$ for a coefficient vector $\boldsymbol{z}_{2m} \in \mathbb{C}^{2m}$, where $\mathscr{K}_{2m}(A, \boldsymbol{r}_0) = \mathrm{span}\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{2m}\}$. Moreover, there exists a sequence $\{\boldsymbol{w}_n\}$ such that $\boldsymbol{r}_0 = \beta \boldsymbol{w}_1$ and

$$AY_n = W_{n+1}\widetilde{H}_n, \qquad n = 1, \ldots, 2L.$$

In terms of this sequence, there holds

$$\boldsymbol{r}_m^{\mathrm{CGS}} = W_{2m+1}\left(\beta \boldsymbol{u}_1 - \widetilde{H}_{2m}\boldsymbol{z}_{2m}\right),$$

where $H_{2m}\boldsymbol{z}_{2m} = \beta \boldsymbol{u}_1$, i.e.,

$$\boldsymbol{r}_m^{\mathrm{CGS}} \perp_W \mathrm{span}\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{2m}\} = A\mathscr{K}_{2m}(A, \boldsymbol{r}_0),$$

which identifies CGS as an OR method. The corresponding MR method is Freund's transpose-free QMR method TFQMR (cf. Freund (1993),Freund (1994)), the iterates of which are defined as

$$\boldsymbol{x}_n^{\mathrm{TFQMR}} = \boldsymbol{x}_0 + Y_n \boldsymbol{z}_n, \quad n = 1, \ldots, 2L,$$

where the coefficient vector $\boldsymbol{z}_n \in \mathbb{C}^n$ solves the least-squares problem

$$\|\beta \boldsymbol{u}_1 - \widetilde{H}_n \boldsymbol{z}_n\|_2 \to \min_{\boldsymbol{z} \in \mathbb{C}^n}.$$

In other words,

$$\|\boldsymbol{r}_n^{\mathrm{TFQMR}}\|_W = \min_{\boldsymbol{x} \in \boldsymbol{x}_0 + \mathscr{K}_n(A, \boldsymbol{r}_0)} \|\boldsymbol{b} - A\boldsymbol{x}\|_W.$$

Comparison of Residuals

The availability of a sequence of vectors $\{\widetilde{\boldsymbol{v}}_j\}_{j \geq 1}$ which is biorthogonal to the Krylov basis $\{\boldsymbol{v}_j\}$ permits a convenient representation of the QOR residual via (2.48). If

$$(\boldsymbol{v}_j, \widetilde{\boldsymbol{v}}_k) = \delta_{jk} d_j, \qquad j, k = 1, \ldots, L,$$

and if both sequences are normalized to one, then, since the MR-residual at step $m-1$ lies in $\mathscr{W}_{m-1} \subset \mathscr{V}_m$, we have $\boldsymbol{r}_{m-1}^{\mathrm{MR}} = \sum_{j=1}^{m} (\boldsymbol{r}_{m-1}^{\mathrm{MR}}, \widetilde{\boldsymbol{v}}_j/d_j)\boldsymbol{v}_j$. Inserting $\boldsymbol{r} - \boldsymbol{w} = \boldsymbol{r}_{m-1}^{\mathrm{MR}}$ in (2.48) then yields

$$\boldsymbol{r}_m^{\mathrm{QOR}} = (\boldsymbol{r}_{m-1}^{\mathrm{MR}}, \widehat{\boldsymbol{w}}_{m+1})_V \boldsymbol{v}_{m+1} = \boldsymbol{v}_{m+1} \sum_{j=1}^{m} \frac{\overline{g}_j}{\overline{d}_j} (\boldsymbol{r}_{m-1}^{\mathrm{MR}}, \widetilde{\boldsymbol{v}}_j)\boldsymbol{v}_j,$$

which, after taking norms and applying the Cauchy-Schwarz inequality, becomes

$$\|\boldsymbol{r}_m^{\mathrm{QOR}}\| \leq \|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\| \sum_{j=1}^{m} \frac{|g_j|}{|d_j|}.$$

This bound is a slightly improved version of a bound by Hochbruck & Lubich (1998).

4.3 The Polynomial Structure

As easily verified, the m-dimensional Krylov space with respect to the operator A and initial vector $\boldsymbol{r}_0$ may be represented as

$$\mathscr{K}_m(A, \boldsymbol{r}_0) = \{q(A)\boldsymbol{r}_0 \,:\, q \in \mathscr{P}_{m-1}\} \qquad m = 1, 2, \ldots ,$$

where $\mathscr{P}_m$ denotes the space of all complex polynomials of degree at most m. The linear map

$$\mathscr{P}_{m-1} \ni q \mapsto q(A)\boldsymbol{r}_0 \in \mathscr{K}_m(A, \boldsymbol{r}_0)$$

is thus always surjective but fails to be an isomorphism if and only if there exists a nonzero polynomial $q \in \mathscr{P}_{m-1}$ with $q(A)\boldsymbol{r}_0 = \boldsymbol{0}$. If such a polynomial exists (e.g., if A has finite rank) then there also exists a (unique) monic polynomial $c = c_{A,\boldsymbol{r}_0}$ of minimal degree with $c(A)\boldsymbol{r}_0 = \boldsymbol{0}$ which is usually called the *minimal polynomial of* $\boldsymbol{r}_0$ *with respect to* A. It is easy to see that the degree of c equals the smallest integer m such that $\mathscr{K}_m = \mathscr{K}_{m+1}$ and thus coincides with the index L introduced in (2.40) (cf. also (3.3) and (4.5)),

$$\begin{aligned} L &= \min\{m \in \mathbb{N}_0 \,:\, \mathscr{K}_m = \mathscr{K}_{m+1}\} = \min\{m \in \mathbb{N}_0 \,:\, A^{-1}\boldsymbol{r}_0 \in \mathscr{K}_m\} \\ &= \min\{\deg q \,:\, q \text{ monic and } q(A)\boldsymbol{r}_0 = \boldsymbol{0}\}. \end{aligned} \tag{4.14}$$

In other words, $\mathscr{P}_{m-1}$ and $\mathscr{K}_m$ are isomorphic linear spaces if and only if $m \leq L$.

The positive semidefinite sesquilinear form

$$(p, q) := (p(A)\boldsymbol{r}_0, q(A)\boldsymbol{r}_0) \qquad (p, q \in \mathscr{P}_\infty := \cup_{m\geq 0}\mathscr{P}_m) \tag{4.15}$$

is therefore positive definite when restricted to $\mathscr{P}_{L-1}$ and thus defines an inner product on this space. We will use the same notation $(\cdot, \cdot)$ for this inner product as for its counterpart on $\mathscr{H}$, as well as for derived quantities such as its induced norm $\|\cdot\| := (\cdot, \cdot)^{1/2}$ and the orthogonality relation $\perp$.

Since every vector $\boldsymbol{x} \in \boldsymbol{x}_0 + \mathscr{K}_m$ is of the form $\boldsymbol{x} = \boldsymbol{x}_0 + q_{m-1}(A)\boldsymbol{r}_0$ for some $q_{m-1} \in \mathscr{P}_{m-1}$, the corresponding residual $\boldsymbol{r} = \boldsymbol{b} - A\boldsymbol{x}$ can be written as

$$\boldsymbol{r} = \boldsymbol{r}_0 - Aq_{m-1}(A)\boldsymbol{r}_0 = p_m(A)\boldsymbol{r}_0, \quad \text{where } p_m(\zeta) := 1 - \zeta q_{m-1}(\zeta) \in \mathscr{P}_m.$$

Note that the *residual polynomial* p_m satisfies the normalization condition $p_m(0) = 1$. Later in this section we will characterize the residual polynomials which belong to the OR and MR iterates as well as their zeros.

First, however, we provide three lemmata for later use. The first recalls a well-known (see, e.g., Paige, Parlett & van der Vorst (1995)) consequence of the Arnoldi decomposition $AV_m = V_m H_m + \eta_{m+1,m} \boldsymbol{v}_{m+1} \boldsymbol{u}_m^\top$ of A (see (4.3)), the second states the conditions under which a Krylov space can have A-invariant subspaces. The third lemma shows that the orthogonal complement of a Krylov space with respect to an A-invariant subspace is itself a Krylov space.

Lemma 4.3.1. *For every polynomial $q(\zeta) = \alpha_m \zeta^m + \cdots + \alpha_1 \zeta + \alpha_0 \in \mathscr{P}_m$, there holds*

$$q(A)\boldsymbol{r}_0 = \beta\, V_m q(H_m)\boldsymbol{u}_1 + \alpha_m\, \beta \prod_{j=1}^{m} \eta_{j+1,j}\, \boldsymbol{v}_{m+1},$$

where $\boldsymbol{u}_1 \in \mathbb{C}^m$ denotes the first unit vector. In particular, $q(A)\boldsymbol{r}_0 = \beta\, V_m q(H_m)\boldsymbol{u}_1$ for every $q \in \mathscr{P}_{m-1}$.

Proof. It is sufficient to prove the assertion for the monomials ζ^k $(k = 0, 1, \ldots, m)$. We have

$$\begin{aligned} \boldsymbol{r}_0 &= \beta\, V_m \boldsymbol{u}_1, \\ A\boldsymbol{r}_0 &= \beta A V_m \boldsymbol{u}_1 = \beta\, V_m H_m \boldsymbol{u}_1 + \beta\, \eta_{m+1,m} \boldsymbol{v}_{m+1} \boldsymbol{u}_m^\top \boldsymbol{u}_1 = \beta\, V_m H_m \boldsymbol{u}_1 \end{aligned}$$

and by induction, for $k < m$,

$$\begin{aligned} A^k \boldsymbol{r}_0 &= A A^{k-1} \boldsymbol{r}_0 = \beta A V_m H_m^{k-1} \boldsymbol{u}_1 \\ &= \beta\, V_m H_m^k \boldsymbol{u}_1 + \beta\, \eta_{m+1,m} \boldsymbol{v}_{m+1} \boldsymbol{u}_m^\top H_m^{k-1} \boldsymbol{u}_1 = \beta\, V_m H_m^k \boldsymbol{u}_1. \end{aligned}$$

The last equality is a consequence of the upper Hessenberg structure of H_m which implies that $H_m^{k-1}\boldsymbol{u}_1$ lies in the span of the first k unit vectors, i.e., $\boldsymbol{u}_m^\top H_m^{k-1} \boldsymbol{u}_1 = 0$ if $k < m$. Finally, for $k = m$, there follows

$$\begin{aligned} A^m \boldsymbol{r}_0 &= \beta A V_m H_m^{m-1} \boldsymbol{u}_1 = \beta\, V_m H_m^m \boldsymbol{u}_1 + \beta\, \eta_{m+1,m} \boldsymbol{v}_{m+1} \boldsymbol{u}_m^\top H_m^{m-1} \boldsymbol{u}_1 \\ &= \beta\, V_m H_m^m \boldsymbol{u}_1 + \beta\, \eta_{m,1}^{(m-1)}\, \eta_{m+1,m} \boldsymbol{v}_{m+1}, \end{aligned}$$

where $\eta_{m,1}^{(m-1)}$ denotes the $(m,1)$-entry of H_m^{m-1} which is the product of the subdiagonal entries of H_m. □

Lemma 4.3.2. *A Krylov space $\mathscr{K}_m(A, \boldsymbol{r}_0)$ contains an A-invariant subspace if and only if it is itself A-invariant.*

Proof. If $\mathscr{U} \subset \mathscr{K}_m(A, \boldsymbol{r}_0)$ is A-invariant, it must contain an eigenvector $\boldsymbol{z}$ of A. As an element of $\mathscr{K}_m$, $\boldsymbol{z}$ has a representation $\boldsymbol{z} = q_{m-1}(A)\boldsymbol{r}_0$ in terms of a nonzero polynomial q_{m-1} of degree at most $m-1$. Moreover, if λ denotes the eigenvalue of A associated with $\boldsymbol{z}$ and $p(\zeta) := (\zeta - \lambda) q_{m-1}(\zeta)$, then $p(A)\boldsymbol{r}_0 = \boldsymbol{0}$ and hence the degree of the minimal polynomial $c_{A,\boldsymbol{r}_0}$ of $\boldsymbol{r}_0$ with respect to A is at most m. Consequently $L = \deg c_{A,\boldsymbol{r}_0} \leq m$ and $\mathscr{K}_m$ is A-invariant (cf. (4.14)). □

Lemma 4.3.3. *Let $\mathscr{U}$ be an A-invariant subspace, $\mathscr{T} = \mathscr{U}^\perp$ its orthogonal complement and set $A_\mathscr{T} := P_\mathscr{T} A P_\mathscr{T}$. Then there holds for $m = 1, 2, \ldots$*

$$P_\mathscr{T} \mathscr{K}_m(A, \boldsymbol{r}_0) = \mathscr{K}_m(P_\mathscr{T} A, P_\mathscr{T} \boldsymbol{r}_0) = \mathscr{K}_m(A_\mathscr{T}, P_\mathscr{T} \boldsymbol{r}_0) \qquad \text{and}$$
$$P_\mathscr{T} A \mathscr{K}_m(A, \boldsymbol{r}_0) = P_\mathscr{T} A \mathscr{K}_m(P_\mathscr{T} A, P_\mathscr{T} \boldsymbol{r}_0) = A_\mathscr{T} \mathscr{K}_m(A_\mathscr{T}, P_\mathscr{T} \boldsymbol{r}_0).$$

Proof. We have $P_\mathscr{T} A P_\mathscr{U} = O$, because $\mathscr{U}$ is A-invariant, and therefore

$$P_\mathscr{T} A = P_\mathscr{T} A P_\mathscr{U} + P_\mathscr{T} A P_\mathscr{T} = P_\mathscr{T} A P_\mathscr{T}.$$

Now an obvious induction shows that for $k = 1, 2, \ldots$

$$P_\mathscr{T} A^k \boldsymbol{r}_0 = [P_\mathscr{T} A]^k \boldsymbol{r}_0 = [P_\mathscr{T} A P_\mathscr{T}]^k \boldsymbol{r}_0$$

which proves the assertions. □

Concerning the notation used in Lemma 4.3.3 we remark that so far $A_\mathscr{T}$ has denoted the orthogonal section $P_\mathscr{T} A|_\mathscr{T}$ of A onto $\mathscr{T}$. We henceforth identify $P_\mathscr{T} A P_\mathscr{T}$ with $A_\mathscr{T}$ since $P_\mathscr{T} A P_\mathscr{T} = P_\mathscr{T} A|_\mathscr{T}$ on $\mathscr{T}$ and $P_\mathscr{T} A P_\mathscr{T} = O$ on $\mathscr{T}^\perp$.

4.3.1 OR Residual Polynomials

We first investigate the residual polynomials p_m^{OR} associated with the OR approach: $\boldsymbol{r}_m^{\mathrm{OR}} = p_m^{\mathrm{OR}}(A)\boldsymbol{r}_0$. The condition $\boldsymbol{r}_m^{\mathrm{OR}} \perp \mathscr{K}_m$ translates to $p_m^{\mathrm{OR}} \perp \mathscr{P}_{m-1}$, i.e., p_m^{OR} is an orthogonal polynomial of degree m, normalized to satisfy $p_m^{\mathrm{OR}}(0) = 1$. This also follows from the fact that $\boldsymbol{r}_m^{\mathrm{OR}}$ is a scalar multiple of $\boldsymbol{v}_{m+1}$, the last element of the orthonormal basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m, \boldsymbol{v}_{m+1}\}$ of $\mathscr{K}_{m+1}$ (cf. Remark 2.3.3): The basis vector $\boldsymbol{v}_{m+1} \in \mathscr{K}_{m+1} \setminus \mathscr{K}_m$ has the form $\boldsymbol{v}_{m+1} = v_m(A)\boldsymbol{r}_0$ for some polynomial v_m of *exact* degree m, and p_m^{OR} must be a scalar multiple of v_m. Next, $\boldsymbol{v}_{m+1} \perp \mathscr{K}_m$, i.e., $v_m \perp \mathscr{P}_{m-1}$, and $\|\boldsymbol{v}_{m+1}\| = \|v_m\| = 1$ show that v_m is an orthogonal polynomial of degree m normalized with respect to the inner product (4.15). The OR residual polynomial is $p_m^{\mathrm{OR}} = v_m / v_m(0)$, a normalization which is of course only possible if v_m does not vanish at the origin.

The close relation of v_m to the characteristic polynomial of the Hessenberg matrix H_m will show that $v_m(0) = 0$ is equivalent to H_m being singular: We know that $\boldsymbol{v}_{m+1} = v_m(A)\boldsymbol{r}_0$ spans the one-dimensional space $\mathscr{K}_{m+1} \cap \mathscr{K}_m^\perp$. If, on the other hand, $h_m(\zeta) := \det(\zeta I - H_m) \in \mathscr{P}_m$ denotes the characteristic polynomial of H_m, then by Lemma 4.3.1 and the Cayley-Hamilton theorem

$$\big(h_m(A)\boldsymbol{r}_0, \boldsymbol{v}_k\big) = \beta \left(V_m h_m(H_m)\boldsymbol{u}_1, \boldsymbol{v}_k\right) + \beta \prod_{j=1}^{m} \eta_{j+1,j} \left(\boldsymbol{v}_{m+1}, \boldsymbol{v}_k\right) = 0 \tag{4.16}$$

$(k = 1, 2, \ldots, m)$. In other words, $\boldsymbol{v} = h_m(A)\boldsymbol{r}_0$ belongs to $\mathscr{K}_{m+1} \cap \mathscr{K}_m^\perp$ and is therefore a scalar multiple of $\boldsymbol{v}_{m+1}$. We have thus shown that the polynomials v_m and h_m can differ only by a scalar factor. We summarize these observations in

Proposition 4.3.4. *The characteristic polynomial h_m of the Hessenberg matrix H_m is the (unique) monic orthogonal polynomial of degree m with respect to the inner product* (4.15)*. The mth OR iterate exists if and only if $h_m(0) \neq 0$ and, in this case, the corresponding residual polynomial is given by $p_m^{OR} = h_m / h_m(0)$.*

We next consider the zeros of p_m^{OR} or, equivalently, the eigenvalues of the matrix $H_m = [(A\boldsymbol{v}_j, \boldsymbol{v}_k)]_{1\le j,k\le m}$, the orthogonal section $A_{\mathscr{K}_m}$ of A onto $\mathscr{K}_m$. Its eigenvalues θ_j,

$$H_m \boldsymbol{y}_j = \theta_j \boldsymbol{y}_j \qquad \text{with } \boldsymbol{y}_j \in \mathbb{C}^m, \|\boldsymbol{y}_j\|_2 = 1, \tag{4.17}$$

are called the *Ritz values* of A (with respect to $\mathscr{K}_m$), while $\boldsymbol{z}_j := V_m \boldsymbol{y}_j$ are the associated *Ritz vectors*.

As the eigenvalues of the nonderogatory matrix H_m, Ritz values have geometric multiplicity one. In case θ_j has algebraic multiplicity $k_j > 1$, we denote by $\boldsymbol{y}_j^{(0)} = \boldsymbol{y}_j, \boldsymbol{y}_j^{(1)}, \ldots, \boldsymbol{y}_j^{(k_j-1)}$ the principal vectors of H_m which belong to the eigenvalue θ_j, so that

$$H_m \boldsymbol{y}_j^{(\ell)} = \theta_j \boldsymbol{y}_j^{(\ell)} + \boldsymbol{y}_j^{(\ell-1)} \qquad (\ell = 1, \ldots, k_j - 1)$$

and define $\boldsymbol{z}_j^{(0)} := V_m \boldsymbol{y}_j^{(0)}$ *and* $\boldsymbol{z}_j^{(\ell)} := V_m \boldsymbol{y}_j^{(\ell)}$ as the associated Ritz vectors.

Although all our conclusions remain valid in this more general case, we will assume in the remaining sections that H_m has m distinct eigenvalues to avoid the (notational) complication of introducing principal vectors.

The Ritz vectors constitute a basis of $\mathscr{K}_m$, and their residual vectors with regard to the eigenvalue problem (4.17) are given by

$$\begin{aligned} A\boldsymbol{z}_j - \theta_j \boldsymbol{z}_j &= AV_m \boldsymbol{y}_j - \theta_j V_m \boldsymbol{y}_j \\ &= V_m H_m \boldsymbol{y}_j + \eta_{m+1,m} \boldsymbol{v}_{m+1} \boldsymbol{u}_m^\top \boldsymbol{y}_j - \theta_j V_m \boldsymbol{y}_j \\ &= \eta_{m+1} (\boldsymbol{u}_m^\top \boldsymbol{y}_j) \boldsymbol{v}_{m+1}. \end{aligned} \tag{4.18}$$

This implies $A\boldsymbol{z}_j - \theta_j \boldsymbol{z}_j \perp \mathscr{K}_m$, which is the commonly used definition of Ritz values and Ritz vectors. We also see that $(A - \theta_j I)\boldsymbol{z}_j \in \operatorname{span}\{\boldsymbol{v}_{m+1}\} = \operatorname{span}\{h_m(A)\boldsymbol{r}_0\}$ for every eigenvalue θ_j of H_m. As an element of $\mathscr{K}_m$, each Ritz vector $\boldsymbol{z}_j$ can be represented as $\boldsymbol{z}_j = z_j(A)\boldsymbol{r}_0$ with a polynomial $z_j \in \mathscr{P}_{m-1}$. Equation (4.18) now implies $(\zeta - \theta_j) z_j(\zeta) = \tau_j h_m(\zeta)$ with $\tau_j \in \mathbb{C} \setminus \{0\}$, which we express as

$$z_j(\zeta) = \tau_j \frac{h(\zeta)}{\zeta - \theta_j}.$$

Proposition 4.3.5. *Let*

$$h_m(\zeta) = \prod_{j=1}^{J} (\zeta - \theta_j)^{k_j} \qquad (\theta_i \neq \theta_j \text{ for } i \neq j)$$

denote the characteristic polynomial of H_m. The Ritz vectors $\boldsymbol{z}_j^{(\ell)}$ $(\ell = 0, \ldots, k_j - 1)$ of A with respect to $\mathscr{K}_m(A, \boldsymbol{r}_0)$ have the form

$$\boldsymbol{z}_j^{(\ell)} = z_j^{(\ell)}(A)\boldsymbol{r}_0, \text{ where } z_j^{(\ell)}(\zeta) = h_m(\zeta) \sum_{i=0}^{\ell} \frac{\tau_{j,i}}{(\zeta - \theta_j)^{i+1}}$$

is a polynomial of exact degree $m - 1$. Moreover, there holds

$$(A - \theta_j I)^{\ell+1} \boldsymbol{z}_j^{(\ell)} \in \operatorname{span}\{\boldsymbol{v}_{m+1}\} = \operatorname{span}\{h_m(A)\boldsymbol{r}_0\} = \operatorname{span}\{\boldsymbol{r}_m^{OR}\},$$

where the last equality assumes that the m-th OR iterate is defined.

Remark 4.3.6. We note that there is also a best approximation property associated with the monic orthogonal polynomials $\{h_m\}$: If we write $h_m(\zeta) = \zeta^m - q_{m-1}(\zeta)$ with $q_{m-1} \in \mathscr{P}_{m-1}$, then $h_m \perp \mathscr{P}_{m-1}$ implies

$$A^m \boldsymbol{r}_0 - q_{m-1}(A)\boldsymbol{r}_0 \perp \mathscr{K}_m(A, \boldsymbol{r}_0) = A^{-1}\mathscr{K}_m(A^{-1}, A^m \boldsymbol{r}_0),$$

i.e., $q_{m-1}(A)\boldsymbol{r}_0$ is the best approximation to $A^m\boldsymbol{r}_0$ from the space $\mathscr{K}_m(A^{-1}, A^m\boldsymbol{r}_0)$, which reveals a connection between m OR steps and m MR steps with A and $\boldsymbol{r}_0$ replaced by A^{-1} and A^m, respectively, which we shall make use of in the next section to determine the zeros of the MR residual polynomials.

4.3.2 MR Residual Polynomials

We now turn to the investigation of the residual polynomials p_m^{MR} associated with the MR residuals $\boldsymbol{r}_m^{\mathrm{MR}} = p_m^{\mathrm{MR}}(A)\boldsymbol{r}_0$. Obviously, these polynomials possess the following minimization property:

$$\|p_m^{\mathrm{MR}}\| = \min\{\|p\| \,:\, \deg p \le m,\, p(0) = 1\}.$$

The condition $\boldsymbol{r}_m^{\mathrm{MR}} \perp A\mathscr{K}_m$ translates into $p_m^{\mathrm{MR}} \perp \zeta\mathscr{P}_{m-1}$ from which we first deduce the *reproducing property* (4.19) below of the MR residual polynomials: For any $q(\zeta) = q(0) + \sum_{j=1}^m \alpha_j\zeta^j \in \mathscr{P}_m$, there holds

$$(q, p_m^{\mathrm{MR}}) = (q(0), p_m^{\mathrm{MR}}) + \left(\zeta\sum_{j=1}^m \alpha_j\zeta^{j-1}, p_m^{\mathrm{MR}}\right) = q(0)\,(1, p_m^{\mathrm{MR}}),$$

and, because this identity is valid for $q = p_m^{\mathrm{MR}}$, i.e., $\|p_m^{\mathrm{MR}}\|^2 = p_m^{\mathrm{MR}}(0)(1, p_m^{\mathrm{MR}}) = (1, p_m^{\mathrm{MR}})$,

$$(q, p_m^{\mathrm{MR}}) = q(0)\,\|p_m^{\mathrm{MR}}\|^2 \qquad \text{for all } q \in \mathscr{P}_m. \tag{4.19}$$

(Had we normalized the polynomials p_m^{MR} by $\|p_m^{\mathrm{MR}}\| = 1$ instead, this reproducing property would have the more familiar form $(q, p_m^{\mathrm{MR}}) = q(0)$ for all q.) The (Fourier) coefficients of p_m^{MR} with respect to the orthonormal basis $\{v_0, v_1, \ldots, v_m\}$ of $\mathscr{P}_m$ are thus given by

$$(p_m^{\mathrm{MR}}, v_j) = \overline{v_j(0)}\,\|p_m^{\mathrm{MR}}\|^2$$

which, in view of $\|p_m^{\mathrm{MR}}\|^2 = \sum_{j=0}^m |v_j(0)|^2\|p_m^{\mathrm{MR}}\|^4$, leads to the expansion

$$p_m^{\mathrm{MR}}(\zeta) = \frac{\sum_{j=0}^m \overline{v_j(0)}\, v_j(\zeta)}{\sum_{j=0}^m |v_j(0)|^2}. \tag{4.20}$$

(note that the denominator $\sum_{j=0}^m |v_j(0)|^2 \ge |v_0(0)|^2$ is always positive since v_0 is a nonzero constant). Moreover, if the OR residual polynomials $p_j^{\mathrm{OR}} = v_j/v_j(0)$ exist for $j = 0, 1, \ldots, m$, we have, since $\|p_j^{\mathrm{OR}}\| = 1/|v_j(0)|$,

$$\overline{v_j(0)}v_j(\zeta) = |v_j(0)|^2 p_j^{\mathrm{OR}}(\zeta) = \frac{p_j^{\mathrm{OR}}(\zeta)}{\|p_j^{\mathrm{OR}}\|^2}.$$

Combined with (4.20), this yields

$$\frac{p_m^{\mathrm{MR}}}{\|p_m^{\mathrm{MR}}\|^2} = \sum_{j=1}^{m} \frac{p_j^{\mathrm{OR}}}{\|p_j^{\mathrm{OR}}\|^2},$$

which is the polynomial analogue of (2.33).

Furthermore, the representation (4.20) shows that, since the polynomials v_j are of exact degree j, p_m^{MR} will have degree m if and only if $v_m(0) \neq 0$, i.e., if the OR polynomial of degree m exists. Otherwise $p_m^{\mathrm{MR}} = p_{m-1}^{\mathrm{MR}} = \cdots = p_k^{\mathrm{MR}}$ and $\deg p_m^{\mathrm{MR}} = k$, if k is the largest index less than m for which $v_k(0) \neq 0$. The MR residual polynomial p_m^{MR} coinciding with p_{m-1}^{MR} or the orthogonal polynomial v_m vanishing at the origin are two further equivalent characterizations of a Galerkin breakdown in step m of the Krylov subspace OR iteration.

To characterize the zeros of the MR residual polynomials in an analogous manner to those of the OR residual polynomials, we begin by identifying them as the eigenvalues of an orthogonal section of A^{-1} onto the Krylov space

$$\mathscr{K}_m(A^{-1}, A^m \boldsymbol{r}_0) = \operatorname{span}\{A^m \boldsymbol{r}_0, A^{m-1}\boldsymbol{r}_0, \ldots, A\boldsymbol{r}_0\} = A\mathscr{K}_m(A, \boldsymbol{r}_0).$$

We denote the associated Arnoldi decomposition by

$$A^{-1}W_m = W_{m+1}\widetilde{G}_m = W_m G_m + \gamma_{m+1,m}\boldsymbol{w}_{m+1}\boldsymbol{u}_m^{\top},$$

in which $\widetilde{G}_m = [\gamma_{j,k}] \in \mathbb{C}^{(m+1)\times m}$ is an upper Hessenberg matrix, G_m the associated square Hessenberg matrix obtained by deleting the last row of $\widetilde{G}_m$ and $W_{m+1} = \begin{bmatrix}\boldsymbol{w}_1 & \ldots & \boldsymbol{w}_{m+1}\end{bmatrix}$ is an orthonormal basis of $\mathscr{K}_{m+1}(A^{-1}, A^m\boldsymbol{r}_0) = \mathscr{K}_{m+1}(A, \boldsymbol{r}_0)$. If we invoke Lemma 4.3.1 applied to this Arnoldi decomposition, we obtain

$$q(A^{-1})A^m\boldsymbol{r}_0 = W_m q(G_m)\beta_m\boldsymbol{u}_1^{\top} + \alpha_m\beta_m\prod_{j=1}^{m}\gamma_{j+1,j}\boldsymbol{w}_{m+1}$$

for any polynomial $q(\zeta) = \alpha_m\zeta^m + \cdots + \alpha_1\zeta + \alpha_0 \in \mathscr{P}_m$, where $\beta_m = \|A^m\boldsymbol{r}_0\|$. Denoting by g_m the characteristic polynomial of G_m, we conclude just as in (4.16) that

$$(g_m(A^{-1})A^m\boldsymbol{r}_0, \boldsymbol{w}_k) = 0, \qquad k = 1, \ldots, m$$

and that $\boldsymbol{w} := g_m(A^{-1})A^m\boldsymbol{r}_0$ belongs to

$$\mathscr{K}_{m+1}(A^{-1}, A^m\boldsymbol{r}_0) \cap \mathscr{K}_m(A^{-1}, A^m\boldsymbol{r}_0)^{\perp} = \mathscr{K}_{m+1}(A, \boldsymbol{r}_0) \cap (A\mathscr{K}_m(A, \boldsymbol{r}_0))^{\perp}.$$

By virtue of its inclusion in the latter space, we conclude that the vector $\boldsymbol{w}$ is a scalar multiple of the MR residual vector $\boldsymbol{r}_m^{\mathrm{MR}}$. Moreover, we observe that $\widehat{g}_m(\zeta) := g_m(\zeta^{-1})\zeta^m$ is a polynomial in ζ of degree at most m, which is sometimes denoted as the *reversed polynomial* of g_m since it is obtained from g_m by reversing the order of the coefficients. Since $\boldsymbol{w} = \widehat{g}_m(A)\boldsymbol{r}_0$ and $\boldsymbol{r}_m^{\mathrm{MR}} = p_m^{\mathrm{MR}}(A)\boldsymbol{r}_0$ are collinear, the same is true for the associated polynomials. Furthermore, since the characteristic polynomial g_m is monic, it follows that $\widehat{g}_m$ has value one at zero, and therefore that $\widehat{g}_m$ coincides with p_m^{MR}. The desired zeros of p_m^{MR} thus coincide with those of $\widehat{g}_m$, which are easily seen to be the reciprocals of the zeros of g_m, which in turn are the eigenvalues of G_m. Since this matrix is not readily

available, we instead derive a matrix which is similar to G_m and therefore has the same characteristic polynomial.

Departing from $AV_m = \widehat{V}_m R_m$ (cf. (3.11)), where $\widehat{V}_m$ denotes the Paige and Saunders basis of $A\mathscr{K}_m(A, \boldsymbol{r}_0)$ and R_m is the triangular factor from the QR-factorization of $\widetilde{H}_m$, we obtain

$$\begin{aligned} A^{-1}\widehat{V}_m &= V_m R_m^{-1} = V_{m+1} \begin{bmatrix} R_m^{-1} \\ \mathbf{0} \end{bmatrix} = V_{m+1} Q_m^H Q_m \begin{bmatrix} R_m^{-1} \\ \mathbf{0} \end{bmatrix} \\ &= \begin{bmatrix} \widehat{V}_m & \widetilde{\boldsymbol{v}}_{m+1} \end{bmatrix} Q_m \begin{bmatrix} R_m^{-1} \\ \mathbf{0} \end{bmatrix} =: \begin{bmatrix} \widehat{V}_m & \widetilde{\boldsymbol{v}}_{m+1} \end{bmatrix} \widetilde{F}_m \qquad (4.21) \\ &= \widehat{V}_m F_m + \widetilde{\boldsymbol{v}}_{m+1} \boldsymbol{f}_m^\top, \quad \text{with } \widetilde{F}_m \text{ partitioned as } \widetilde{F}_m = \begin{bmatrix} F_m \\ \boldsymbol{f}_m^\top \end{bmatrix}. \end{aligned}$$

We note that both $\widehat{V}_m$ and W_m are orthonormal bases of the same space $A\mathscr{K}_m$, which implies a relation of the form $\widehat{V}_m = W_m T$ with a unitary matrix $T \in \mathbb{C}^{m\times m}$. Therefore,

$$F_m = [(A^{-1}\widehat{\boldsymbol{v}}_j, \widehat{\boldsymbol{v}}_k)]_{1\le j,k\le m} = T^H G_m T$$

which identifies the matrix F_m as similar to G_m. The zeros $\widetilde{\theta}_j$ of p_m^{MR} are therefore the reciprocals of the eigenvalues of F_m, determined by

$$\begin{aligned} \frac{1}{\widetilde{\theta}_j}\widehat{\boldsymbol{y}}_j = F_m\widehat{\boldsymbol{y}}_j &= \begin{bmatrix} I_m & \mathbf{0} \end{bmatrix} Q_m \begin{bmatrix} R_m^{-1} \\ \mathbf{0} \end{bmatrix} \widehat{\boldsymbol{y}}_j \\ &= \begin{bmatrix} I_m & \mathbf{0} \end{bmatrix} Q_m \begin{bmatrix} I_m \\ \mathbf{0} \end{bmatrix} R_m^{-1}\widehat{\boldsymbol{y}}_j =: \widehat{Q}_m R_m^{-1}\widehat{\boldsymbol{y}}_j, \end{aligned}$$

or, equivalently, as solution of the generalized eigenvalue problem

$$R_m\widehat{\boldsymbol{y}}_j = \widetilde{\theta}_j \widehat{Q}_m \widehat{\boldsymbol{y}}_j.$$

The matrix $\widehat{Q}_m$ is obtained by deleting the last row and column of Q_m, which, by (2.53) and (2.54), yields

$$\widehat{Q}_m = \begin{bmatrix} I_{m-1} & \mathbf{0} \\ \mathbf{0} & c_m \end{bmatrix} G_{m-1} \begin{bmatrix} G_{m-2} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \cdots \begin{bmatrix} G_1 & O \\ O & I_{m-2} \end{bmatrix}.$$

Equation (4.21) shows that F_m represents the orthogonal section of A^{-1} onto $A\mathscr{K}_m$ with respect to $\widehat{V}_m$. Its eigenvalues $1/\widetilde{\theta}_j$ are therefore the Ritz values of A^{-1} with respect to this space, and thus satisfy

$$\mathbf{0} = \widehat{V}_m^* \left(A^{-1}\widehat{V}_m\widehat{\boldsymbol{y}}_j - \frac{1}{\widetilde{\theta}_j}\widehat{V}_m\widehat{\boldsymbol{y}}_j \right) = \widehat{V}_m^* \left(A^{-1}\widehat{\boldsymbol{z}}_j - \frac{1}{\widetilde{\theta}_j}\widehat{\boldsymbol{z}}_j \right)$$

with Ritz vectors $\widehat{\boldsymbol{z}}_j := \widehat{V}_m\widehat{\boldsymbol{y}}_j$, which, upon multiplication by $\widetilde{\theta}_j$, substituting $A^{-1}\widehat{V}_m = V_m R_m^{-1}$ and multiplication by R_m^H, becomes

$$\mathbf{0} = R_m^H \widehat{V}_m^* \left(AV_m R_m^{-1}\widehat{\boldsymbol{y}}_j - \widetilde{\theta}_j V_m R_m^{-1}\widehat{\boldsymbol{y}}_j \right) = (AV_m)^* \left(A\widetilde{\boldsymbol{z}}_j - \widetilde{\theta}_j \widetilde{\boldsymbol{z}}_j \right), \qquad (4.22)$$

where $\widetilde{\boldsymbol{z}}_j := V_m \widetilde{\boldsymbol{y}}_j := V_m R_m^{-1} \widehat{\boldsymbol{y}}_j = A^{-1} \widehat{\boldsymbol{z}}_j$. Vectors $\widetilde{\boldsymbol{z}}_j$ and numbers $\widetilde{\theta}_j$ which satisfy the previous equation are called *harmonic Ritz vectors and values* with respect to A and $\mathscr{K}_m$ (cf. Paige et al. (1995)). A better known characterization of these quantities is

$$(AV_m)^* V_m \widetilde{\boldsymbol{y}}_j = \frac{1}{\widetilde{\theta}_j} (AV_m)^* A V_m \widetilde{\boldsymbol{y}}_j \quad \text{i.e.,} \quad H_m^H \widetilde{\boldsymbol{y}}_j = \frac{1}{\widetilde{\theta}_j} \widetilde{H}_m^H \widetilde{H}_m \widetilde{\boldsymbol{y}}_j.$$

That this formulation gives rise to the same set of eigenvalues can be seen from the similarity transformation

$$\begin{aligned}(\widetilde{H}_m^H \widetilde{H}_m)^{-1} H_m^H &= \begin{bmatrix} R_m^{-1} & \mathbf{0} \end{bmatrix} Q_m \begin{bmatrix} I_m \\ \mathbf{0} \end{bmatrix} \\ &= R_m^{-1} \begin{bmatrix} I_m & \mathbf{0} \end{bmatrix} Q_m \begin{bmatrix} R_m^{-1} \\ \mathbf{0} \end{bmatrix} R_m = R_m^{-1} F_m R_m.\end{aligned}$$

The harmonic Ritz vectors lie in $\mathscr{K}_m$ and, in view of (4.22), satisfy

$$(A - \widetilde{\theta}_j I) \widetilde{\boldsymbol{z}}_j \perp A\mathscr{K}_m.$$

In other words, $(A - \widetilde{\theta}_j I)\widetilde{\boldsymbol{z}}_j \in \mathscr{K}_{m+1} \cap (A\mathscr{K}_m)^\perp = \operatorname{span}\{\boldsymbol{r}_m^{\mathrm{MR}}\}$ and therefore, if the polynomials $\widetilde{z}_j \in \mathscr{P}_{m-1}$ are defined by $\widetilde{\boldsymbol{z}}_j = \widetilde{z}_j(A)\boldsymbol{r}_0$, there holds

$$\widetilde{z}_j(\zeta) = \tau_j \frac{p_m^{\mathrm{MR}}(\zeta)}{\zeta - \widetilde{\theta}_j} = \tau_j \frac{\widehat{g}_m(\zeta)}{\zeta - \widetilde{\theta}_j} \tag{4.23}$$

for some normalization factor $\tau_j \neq 0$.

Remark 4.3.7. Polynomials which possess the reproducing property (4.19) are called *kernel polynomials*. Their role in Krylov subspace methods was first explored by Stiefel (1958) in the Hermitian case and later extended to the non-Hermitian case by Freund (Freund 1992*b*, Freund 1992*a*). See also Manteuffel & Otto (1994).

4.3.3 The Implicitly Restarted Arnoldi Process

When manipulating Krylov subspaces, the following fundamental task often arises: given a Krylov space $\mathscr{K}_m(A, \boldsymbol{v}_1)$ which is not A-invariant, along with the associated Arnoldi factorization

$$AV_m = V_m H_m + \eta_{m+1,m} \boldsymbol{v}_{m+1} \boldsymbol{u}_m^\top \qquad (\eta_{m+1,m} \neq 0) \tag{4.24}$$

and given an arbitrary vector $\boldsymbol{v} \in \mathscr{K}_{m-1}(A, \boldsymbol{v}_1)$, generate the Arnoldi factorization associated with $\mathscr{K}_p(A, \boldsymbol{v})$ with p as large as possible *without* performing additional multiplications with A. The technique which accomplishes this task is known as the *implicitly restarted Arnoldi (IRA)* process and is due to Sorensen (1992).

As a member of $\mathscr{K}_{m-1}$, $\boldsymbol{v}$ has the representation $\boldsymbol{v} = q_{k-1}(A)\boldsymbol{v}_1$ with q_{k-1} of exact degree $k-1$, $1 \leq k < m$. In other words, $\boldsymbol{v} \in \mathscr{K}_k \setminus \mathscr{K}_{k-1}$. We will show that $p = m - k$ is maximal and the resulting Arnoldi factorization has the form

$$A\check{V}_p = \check{V}_p \check{H}_p + \check{\eta}_{p+1,p} \check{\boldsymbol{v}}_{p+1} \boldsymbol{u}_p^\top \tag{4.25}$$

with $\check{\boldsymbol{v}}_1 = \boldsymbol{v}/\|\boldsymbol{v}\|$. That $p = m - k$ holds should not come as a surprise because the construction of the factorization (4.24) requires m multiplications by A, whereas $\boldsymbol{v}$ can be computed by only $k - 1$ matrix-vector products. Exactly $p + 1 = m - k + 1$, i.e., the number of the 'remaining' multiplications by A are needed to establish (4.25) in the conventional way.

We assume that the polynomial q_{k-1} is given in factored form

$$q_{k-1}(\zeta) = \prod_{j=1}^{k-1}(\zeta - \theta_j)$$

as this is how it will arise in our applications. It is obviously sufficient to show how the decomposition (4.25) can be established in the case $k = 2$, i.e., if $\boldsymbol{v} = (A - \theta I)\boldsymbol{v}_1$. Polynomials of higher degree can then be handled by repeated application of the procedure below.

Each step of the IRA method is based on one step of the shifted QR algorithm. Following Sorensen (1992, p. 363), we begin by subtracting θV_m on both sides of the Arnoldi decomposition (4.24)

$$(A - \theta I)V_m = V_m(H_m - \theta I) + \eta_{m+1,m}\boldsymbol{v}_{m+1}\boldsymbol{u}_m^\top,$$

then form the QR factorization of $H_m - \theta I$,

$$(A - \theta I)V_m = V_m QR + \eta_{m+1,m}\boldsymbol{v}_{m+1}\boldsymbol{u}_m^\top, \tag{4.26}$$

multiply by Q from the right,

$$(A - \theta I)V_m Q = (V_m Q)(RQ) + \eta_{m+1,m}\boldsymbol{v}_{m+1}\boldsymbol{u}_m^\top Q,$$

and add $\theta V_m Q$ on both sides to obtain

$$A(V_m Q) = (V_m Q)(RQ + \theta I) + \eta_{m+1,m}\boldsymbol{v}_{m+1}\boldsymbol{u}_m^\top Q. \tag{4.27}$$

We rewrite (4.27) to introduce some extra notation:

$$\begin{aligned} A\begin{bmatrix}\check{\boldsymbol{v}}_1 & \dots & \check{\boldsymbol{v}}_{m-1} & \boldsymbol{v}_m^+\end{bmatrix} &= \begin{bmatrix}\check{\boldsymbol{v}}_1 & \dots & \check{\boldsymbol{v}}_{m-1} & \boldsymbol{v}_m^+\end{bmatrix}\begin{bmatrix}\check{H}_{m-1} & * \\ \eta^+\boldsymbol{u}_{m-1}^\top & *\end{bmatrix} \\ &\quad + \eta_{m+1,m}\boldsymbol{v}_{m+1}\begin{bmatrix}0 & \dots & 0 & q_{m,m-1} & q_{m,m}\end{bmatrix} \end{aligned}$$

where we have made use of the fact that

$$RQ + \theta I = \begin{bmatrix}\check{H}_{m-1} & * \\ \eta^+\boldsymbol{u}_{m-1}^\top & *\end{bmatrix} \in \mathbb{C}^{m\times m}$$

is again an upper Hessenberg matrix due to the upper Hessenberg form of Q. We note in passing that, in case θ happens to be an eigenvalue of H_m (and only then), the last row of R is zero (and only the last row since H_m is nonderogatory) and therefore $\eta^+ = 0$.

We now omit the last column in (4.27), giving

$$\begin{aligned} A\begin{bmatrix}\check{\boldsymbol{v}}_1 & \dots & \check{\boldsymbol{v}}_{m-1}\end{bmatrix} &= \begin{bmatrix}\check{\boldsymbol{v}}_1 & \dots & \check{\boldsymbol{v}}_{m-1}\end{bmatrix}\check{H}_{m-1} \\ &\quad + (\eta^+\boldsymbol{v}_m^+ + \eta_{m+1,m}q_{m,m-1}\boldsymbol{v}_{m+1})\boldsymbol{u}_{m-1}^\top, \end{aligned}$$

which, setting $\check{\eta}_{m,m-1} := \|\eta^+\boldsymbol{v}_m^+ + \eta_{m+1,m}q_{m,m-1}\boldsymbol{v}_{m+1}\|$, becomes

$$A\check{V}_{m-1} = \check{V}_{m-1}\check{H}_{m-1} + \check{\eta}_{m,m-1}\check{\boldsymbol{v}}_m\boldsymbol{u}_{m-1}^\top. \tag{4.28}$$

Theorem 4.3.8. *With the notation introduced above, the decomposition* (4.28) *is an Arnoldi factorization of A with respect to the Krylov space $\mathscr{K}_{m-1}(A,(A-\theta I)\boldsymbol{v}_1)$.*

Proof. Since Q is unitary, it follows that the elements of $\check{V}_{m-1} = \begin{bmatrix}\check{\boldsymbol{v}}_1 & \dots & \check{\boldsymbol{v}}_{m-1}\end{bmatrix}$ are orthonormal as the first $m-1$ elements of V_mQ. Next, the vector

$$\check{\boldsymbol{v}}_m = (\eta^+\boldsymbol{v}_m^+ + \eta_{m+1,m}q_{m,m-1}\boldsymbol{v}_{m+1})/\check{\eta}_{m,m-1}$$

has unit norm and is orthogonal to $\check{\boldsymbol{v}}_1,\dots,\check{\boldsymbol{v}}_{m-1}$ since $\boldsymbol{v}_m^+$, as the last element V_mQ, is orthogonal to the previous elements $\check{\boldsymbol{v}}_1,\dots,\check{\boldsymbol{v}}_{m-1}$ and since $\boldsymbol{v}_{m+1}$ is orthogonal to V_m and hence also to V_mQ. That the new first basis vector $\check{\boldsymbol{v}}_1$ is a multiple of $(A-\theta I)\boldsymbol{v}_1$ follows by equating the first vector on both sides of (4.26). It remains to show that the Hessenberg matrix $\check{H}_{m-1}$ is unreduced. If, say, $\check{\eta}_{k+1,k} = 0$ for some $k < m$ this would imply that $\mathscr{K}_k(A,\check{\boldsymbol{v}}_1)$ is a proper A-invariant subspace of $\mathscr{K}_m(A,\boldsymbol{v}_1)$, which, in view of Lemma 4.3.2, contradicts the assumption $\eta_{m+1,m} \neq 0$. □

As mentioned previously, the decomposition (4.25) involving a new starting vector $\check{\boldsymbol{v}}_1 = q_{k-1}(A)\boldsymbol{v}_1$ is effected by $k-1$ steps of the procedure outlined above. For later use, we note that the associated Krylov space is given by

$$\mathscr{K}_p(A,\check{\boldsymbol{v}}_1) = \{r(A)q_{k-1}(A)\boldsymbol{v}_1 \,:\, r \in \mathscr{P}_{p-1}\} \subset \mathscr{K}_{p+k-1}(A,\boldsymbol{v}_1).$$

4.4 The Conjugate Gradient Method and its Descendants

The ancestor of all Krylov subspace methods is the conjugate gradient method of Hestenes & Stiefel (1952), which in our terminology is the OR method for solving (1.1) when A is selfadjoint and positive definite. As we shall see below, these additional assumptions permit the approximations to be generated at substantially less expense than in the gerneral case.

Since it is fair to say that all subsequently developed Krylov subspace methods can be seen as attempts to generalize CG to broader classes of operator equations, we provide in this section a brief overview of the most important developments, highlighting in the process some of the algorithms proposed along the way. Upon its introduction in the early 1950s, CG was viewed as a direct solution method, terminating in a finite number of steps and, in contrast with factorization procedures, working only with the original matrix A. As such, however, it was found lacking, since the finite termination property is usually destroyed by roundoff error, and, despite some activity in the field of optimization, the CG method was discarded as a solution approach for linear equations. The rekindling of interest in CG as an equation-solver is usually attributed to Reid (1971), which points out that CG offers many advantages over other solvers when applied to large, sparse and reasonably well-conditioned systems of equations, providing a sufficiently accurate approximation after a number of steps which is usually much smaller than the dimension of the system. In 1975, a major innovation was achieved by Concus, Golub and O'Leary, who showed how what is now called *preconditioning* can be incorporated into the CG algorithm, making CG an effective solver also for less well-conditioned problems such as

those resulting from the discretization of second or fourth-order elliptic boundary value problems, whenever a suitable preconditioner can be found. The point of view taken in these early papers, however, was that of CG being used to *accelerate* a given stationary iteration, in which the splitting operator plays the role of the preconditioner. Around the same time, Paige and Saunders proposed their algorithms SYMMLQ and MINRES[1], which constitute the first stable Krylov subspace methods for solving Hermitian indefinite systems. The methods of Paige and Saunders employ the (Hermitian) Lanczos process to generate an orthogonal basis of the Krylov space and generate the approximations using a QR factorization of the associated tridiagonal matrix. This is exactly the approach used many years later in the state-of-the-art Krylov solvers for general non-Hermitian problems, namely GMRES (Saad & Schultz 1986) and QMR (Freund & Nachtigal 1991). Another line of research in the early 1980s attempted to generalize the *error-minimizing* property of CG, and this led to methods known as ORTHOMIN, ORTHODIR and ORTHORES. The first of these, ORTHOMIN, was first proposed by Vinsome (1976) in a conference on oil reservoir simulation. These methods were later proposed by (Young & Jea 1980) as three canonical meta-algorithms for generalizing CG to non-Hermitian problems, and subsequently analyzed by many researchers, among them Elman (1982), Eisenstat et al. (1983), Young & Jea (1980), Faber & Manteuffel (1984), Faber & Manteuffel (1987), Joubert & Young (1987), and Ashby & Gutknecht (1993).

We proceed by first restricting A to be Hermitian, resulting in a version of MR/OR with short recurrences, in which each new basis vector of the Krylov space is generated using only its two predecessors. We then add the further assumption of positive definiteness and derive the CG algorithm as well as its MR counterpart, which is known as the *conjugate residual (CR) method*. We then derive the methods known as ORTHOMIN, ORTHORES and ORTHODIR.

4.4.1 The Hermitian Case

Under the additional assumption $A = A^*$, the entries of the square Hessenberg matrix H_m (cf. (2.39),(4.3)) satisfy

$$\eta_{j,i} = (A\boldsymbol{v}_i, \boldsymbol{v}_j) = (\boldsymbol{v}_i, A\boldsymbol{v}_j) = \overline{(A\boldsymbol{v}_j, \boldsymbol{v}_i)} = \overline{\eta}_{i,j}$$

i.e., H_m is Hermitian, which for a Hessenberg matrix means tridiagonal. Moreover, since the subdiagonal elements of H_m are nonnegative, the Hermitian tridiagonal matrix H_m contains only real entries, suggesting the simpler notation

$$\widetilde{H}_m = \begin{bmatrix} \gamma_1 & \delta_2 & 0 & \dots & 0 \\ \delta_2 & \gamma_2 & \delta_3 & \ddots & \vdots \\ 0 & \delta_3 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \delta_m \\ 0 & \dots & 0 & \delta_m & \gamma_m \\ 0 & \dots & & 0 & \delta_{m+1} \end{bmatrix}, \qquad m = 1, \dots L, \quad (\delta_{L+1} = 0). \tag{4.29}$$

[1]More precisely, MINRES computes the MR approximations with respect to the Euclidean inner product and the associated OR approximations when they exist. SYMMLQ minimizes the *error* over the slightly different shifted Krylov space $\boldsymbol{x}_0 + \mathscr{K}_m(A, A\boldsymbol{r}_0)$, a scheme which only works in the Hermitian case.

The resulting version of the Arnoldi-process for selfadjoint A is known as the *(Hermitian) Lanczos process* (Lanczos 1950, Lanczos 1952) and is summarized in Algorithm 4.4.1.

Algorithm 4.4.1 The Hermitian Lanczos algorithm.

$\boldsymbol{r}_0$ given, $\delta_1 = \|\boldsymbol{r}_0\|$, $\boldsymbol{v}_1 = \boldsymbol{r}_0/\delta_1$, $\boldsymbol{v}_0 = \boldsymbol{0}$
for $m = 1, 2, \ldots$
  $\gamma_m = (A\boldsymbol{v}_m, \boldsymbol{v}_m)$
  $\widehat{\boldsymbol{v}}_{m+1} = A\boldsymbol{v}_m - \gamma_m \boldsymbol{v}_m - \delta_m \boldsymbol{v}_{m-1}$
  $\delta_{m+1} = \|\widehat{\boldsymbol{v}}_{m+1}\|$
  $\boldsymbol{v}_{m+1} = \widehat{\boldsymbol{v}}_{m+1}/\delta_{m+1}$
end

The practical consequence of the tridiagonal form of $\widetilde{H}_m$ is that the orthogonalization of $A\boldsymbol{v}_j$ against $\mathscr{V}_j$, $j = 1, \ldots L$, entails only inner products and linear combinations involving the two previous basis vectors, hence earlier basis vectors need not be stored. The Hermitian case thus allows the orthogonalization procedure for generating the basis of $\mathscr{K}_m(A, \boldsymbol{r}_0)$ to be carried out with a constant amount of work and storage per step.

There are several equivalent formulations of the Lanczos process (see Gutknecht (1997) for a discussion of the more general non-Hermitian Lanczos process). In particular, in the literature on Krylov subspace methods for Hermitian positive definite systems the variant given in Algorithm 4.4.2, in which the basis vectors remain unnormalized, is often used.

Algorithm 4.4.2 The unnormalized Hermitian Lanczos process.

$\boldsymbol{r}_0$ given, $\check{\boldsymbol{v}}_1 = \boldsymbol{r}_0$, $\check{\delta}_1 = 0$, $\check{\boldsymbol{v}}_0 = \boldsymbol{0}$
for $m = 1, 2, \ldots$
  $\check{\gamma}_m = (A\check{\boldsymbol{v}}_m, \check{\boldsymbol{v}}_m)/(\check{\boldsymbol{v}}_m, \check{\boldsymbol{v}}_m)$
  $\check{\delta}_m = (A\check{\boldsymbol{v}}_m, \check{\boldsymbol{v}}_{m-1})/(\check{\boldsymbol{v}}_{m-1}, \check{\boldsymbol{v}}_{m-1})$
  $\check{\boldsymbol{v}}_{m+1} = A\check{\boldsymbol{v}}_m - \check{\gamma}_m \check{\boldsymbol{v}}_m - \check{\delta}_m \check{\boldsymbol{v}}_{m-1}$
end

As is easily verified, the quantities of the two algorithms are related by

$$\check{\boldsymbol{v}}_m = \pi_m \boldsymbol{v}_m, \quad \check{\gamma}_m = \gamma_m, \quad \check{\delta}_m = \frac{\pi_{m-1}}{\pi_m}\delta_m, \qquad m = 1, 2, \ldots$$

where $\pi_m := \delta_1 \delta_2 \cdots \delta_m$.

GMRES for Hermitian Systems: MINRES

Since $\widetilde{H}_m$ is a real matrix, its QR-factors (2.49) may also be chosen real, hence no phase factors occur in the Givens parameters (2.57). The tridiagonal form of $\widetilde{H}_m$ leads to a

triangular factor R_m containing only three nonvanishing diagonals

$$\widetilde{H}_m = Q_m \begin{bmatrix} R_m \\ \mathbf{0} \end{bmatrix}, \qquad R_m = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & 0 & \dots & 0 \\ 0 & r_{2,2} & r_{2,3} & \ddots & \ddots & \vdots \\ & & r_{3,3} & \ddots & \ddots & 0 \\ & & & \ddots & \ddots & r_{m-2,m} \\ \vdots & & & \ddots & \ddots & r_{m-1,m} \\ 0 & & & \dots & 0 & r_{m,m} \end{bmatrix}. \tag{4.30}$$

As a result, the entries of the triangular factor have the following simple expressions in terms of the entries of $\widetilde{H}_m$ and the trigonometric quantities c_m and s_m defined in (2.57).

Proposition 4.4.1. *The nonzero entries of the triangular factor R_m in* (4.30) *for the case of Hermitian A are given by*

$$r_{j-2,j} = s_{j-2}\delta_j, \qquad j \geq 3,$$

$$r_{j-1,j} = \begin{cases} c_1\delta_2 + s_1\gamma_2, & j = 2, \\ c_{j-1}c_{j-2}\delta_j + s_{j-2}\gamma_j, & j \geq 3, \end{cases}$$

$$r_{j,j} = (\widetilde{r}_{j,j}^2 + \delta_{j+1}^2)^{1/2}, \quad j \geq 1, \qquad \widetilde{r}_{j,j} = \begin{cases} \gamma_1, & j = 1, \\ -s_1\delta_2 + c_1\gamma_2, & j = 2, \\ -s_{j-1}c_{j-2}\delta_j + c_{j-1}\gamma_j, & j \geq 3. \end{cases}$$

In order to take advantage of the short recurrences for the basis vectors it is necessary to develop recurrence formulas also for the iterates and residuals which do not involve basis vectors no longer needed in the basis generation. We shall derive such update formulas for the GMRES algorithm, resulting in an algorithm known as MINRES (see also (Paige & Saunders 1975, Fischer 1996)).

The coefficient vector $\boldsymbol{y}_m^{\mathrm{MR}}$ of the MR approximation $\boldsymbol{x}_m^{\mathrm{MR}} = \boldsymbol{x}_0 + V_m\boldsymbol{y}_m^{\mathrm{MR}}$ satisfies

$$R_m\boldsymbol{y}_m^{\mathrm{MR}} = \beta\widetilde{Q}_m^H\boldsymbol{u}_1^{(m+1)} = \beta\boldsymbol{z}_m,$$

(cf. (2.65)), hence $\boldsymbol{x}_m^{\mathrm{MR}} = \boldsymbol{x}_0 + \beta V_m R_m^{-1}\boldsymbol{z}_m$, which, upon setting $P_m = [\boldsymbol{p}_1, \dots, \boldsymbol{p}_m] := V_m R_m^{-1}$, yields the update formula

$$\boldsymbol{x}_m^{\mathrm{MR}} = \boldsymbol{x}_{m-1}^{\mathrm{MR}} + \beta c_m\zeta_m\boldsymbol{p}_m.$$

Because the update of the approximation is in the direction of $\boldsymbol{p}_m$, the vectors P_m are called *direction vectors*. Comparing coefficients in $V_m = P_m R_m$, we see that the direction vectors obey the recurrence formulas

$$\begin{aligned} \boldsymbol{p}_1 &= r_{1,1}^{-1}\boldsymbol{v}_1 \\ \boldsymbol{p}_2 &= r_{2,2}^{-1}(\boldsymbol{v}_2 - r_{1,2}\boldsymbol{p}_1) \\ \boldsymbol{p}_m &= r_{m,m}^{-1}(\boldsymbol{v}_m - r_{m-1,m}\boldsymbol{p}_{m-1} - r_{m-2,m}\boldsymbol{p}_{m-2}), \qquad m \geq 3. \end{aligned}$$

In view of $AP_m = AV_m R_m^{-1} = \widehat{V}_m R_m R_m^{-1} = \widehat{V}_m$. the residual may be expressed as

$$\boldsymbol{r}_m^{\mathrm{MR}} = \boldsymbol{r}_{m-1}^{\mathrm{MR}} - \beta c_m\zeta_m\widehat{\boldsymbol{v}}_m.$$

Usually, however, it is sufficient to compute the residual norm, which is updated as $\|\boldsymbol{r}_m^{\mathrm{MR}}\| = s_m\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|$ (cf. (2.24). This concludes the update formulas for the MINRES algorithm. The corresponding OR quantities may be computed using (3.4), (2.31) and (2.26).

GCR for Hermitian Systems

As in the GMRES algorithm, GCR (cf. Algorithm 4.2.2) undergoes considerable simplification when A is Hermitian. First, in view of the identity

$$A\boldsymbol{r}_{m-1} = \mu_{m,m}\boldsymbol{w}_m + \sum_{j=1}^{m-1}\mu_{j,m}\boldsymbol{w}_j,$$

where we have set $\mu_{m,m} := \|\widehat{\boldsymbol{w}}_m\|$ and $\mu_{j,m} := (A\boldsymbol{r}_{m-1}, \boldsymbol{w}_j), j = 1, \ldots, m-1$, we write all these equations up to step m in vector form as

$$A\left[\boldsymbol{r}_0, \ldots, \boldsymbol{r}_{m-1}\right] = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m]\begin{bmatrix}\mu_{1,1} & \cdots & \mu_{1,m}\\ & \ddots & \vdots\\ & & \mu_{m,m}\end{bmatrix}. \tag{4.31}$$

The MR condition states that $\boldsymbol{r}_m \perp \mathscr{W}_m = A\mathscr{C}_m$, and if $A = A^*$ this is equivalent with $A\boldsymbol{r}_m \perp \mathscr{C}_m$. Moreover, in the Krylov space case $\mathscr{C}_m = \mathscr{K}_m(A, \boldsymbol{r}_0)$, we also have

$$A\boldsymbol{r}_m \perp \mathscr{K}_m(A, \boldsymbol{r}_0) \supset A\mathscr{K}_{m-1}(A, \boldsymbol{r}_0) \ = \mathrm{span}\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{m-1}\}.$$

But this implies that $\mu_{j,m} = 0$ for $j = 1, \ldots, m-1$, and (4.31) simplifies to

$$A\left[\boldsymbol{r}_0, \ldots, \boldsymbol{r}_{m-1}\right] = \left[\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\right] B_m \tag{4.32}$$

with the bidiagonal matrix

$$B_m = \begin{bmatrix}\mu_1 & \nu_2 & & \\ & \mu_2 & \ddots & \\ & & \ddots & \nu_m\\ & & & \mu_m\end{bmatrix}, \qquad \mu_j := \mu_{j,j}, \quad \nu_j := \mu_{j-1,j}, \quad j = 1, \ldots, m.$$

This results in the decomposition $AC_m = W_m$ with an orthonormal basis W_m of the approximation space $\mathscr{W}_m = A\mathscr{K}_m(A, \boldsymbol{r}_0)$ with pre-images $C_m = [\boldsymbol{r}_0, \ldots, \boldsymbol{r}_{m-1}]B_m^{-1}$. Alternatively, we may write this relation in unnormalized form as

$$A\left[\boldsymbol{r}_0, \ldots, \boldsymbol{r}_{m-1}\right] = \left[\widetilde{\boldsymbol{w}}_1, \ldots, \widetilde{\boldsymbol{w}}_m\right]\widetilde{B}_m, \qquad \widetilde{B}_m = D_m^{-1}B_m, \quad D_m = \mathrm{diag}(\mu_1, \ldots, \mu_m).$$

As a consequence of the orthonormality of the $\boldsymbol{w}_j$ and the decomposition (4.32), the Hermitian matrix $[\boldsymbol{r}_0, \ldots, \boldsymbol{r}_{m-1}]^*A[\boldsymbol{r}_0, \ldots, \boldsymbol{r}_{m-1}]$ is upper triangular, hence diagonal, implying that the residuals are orthogonal with respect to the sesquilinear form associated with A, which is also referred to as being *A-conjugate*. We define

$$\rho_j := (A\boldsymbol{r}_j, \boldsymbol{r}_j), \qquad j = 0, \ldots, m-1.$$

As a result, we obtain the following update formulas, in which we use the unnormalized bases $\widetilde{W}_m = W_m D_m$ and $\widetilde{C}_m := C_m D_m$ in place of W_m and C_m:

$$\begin{aligned}
\widetilde{\boldsymbol{w}}_m &= A\boldsymbol{r}_{m-1} - \frac{(A\boldsymbol{r}_{m-1}, \widetilde{\boldsymbol{w}}_{m-1})}{\mu_{m-1}^2}\widetilde{\boldsymbol{w}}_{m-1} \\
\widetilde{\boldsymbol{c}}_m &= \boldsymbol{r}_{m-1} - \frac{(A\boldsymbol{r}_{m-1}, \widetilde{\boldsymbol{w}}_{m-1})}{\mu_{m-1}^2}\widetilde{\boldsymbol{c}}_{m-1} \\
\boldsymbol{r}_m &= \boldsymbol{r}_{m-1} - \frac{(\boldsymbol{r}_0, \widetilde{\boldsymbol{w}}_m)}{\mu_m^2}\widetilde{\boldsymbol{w}}_m \\
\boldsymbol{x}_m &= \boldsymbol{x}_{m-1} - \frac{(\boldsymbol{r}_0, \widetilde{\boldsymbol{w}}_m)}{\mu_m^2}\widetilde{\boldsymbol{c}}_m
\end{aligned}$$

Three further observations now result in the Hermitian GCR algorithm: first, from

$$(\boldsymbol{r}_{m-1}, \boldsymbol{w}_m) = (\boldsymbol{r}_0 - \sum_{j=1}^{m-1}(\boldsymbol{r}_0, \boldsymbol{w}_j)\boldsymbol{w}_j, \boldsymbol{w}_m) = (\boldsymbol{r}_0, \boldsymbol{w}_m)$$

we conclude $(\boldsymbol{r}_0, \widetilde{\boldsymbol{w}}_m) = (\boldsymbol{r}_{m-1}, \widetilde{\boldsymbol{w}}_m)$. Next, from the update formula for $\widetilde{\boldsymbol{w}}_m$ we obtain

$$(\boldsymbol{r}_{m-1}, \widetilde{\boldsymbol{w}}_m) = (\boldsymbol{r}_{m-1}, A\boldsymbol{r}_{m-1}) - \overline{(A\boldsymbol{r}_{m-1}, \boldsymbol{w}_{m-1})}(\boldsymbol{r}_{m-1}, \boldsymbol{w}_{m-1}) = (A\boldsymbol{r}_{m-1}, \boldsymbol{r}_{m-1}) = \rho_{m-1}$$

since $\boldsymbol{r}_{m-1} \perp \mathscr{W}_{m-1}$. Finally, we note

$$(A\boldsymbol{r}_{m-1}, \boldsymbol{r}_{m-1}) = (A\boldsymbol{r}_{m-1}, \boldsymbol{r}_{m-2}) - \overline{(\boldsymbol{r}_{m-2}, \boldsymbol{w}_{m-1})}(A\boldsymbol{r}_{m-1}, \boldsymbol{w}_{m-1})$$

in which the first term on the right vanishes due to the A-conjugacy of the residuals, yielding

$$(A\boldsymbol{r}_{m-1}, \widetilde{\boldsymbol{w}}_{m-1}) = \frac{-\mu_{m-1}^2}{\overline{(\boldsymbol{r}_{m-2}, \widetilde{\boldsymbol{w}}_{m-1})}}(A\boldsymbol{r}_{m-1}, \boldsymbol{r}_{m-1}) = -\mu_{m-1}^2\frac{(A\boldsymbol{r}_{m-1}, \boldsymbol{r}_{m-1})}{(A\boldsymbol{r}_{m-2}, \boldsymbol{r}_{m-2})} = -\mu_{m-1}\frac{\rho_{m-1}}{\rho_{m-2}}.$$

Incorporating these results into the update formulas, we obtain the Hermitian GCR algorithm summarized in Algorithm 4.4.1.

Each step of the Hermitian GCR algorithm requires one matrix-vector multiplication, two inner products and four vector updates. The storage requirements consist of the five vectors $\widetilde{\boldsymbol{w}}_m$, $\widetilde{\boldsymbol{c}}_m$, $A\boldsymbol{r}_m$, $\boldsymbol{r}_m$ and $\boldsymbol{x}_m$.

The associated OR approximations can be computed from the MR approximations as usual. As mentioned in Section 4.2.1, this version of GCR will not break down prematurely for any $\boldsymbol{r}_0$ only if $0 \notin W(A)$, which for Hermitian matrices means that they be positive or negative definite. Otherwise the only safe algorithm is MINRES. For positive definite problems the conjugate gradient methods discussed in the following section offer more efficient schemes for computing OR approximations.

4.4.2 The Hermitian Positive Definite Case

When A is positive definite in addition to being Hermitian, this implies that its field of values does not contain the origin, and therefore the OR approximation exists in each

Algorithm 4.4.3 The Hermitian GCR algorithm.

$\boldsymbol{x}_0$ given, $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$
$\widetilde{\boldsymbol{w}}_0 = \boldsymbol{0}, \widetilde{\boldsymbol{c}}_0 = \boldsymbol{0}, \rho_0 = 0, \rho_{-1} = 1$
for $j = 1, 2 \ldots$
  $\widetilde{\boldsymbol{w}}_m = A\boldsymbol{r}_{m-1} + \frac{\rho_{m-1}}{\rho_{m-2}}\widetilde{\boldsymbol{w}}_{m-1}$
  $\mu_m^2 = (\widetilde{\boldsymbol{w}}_m, \widetilde{\boldsymbol{w}}_m)$
  $\widetilde{\boldsymbol{c}}_m = \boldsymbol{r}_{m-1} + \frac{\rho_{m-1}}{\rho_{m-2}}\widetilde{\boldsymbol{c}}_{m-1}$
  $\boldsymbol{r}_m = \boldsymbol{r}_{m-1} - \frac{\rho_{m-1}}{\mu_m^2}\widetilde{\boldsymbol{w}}_m$
  $\rho_m = (A\boldsymbol{r}_m, \boldsymbol{r}_m)$
  $\boldsymbol{x}_m = \boldsymbol{x}_{m-1} + \frac{\rho_{m-1}}{\mu_m^2}\widetilde{\boldsymbol{c}}_m$
end

iteration and can be computed stably without first computing the associated MR approximation.

Even more importantly, the fact that A is Hermitian and positive definite implies that all powers $A^s, s \in \mathbb{R}$, of A induce an inner product $(A^s\cdot, \cdot)$ with associated norm $\|\cdot\|_{A^s}$. This immediately permits new interpretations of the OR property: First, since

$$\boldsymbol{r}_m \perp \mathscr{K}_m(A, \boldsymbol{r}_0) \Leftrightarrow \boldsymbol{r}_m \perp_{A^{-1}} A\mathscr{K}_m(A, \boldsymbol{r}_0),$$

we can say that the OR approximation minimizes the residual in the A^{-1}-norm, i.e., it is the MR method with respect to the A^{-1}-inner product. By analogous reasoning, the OR approximation minimizes a natural norm of the *error* $\boldsymbol{e}_m := A^{-1}\boldsymbol{b} - \boldsymbol{x}_m = A^{-1}\boldsymbol{r}_m$ as seen from

$$\boldsymbol{r}_m = A\boldsymbol{e}_m \perp \mathscr{K}_m(A, \boldsymbol{r}_0) \Leftrightarrow \boldsymbol{e}_m \perp_A \mathscr{K}_m(A, \boldsymbol{r}_0),$$

i.e., the OR approximation also minimizes the error in the A-norm over all correction vectors in $\mathscr{K}_m(A, \boldsymbol{r}_0)$. When the operator A is associated with a symmetric positive definite sesquilinear form arising from the variational formulation of an elliptic boundary value problem (as e.g. in the Galerkin finite element method) then the A-norm is known as the *energy norm*. In particular, in case of the Galerkin finite element method the OR approximation minimizes the energy norm of the error in the discrete solution over the Krylov space in the same fashion as the Galerkin approximation minimizes the error with respect to the continuous solution over the finite element space.

We shall require a third form of the Hermitian Lanczos process which underlies the most commonly used implementation of the conjugate gradient method. Since the (real) matrix $H_m = [I_m \ \boldsymbol{0}]\widetilde{H}_m$ in (4.29) is symmetric and positive definite, $\widetilde{H}_m$ possesses the factorization $\widetilde{H}_m = \widetilde{L}_m D_m L_m^\top$ (see (Golub & van Loan 1996, pp. 135)) with the unit lower triangular matrix

$$\widetilde{L}_m = \begin{bmatrix} 1 & & & & \\ \lambda_2 & 1 & & & \\ & \ddots & \ddots & & \\ & & \lambda_m & 1 & \\ & & & \lambda_{m+1} & \end{bmatrix} \in \mathbb{R}^{(m+1)\times m},$$

a diagonal matrix $D_m = \mathrm{diag}(\mu_1, \ldots, \mu_m)$ and $L_m = [I_m \; \mathbf{0}]\, \widetilde{L}_m$. Comparing coefficients yields the relations

$$\mu_1 = \gamma_1, \qquad \mu_{m+1} = \gamma_{m+1} - \mu_m \lambda_{m+1}^2, \qquad m = 1, 2, \ldots, \tag{4.33}$$

$$\lambda_{m+1} = \delta_{m+1}/\mu_m, \qquad m = 1, 2, \ldots. \tag{4.34}$$

Defining the new vector sequence $P_m := [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_m]$ by $P_m = V_m L_m^{-\top}$ now permits us to split the Lanczos decomposition $AV_m = V_{m+1}\widetilde{H}_m$ into a pair of coupled two-term recurrences

$$AP_m = V_{m+1}\widetilde{L}_m D_m, \quad P_m L_m^{\top} = V_m, \qquad m = 1, 2, \ldots. \tag{4.35}$$

From

$$P_m^* A P_m = L_m^{-1} V_m^* A V_m L_m^{-\top} = L_m^{-1} H_m L_m^{-\top} = L_m^{-1}(L_m D_m L_m^{\top}) L_m^{-\top} = D_m,$$

we observe that the vectors $\boldsymbol{p}_j$ are A-orthogonal and $\mu_j = (A\boldsymbol{p}_j, \boldsymbol{p}_j) > 0$. Comparing $\delta_{m+1}\boldsymbol{v}_{m+1} = A\boldsymbol{v}_m - \gamma_m \boldsymbol{v}_m - \delta_m \boldsymbol{v}_{m-1}$ (cf. Algorithm 4.4.1) with the first of the recurrences in (4.35), we find that

$$A\boldsymbol{p}_m = \mu_m \boldsymbol{v}_m + \delta_{m+1}\boldsymbol{v}_{m+1},$$

where we have used (4.34), and thus conclude

$$\delta_{m+1}\boldsymbol{v}_{m+1} = A\boldsymbol{p}_m - \mu_m \boldsymbol{v}_m, \quad \delta_{m+1} = \|A\boldsymbol{v}_m - \mu_m \boldsymbol{v}_m\|, \qquad m = 1, 2 \ldots.$$

This yields the coupled two-term Hermitian Lanczos algorithm summarized in Algorithm 4.4.2.

Algorithm 4.4.4 The coupled two-term Hermitian Lanczos algorithm.

```
1 v̂_1 given, δ_1 = ‖v̂_1‖, v_1 = v̂_1/δ_1, v_0 = 0
2 p_0 = 0, μ_0 = 1
3 for m = 1, 2, ...
4     p_m = v_m − δ_m/μ_{m−1} p_{m−1}
5     μ_m = (A p_m, p_m)
6     v̂_{m+1} = A p_m − μ_m v_m
7     δ_{m+1} = ‖v̂_{m+1}‖
8     v_{m+1} = v̂_{m+1}/δ_{m+1}
9 end
```

The Conjugate Gradient Algorithm

The coefficient vector $\boldsymbol{y}_m \in \mathbb{C}^m$ of the OR approximation $\boldsymbol{x}_m = \boldsymbol{x}_0 + V_m \boldsymbol{y}_m$ solves the system $V_m^* A V_m = V_m^* \boldsymbol{r}_0$ (cf. (2.45)), which suggests using the A-orthogonal direction vectors $\boldsymbol{p}_m$ to construct the solution instead. Setting $\boldsymbol{x}_m = \boldsymbol{x}_0 + P_m \boldsymbol{z}_m$, $\boldsymbol{z}_m \in \mathbb{C}_m$, and noting that

$\text{span}\{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_m\} = \text{span}\{\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m\} = \mathscr{K}_m(A,\boldsymbol{r}_0)$, we find that the coordinate vector $\boldsymbol{z}_m$ must satisfy

$$\boldsymbol{0} = P_m^*(\boldsymbol{r}_0 - AP_m\boldsymbol{z}_m), \quad \text{i.e.,} \quad D_m\boldsymbol{z}_m = P_m^*\boldsymbol{r}_0,$$

hence the components $\{\zeta_j\}_{j=1}^m$ of $\boldsymbol{z}_m$ are given by $\zeta_j = (\boldsymbol{r}_0,\boldsymbol{p}_j)/(A\boldsymbol{p}_j,\boldsymbol{p}_j)$. In view of $\boldsymbol{r}_m = \boldsymbol{r}_0 - AP_m\boldsymbol{z}_m$, the A-orthogonality of the direction vectors yields $(\boldsymbol{r}_0,\boldsymbol{p}_j) = (\boldsymbol{r}_{j-1},\boldsymbol{p}_j)$, and we may compute each new component ζ_j without storing $\boldsymbol{r}_0$.

Thus, setting $\widehat{\boldsymbol{v}}_1 = \boldsymbol{r}_0$ in the coupled two-term formulation of the Hermitian Lanczos process (Algorithm 4.4.2) augmented with the updates

$$\zeta_m = (\boldsymbol{r}_{m-1},\boldsymbol{p}_m)/(A\boldsymbol{p}_m,\boldsymbol{p}_m), \quad \boldsymbol{x}_m = \boldsymbol{x}_{m-1} + \zeta_m\boldsymbol{p}_m, \quad \boldsymbol{r}_m = \boldsymbol{r}_{m-1} - \zeta_m A\boldsymbol{p}_m \tag{4.36}$$

in each step yields an implementation of the OR method. There is, however, still redundancy in generating both the Lanczos vectors and the residuals, since for the OR method these vectors are collinear. By suitably rescaling the direction vectors we may, to this end, replace the update for $\boldsymbol{p}_m$ in line 4 of Algorithm 4.4.2 by

$$\boldsymbol{p}_m = \boldsymbol{r}_{m-1} + \beta_{m-1}\boldsymbol{p}_{m-1}, \quad m = 2,3,\ldots. \tag{4.37}$$

Orthogonality of $\boldsymbol{r}_{m-1}$ against $\mathscr{K}_{m-1}(A,\boldsymbol{r}_0) = \text{span}\{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_{m-1}\}$ then implies that $(\boldsymbol{r}_{m-1},\boldsymbol{p}_m) = (\boldsymbol{r}_{m-1},\boldsymbol{r}_{m-1})$ in the numerator of μ_m. To obtain an expression for β_{m-1} we form the inner product of (4.37) with $A\boldsymbol{p}_{m-1}$, which yields

$$0 = (\boldsymbol{r}_{m-1},A\boldsymbol{p}_{m-1}) + \beta_{m-1}(A\boldsymbol{p}_{m-1},\boldsymbol{p}_{m-1}) = \frac{(\boldsymbol{r}_{m-1},\boldsymbol{r}_{m-2}-\boldsymbol{r}_{m-1})}{\zeta_{m-1}} + \beta_{m-1}(A\boldsymbol{p}_{m-1},\boldsymbol{p}_{m-1})$$

and, by orthogonality of the residuals,

$$\beta_m = \frac{(\boldsymbol{r}_m,\boldsymbol{r}_m)}{(\boldsymbol{r}_{m-1},\boldsymbol{r}_{m-1})}, \qquad m = 1,2\ldots.$$

These further simplifications now result in the most popular version of the conjugate gradient method as proposed by Hestenes & Stiefel (1952), which is given in Algorithm 4.4.5, in which we have, following convention, renamed the quantities ζ_m as α_m. This algo-

Algorithm 4.4.5 CG with coupled two-term recurrences.

1 $\boldsymbol{x}_0$ given, $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$, $\boldsymbol{p}_1 = \boldsymbol{r}_0$
2 **for** $m = 1,2,\ldots$
3 $\quad\alpha_m = (\boldsymbol{r}_{m-1},\boldsymbol{r}_{m-1})/(A\boldsymbol{p}_m,\boldsymbol{p}_m)$
4 $\quad\boldsymbol{x}_m = \boldsymbol{x}_{m-1} + \alpha_m\boldsymbol{p}_m$
5 $\quad\boldsymbol{r}_m = \boldsymbol{r}_{m-1} - \alpha_m A\boldsymbol{p}_m$
6 $\quad\beta_m = (\boldsymbol{r}_m,\boldsymbol{r}_m)/(\boldsymbol{r}_{m-1},\boldsymbol{r}_{m-1})$
7 $\quad\boldsymbol{p}_{m+1} = \boldsymbol{r}_m + \beta_m\boldsymbol{p}_m$
8 **end**

rithm, which performs one matrix-vector multiplication, two inner products and three vector updates in each step, requires the storage of the four vectors $\boldsymbol{x}_m$, $\boldsymbol{r}_m$, $\boldsymbol{p}_m$, and $A\boldsymbol{p}_m$.

An alternative approach for implementing the conjugate gradient method is to have the Lanczos process generate only the A-conjugate direction vectors to begin with. When A is Hermitian and positive definite, this can be achieved by carrying out the Lanczos process with the A-inner product in place of the original inner product $(\cdot, \cdot)$. Using the unnormalized Lanczos process (Algorithm 4.4.2), this leads to Algorithm 4.4.6, the conjugate gradient method based on a three-term recurrence for the direction vectors. This algorithm is somewhat more expensive than the two-term variant, performing one

Algorithm 4.4.6 CG with three-term recurrence for the direction vectors.

$\boldsymbol{x}_0$ given,$\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$, $\boldsymbol{p}_1 = \boldsymbol{r}_0$
for $m = 1, 2, \ldots$
  $\alpha_m = (\boldsymbol{r}_{m-1}, \boldsymbol{p}_m)/(A\boldsymbol{p}_m, \boldsymbol{p}_m)$
  $\boldsymbol{x}_m = \boldsymbol{x}_{m-1} + \alpha_m \boldsymbol{p}_m$
  $\boldsymbol{r}_m = \boldsymbol{r}_{m-1} - \alpha_m A\boldsymbol{p}_m$
  $\gamma_m = (A\boldsymbol{p}_m, A\boldsymbol{p}_m)/(A\boldsymbol{p}_m, \boldsymbol{p}_m)$
  $\delta_m = (A\boldsymbol{p}_m, \boldsymbol{p}_{m-1})/(A\boldsymbol{p}_{m-1}, \boldsymbol{p}_{m-1})$
  $\boldsymbol{p}_{m+1} = A\boldsymbol{p}_m - \gamma_m \boldsymbol{p}_m - \delta_m \boldsymbol{p}_{m-1}$
end

matrix-vector product, four inner products and three vector updates per step and requiring the storage of $\boldsymbol{x}_m, \boldsymbol{r}_m, \boldsymbol{p}_m, \boldsymbol{p}_{m-1}$, and $A\boldsymbol{p}_m$

Remark 4.4.2. As is easily verified, running the Lanczos process with the A-inner product generates a decomposition $AP_m = P_{m+1}\widetilde{H}_m^{(1)}$, with $\widetilde{H}_m^{(1)} = L_{m+1}^{\top}\widetilde{L}_m D_m$. This corresponds to applying one step of Rutishauser's LR-algorithm to the square tridiagonal matrix H_L associated with the termination index L. This is the key observation to relating the OR iterates with respect to the scale of inner products $(A^s\cdot, \cdot)$, $s = 0, 1, 2, \ldots$ as discovered by Gutknecht (1993*a*).

Finally, there is yet a third variant of the conjugate gradient method which is based on a three-term recurrence for the iterates and residuals which is given below as Algorithm, 4.4.7. This form was derived in Stiefel (1955) by using the three-term recurrence

Algorithm 4.4.7 CG with three-term recurrence for iterates and residuals.

$\boldsymbol{x}_0$ given, $\boldsymbol{x}_{-1} = \boldsymbol{x}_0, \boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0, \boldsymbol{r}_{-1} = \boldsymbol{r}_0, e_{-1} = 0$
for $m = 1, 2, \ldots$
  $q_k = (A\boldsymbol{r}_k, \boldsymbol{r}_k)/(\boldsymbol{r}_k, \boldsymbol{r}_k) - e_{k-1}$
  $\boldsymbol{r}_{k+1} = \boldsymbol{r}_k + 1/q_k\,[-A\boldsymbol{r}_k + e_{k-1}(\boldsymbol{r}_k - \boldsymbol{r}_{k-1})]$
  $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + 1/q_k\,[\boldsymbol{r}_k + e_{k-1}(\boldsymbol{x}_k - \boldsymbol{x}_{k-1})]$
  $e_k = q_k(\boldsymbol{r}_{k+1}, \boldsymbol{r}_{k+1})/(\boldsymbol{r}_k, \boldsymbol{r}_k)$
end

which generates the sequence of orthogonal polynomials associated with the inner product

$$(p, q) = (p(A)\boldsymbol{r}_0, q(A)\boldsymbol{r}_0)$$

on the space of polynomials. For an up-to-date account of the connections between orthogonal polynomials and Krylov subspace methods for Hermitian matrices, see the recent monograph by Fischer (1996). This form of the conjugate gradient method was also used in Rutishauser (1959), Reid (1971), Concus, Golub & O'Leary (1976), and Concus & Golub (1976). In particular, Rutishauser (1959) emphasized how this form reveals that CG belongs to the class of *gradient methods*, in which each correction vector is chosen from the span of all previous residuals. As we have mentioned before, this span coincides with the usual Krylov space unless a Galerkin breakdown occurs, which can be excluded in the Hermitian positive definite case. Algorithm 4.4.7 requires storing $\boldsymbol{x}_{k-1}$, $\boldsymbol{x}_k$, $\boldsymbol{r}_{k-1}$, $\boldsymbol{r}_k$, and $A\boldsymbol{r}_k$ and performs one matrix-vector multiplication, two inner products and two extended vector updates. It is thus roughly as expensive as Algorithm 4.4.6.

The Conjugate Residual Algorithm

Although the conjugate gradient method was by far the more popular method, an implementation of the MR method was already implicitly contained in the original Hestenes and Stiefel paper (Hestenes & Stiefel 1952). The first formulations as a method in its own right seems to be in Stiefel (1955), and later Rutishauser (1959). It has since then come to be known as the *conjugate residual (CR) method*, and it is the predecessor of the GCR method. The algorithm can be obtained by replacing the A-inner product in the CG algorithm by the A^2-inner product. In this case the error $A^{-1}\boldsymbol{b} - \boldsymbol{x}_m$ minimizes the A^2-norm over $\mathscr{K}_m(A, \boldsymbol{r}_0)$, which is equivalent with minimizing the original norm $(\cdot,\cdot)^{1/2}$ of the residual, hence this is the MR method with respect to the original inner product.

Making these substitutions in Algorithm 4.4.5 yields the conjugate residual method, depicted in Algorithm 4.4.8. The last line in Algorithm 4.4.8 indicates that the vectors

Algorithm 4.4.8 The conjugate residual method.

$\boldsymbol{x}_0$ given,$\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$, $\boldsymbol{p}_1 = \boldsymbol{r}_0$
for $m = 1, 2, \ldots$
  $\alpha_m = (A\boldsymbol{r}_{m-1}, \boldsymbol{r}_{m-1})/(A\boldsymbol{p}_m, A\boldsymbol{p}_m)$
  $\boldsymbol{x}_m = \boldsymbol{x}_{m-1} + \alpha_m \boldsymbol{p}_m$
  $\boldsymbol{r}_m = \boldsymbol{r}_{m-1} - \alpha_m A\boldsymbol{p}_m$
  $\beta_m = (A\boldsymbol{r}_m, \boldsymbol{r}_m)/(A\boldsymbol{r}_{m-1}, \boldsymbol{r}_{m-1})$
  $\boldsymbol{p}_{m+1} = \boldsymbol{r}_m + \beta_m \boldsymbol{p}_m$
  $A\boldsymbol{p}_{m+1} = A\boldsymbol{r}_m + \beta_m A\boldsymbol{p}_m$
end

$A\boldsymbol{p}_m$ can be updated using the vectors $A\boldsymbol{r}_m$ and therefore do not incur an additional matrix-vector multiplication per step. In addition, CR requires two inner products and four vector updates in each iteration and the storage of the vectors $\boldsymbol{x}_m, \boldsymbol{r}_m, \boldsymbol{p}_m, A\boldsymbol{p}_m$ and $A\boldsymbol{r}_m$. This is one more update and one additional vector to be stored than in the standard CG algorithm, hence, unless an application calls for the minimization of the original norm, the CG method is to be preferred. Comparison with Algorithm 4.4.1, derived by specializing the GCR method for Hermitian A, shows that the CR algorithm is identical.

We recall, however, that CR (and therefore Hermitian GCR) as well as the three variants of CG we have discussed may all fail if A Hermitian but not positive definite. In this case the only stable approach is the MINRES algorithm, which computes the MR approximations and allows the stable generation of the OR approximations whenever these exist.

4.4.3 The non-Hermitian Case

The attractive features of the conjugate gradient method are its short vector recurrence and its norm minimization property, and from the late 1970s onward much research was devoted to extending CG to non-Hermitian problems, until it was shown by Faber & Manteuffel (1984) that full recurrences were essentially unavoidable when minimizing with respect to a fixed norm, hence either one or the other property must be sacrificed when A is non-Hermitian.

The earliest reference in this context is Vinsome (1976), in which ORTHOMIN, a Krylov subspace method for non-Hermitian problems is proposed which features a full recurrence which is truncated after a small number of terms. Young & Jea (1980) propose an *idealized generalized conjugate gradient method* for non-Hermitian problems, which selects approximations $\boldsymbol{x}_m$ from $\boldsymbol{x}_0 + \mathscr{K}_m(A, \boldsymbol{r}_0)$ such that the associated error $\boldsymbol{e}_m = A^{-1}\boldsymbol{b} - \boldsymbol{x}_m$ satisfies the *Galerkin condition*

$$\langle \boldsymbol{e}_m, \boldsymbol{v} \rangle = 0 \qquad \text{for all } \boldsymbol{v} \in \mathscr{K}_m(A, \boldsymbol{r}_0).$$

Here $\langle \cdot, \cdot \rangle$ is an arbitrary sesquilinear form, which could be an inner product, in which case the error vectors minimize the associated norm over the shifted Krylov space. Besides ORTHOMIN, Young & Jea (1980) present two further algorithms ORTHODIR and ORTHORES which implement their idealized generalized conjugate gradient method under the general assumption that A is *positive real* with respect to $\langle \cdot, \cdot \rangle$, i.e., that

$$\langle A\boldsymbol{v}, \boldsymbol{v} \rangle > 0 \qquad \text{for all } \boldsymbol{v} \in \mathscr{H}.$$

In addition, ORTHOMIN and ORTHORES require that also $\langle \boldsymbol{v}, \boldsymbol{v} \rangle > 0$ for all $\boldsymbol{v} \in \mathscr{H}$.

Our emphasis in this section is the relation of these three algorithms to the three variants of CG and the situations in which they may break down before finding the solution. To simplify matters, we restrict ourselves to the case when $\langle \cdot, \cdot \rangle$ is an inner product. In the more general case of an arbitrary sesquilinear form the concept of orthogonality must be generalized to *one-sided* or *semi-orthogonality*, and the orthogonalization steps become slightly more complicated.

The ORTHOMIN algorithm is given in Algorithm 4.4.9. We observe that, as in Algorithm 4.4.5, the direction vectors $\{\boldsymbol{p}_m\}$ in ORTHOMIN are obtained by orthogonalizing the current residual against the previous direction vectors. In order for this procedure to generate a basis of $\mathscr{K}_m(A, \boldsymbol{r}_0)$, it is necessary that the residuals span the entire Krylov space at each step, otherwise the algorithm stalls because the Krylov space cannot be extended. This occurs whenever $\alpha_m = 0$, which, from lines 3 and 7 of Algorithm 4.4.9, occurs whenever

$$0 = \langle \boldsymbol{e}_{m-1}, \boldsymbol{p}_m \rangle = \langle \boldsymbol{e}_{m-1}, \boldsymbol{r}_{m-1} \rangle = \langle \boldsymbol{e}_{m-1}, A\boldsymbol{e}_{m-1} \rangle.$$

Algorithm 4.4.9 The ORTHOMIN algorithm for an inner product.

$\boldsymbol{x}_0$ given,$\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0, \boldsymbol{p}_1 = \boldsymbol{r}_0$
for $m = 1, 2, \ldots$
  $\alpha_m = \langle \boldsymbol{e}_{m-1}, \boldsymbol{p}_m \rangle / \langle \boldsymbol{p}_m, \boldsymbol{p}_m \rangle$
  $\boldsymbol{x}_m = \boldsymbol{x}_{m-1} + \alpha_m \boldsymbol{p}_m$
  $\boldsymbol{r}_m = \boldsymbol{r}_{m-1} - \alpha_m A\boldsymbol{p}_m$
  $\beta_{m,j} = \langle \boldsymbol{r}_m, \boldsymbol{p}_j \rangle / \langle \boldsymbol{p}_j, \boldsymbol{p}_j \rangle, \quad j = 1, \ldots, m$
  $\boldsymbol{p}_{m+1} = \boldsymbol{r}_m - \sum_{j=1}^{m} \beta_{m,j} \boldsymbol{p}_j$
end

By Theorem 4.2.2 this can be excluded for all initial residuals if and and only if $0 \notin W(A)$, where $W(A)$ is understood to be the field of values of A with respect to $\langle \cdot, \cdot \rangle$.

The algorithm ORTHODIR, given in Algorithm 4.4.10, computes the direction vectors by multiplying the last with A and orthogonalizing against all previous ones, which amounts to an unnormalized Arnoldi process. This way, an orthogonal basis of the Krylov space is always obtained, and no restriction is required of A. We also note that OR-

Algorithm 4.4.10 ORTHODIR.

$\boldsymbol{x}_0$ given,$\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0, \boldsymbol{p}_1 = \boldsymbol{r}_0$
for $m = 1, 2, \ldots$
  $\alpha_m = \langle \boldsymbol{e}_{m-1}, \boldsymbol{p}_m \rangle / \langle \boldsymbol{p}_m, \boldsymbol{p}_m \rangle$
  $\boldsymbol{x}_m = \boldsymbol{x}_{m-1} + \alpha_m \boldsymbol{p}_m$
  $\boldsymbol{r}_m = \boldsymbol{r}_{m-1} - \alpha_m A\boldsymbol{p}_m$
  $\beta_{m,j} = \langle A\boldsymbol{p}_m, \boldsymbol{p}_j \rangle / \langle \boldsymbol{p}_j, \boldsymbol{p}_j \rangle, \quad j = 1, \ldots, m$
  $\boldsymbol{p}_{m+1} = A\boldsymbol{p}_m - \sum_{j=1}^{m} \beta_{m,j} \boldsymbol{p}_j$
end

THOMIN is the straightforward generalization of the second variant of CG given in Algorithm 4.4.6.

In the same manner, ORTHORES, given below in Algorithm 4.4.11, is the generalization of the third CG variant (Algorithm 4.4.7). Instead of explicitly generating an orthogonal basis of $\mathscr{K}_m(A, \boldsymbol{r}_0)$, ORTHORES generates the new approximation from previous approximation and the last residual in such a way that $\langle \boldsymbol{e}_{m+1}, \boldsymbol{r}_j \rangle = 0$ for $j \leq m$. Just as ORTHOMIN, ORTHORES can stall if the residuals fail to span the maximal Krylov space, hence it must be required that A be definite with respect to $\langle \cdot, \cdot \rangle$.

The formulation of these three algorithms are those given in Ashby & Gutknecht (1993), and they correspond to those introduced in Young & Jea (1980) for the case that the sesquilinear form is an inner product. Of course, the inner product must be chosen in order that expressions like $\langle \boldsymbol{e}_{m-1}, \boldsymbol{p}_m \rangle$, which contain the unknown error vector, can be computed. The most popular choice (see e.g. (Elman 1982, Eisenstat et al. 1983, Saad & Schultz 1985) is $\langle \boldsymbol{u}, \boldsymbol{v} \rangle := (A\boldsymbol{u}, A\boldsymbol{v})$, in which case the error is A^*A-orthogonal to $\mathscr{K}_m$, or equivalently, the residual is orthogonal to $A\mathscr{K}_m$ with respect to the given inner product $(\cdot, \cdot)$, resulting in the MR method with respect to this inner product. In this case

Algorithm 4.4.11 ORTHORES.

1 $\boldsymbol{x}_0$ given, $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$
2 **for** $m = 0, 1, \ldots$
3 $\quad \gamma_{m,j} = \left[\langle \boldsymbol{r}_m, \boldsymbol{r}_j\rangle - \sum_{i=0}^{j-1} \gamma_{m,i}\langle \boldsymbol{e}_i, \boldsymbol{r}_j\rangle\right] / \langle \boldsymbol{e}_j, \boldsymbol{r}_j\rangle, \quad j = 0, \ldots, m$
4 $\quad \sigma_m = \left(\sum_{j=0}^{m} \gamma_{m,j}\right)^{-1}, \quad j = 0, \ldots, m$
5 $\quad \rho_{m,j} = \sigma_m \gamma_{m,j}$
6 $\quad \boldsymbol{x}_{m+1} = \sigma_m \boldsymbol{r}_m + \sum_{j=0}^{m} \rho_{m,j} \boldsymbol{x}_j$
7 $\quad \boldsymbol{r}_{m+1} = -\sigma_m A\boldsymbol{r}_m + \sum_{j=0}^{m} \rho_{m,j} \boldsymbol{r}_j$
8 **end**

ORTHOMIN algorithm coincides with the GCR algorithm (Algorithm 4.2.2) for Krylov spaces, and ORTHODIR with the modification of GCR which uses the Arnoldi basis of $\mathscr{K}_m(A, A\boldsymbol{r}_0)$ and is therefore immune to premature termination. Here vectors $A\boldsymbol{p}_m$ form the orthogonal basis of $A\mathscr{K}_m(A, \boldsymbol{r}_0)$; the vectors, however, are not normalized in ORTHODIR. A stabilized version of this algorithm is discussed in Rozložník & Strakoš (1996).

A matrix formulation of ORTHOMIN, ORTHODIR and ORTHORES for the case of a general inner product is given in Ashby & Gutknecht (1993).

These algorithms have been more or less abandoned in favor of GMRES, FOM and their restarted variants (see Chapter 5), as well as QMR and other methods with short recurrences. The advantage of GMRES lies in the fact that only the basis generated by the Arnoldi process needs to be generated and stored, and that the algorithm does not terminate prematurely when two residuals coincide, as does GCR. Numerical results contained in Rozložník & Strakoš (1996) indicate that ORTHOMIN can be made as numerically stable as GMRES provided the basis vectors are normalized and a stable orthogonalization procedure such as modified Gram-Schmidt is used in place of the classical Gram-Schmidt procedure used in the original formulation.

4.5 Inner Products and Short Recurrences

We have seen in Chapter 2 that computing MR and OR approximations requires the construction of an orthogonal basis of the approximation or residual space, and that the orthogonalization procedure requires storing and forming inner products with the entire basis. In Section 2.3.3 we saw that an alternative is to settle for a possibly nonorthogonal basis of the residual space, which leads to the QMR and QOR methods, which are MR and OR methods with respect to an inner product depending on the basis.

For Krylov subspace methods this reduces to the task of constructing a basis of the Krylov space. As mentioned in Section 4.2.2, the QMR and QOR methods use the look-ahead Lanczos process for this purpose.

If, however, the MR and OR approximations are desired with respect to a fixed inner product, then there is generally no alternative to the (full) Arnoldi process. The fact that such an orthogonal basis can be generated with a three-term recursion when A

is Hermitian raises the question of what class of matrices permits a shorter than full recursion length in the Arnoldi process, and this question was answered by Faber & Manteuffel (1984). In the remainder of this section we shall present the main ideas of this result and relate these to our framework for MR and OR methods.

In the Arnoldi decomposition $AV_m = V_{m+1}\widetilde{H}_m$ the orthonormal basis vectors $\boldsymbol{v}_m$ satisfy the recursion

$$\eta_{m+1,m}\boldsymbol{v}_{m+1} = A\boldsymbol{v}_m - \eta_{m,m}\boldsymbol{v}_m - \cdots - \eta_{1,m}\boldsymbol{v}_1, \qquad m = 1, 2, \ldots, L-1$$

(cf. (2.39), (2.42) and (4.3)) with $\eta_{j,m} = (A\boldsymbol{v}_m, \boldsymbol{v}_j), j = 1, \ldots, m$. Generating $\boldsymbol{v}_{m+1}$ thus requires evaluating a recursion containing $m+1$ terms, and for this recursion to reduce contain to at most, say, s terms, one would have to guarantee $(A\boldsymbol{v}_m, \boldsymbol{v}_j) = 0$ for $j < m-(s-2)$. A sufficient condition for this to happen is given by the following Lemma.

Lemma 4.5.1. *If A is such that*

$$A^*\boldsymbol{v} \in \mathscr{K}_{s-1}(A, \boldsymbol{v}) \qquad \textit{for all } \boldsymbol{v} \in \mathscr{H}, \tag{4.38}$$

then the Arnoldi process for A simplifies to an s-term recursion.

Proof. Under the given assumption, we have at step m of the Arnoldi process

$$(A\boldsymbol{v}_m, \boldsymbol{v}_j) = (\boldsymbol{v}_m, A^*\boldsymbol{v}_j) = (\boldsymbol{v}_m, q(A)\boldsymbol{v}_j), \qquad \text{with } q \in \mathscr{P}_{s-2}.$$

From $\boldsymbol{v}_j \in \mathscr{K}_j(A, \boldsymbol{r}_0)$ we conclude $q(A)\boldsymbol{v}_j \in \mathscr{K}_{j+s-2}(A, \boldsymbol{r}_0)$, and therefore $(\boldsymbol{v}_m, A^*\boldsymbol{v}_j) = 0$ for $j + s - 2 < m$, i.e., for $j < m-(s-2)$. □

Lemma 4.5.2 shows that (4.38) is a rather strong requirement on A.

Lemma 4.5.2. *Condition* (4.38) *is satisfied if and only if A is normal and $A^* = q(A)$ for some $q \in \mathscr{P}_{s-2}$.*

Proof. Only necessity needs to be shown. To this end, note first that $A^*\boldsymbol{v} = q(A)\boldsymbol{v}$ for all $\boldsymbol{v} \in \mathscr{H}$ implies that eigenvectors of A are also eigenvectors of A^*, which is a characterization of normality (see Horn & Johnson (1985, Section 2.5)). If $d = d(A)$ denotes the degree of the minimal polynomial of A, then A, as a normal matrix, possesses d distinct eigenvalues $\{\lambda_j\}_{j=1}^d$. As is easily verified, the interpolating polynomial $q \in \mathscr{P}_{d-1}$ such that $q(\lambda_j) = \overline{\lambda_j}$, $j = 1, \ldots, d$, satisfies $A^* = q(A)$, and it remains to show that $d-1 \le s-2$. This can be seen by considering any vector $\boldsymbol{v}$ which generates a Krylov space with respect to A of maximal dimension d, in which case $A^*\boldsymbol{v} = q(A)\boldsymbol{v}$ is a representation of $A^*\boldsymbol{v}$ in terms of the linearly independent vectors $\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^{d-1}\boldsymbol{v}$. Since the degree of q is exactly $d-1$, the coefficient of $A^{d-1}\boldsymbol{v}$ herein is nonzero. By (4.38) we also have $A^*\boldsymbol{v} = \widetilde{q}(A)\boldsymbol{v}$ with $\widetilde{q} \in \mathscr{P}_{s-2}$ and, if $s-2 < d-1$, this would imply that $A^*\boldsymbol{v}$ could be represented without $A^{d-1}\boldsymbol{v}$, a contradiction. □

Besides the sufficient condition given in Lemma 4.5.2, it is clear that the Arnoldi process reduces to at most an s-term recursion if the degree of the minimal polynomial of A is at most s. Faber & Manteuffel (1984) show by way of a continuity argument that these conditions are also necessary for an s-term Arnoldi procedure, and we cite their important result in

Theorem 4.5.3. *The Arnoldi process for a matrix $A \in \mathbb{C}^{n\times n}$ reduces to an s-step recursion for all initial vectors if and only if either*

(a) the minimal polynomial of A has degree at most s, or

(b) A^* is a polynomial in A of degree at most $s-2$.

It is interesting, in light of this result, to characterize the matrices for which a three-term recursion exists as in the Hermitian case. By Theorem 4.5.3 and Lemma 4.5.2, it is necessary that A be normal and A^* be a linear polynomial in A. A simple calculation (we refer to (Faber & Manteuffel 1984) or (Greenbaum et al. 1997, page 101)) shows this implies the following three alternatives for A: either A is a multiple of the identity, or A is Hermitian, or A is of the form

$$A = e^{i\theta}\left(\alpha I + B\right), \qquad \theta \in [0, 2\pi), \quad \alpha \in \mathbb{R}, \quad B = -B^*.$$

This class of matrices is usually characterized as all normal matrices whose spectrum lies on a line in the complex plane.

Thus, the generalizations of CG with three-term recurrences had already been discovered by the time the Faber-Manteuffel result appeared: besides the methods for Hermitian A, methods based on the Lanczos process for matrices of the type $I + B$ with B skew-Hermitian were proposed by Concus & Golub (1976) and Widlund (1978).

4.6 Inner Products and Preconditioning

Preconditioning refers to the practice of solving a modified but equivalent linear system to (1.1) for which the iterative solution method converges more quickly. In practice MR and OR methods are nearly always applied in combination with preconditioning, as the extra effort this requires is usually outweighed by savings due to faster convergence. A discussion of preconditioning strategies is beyond the scope of the present work (we refer to the monographs Saad (1996) and Greenbaum (1997) and the references given there), but we do point out in this section how the Krylov space structure as well as the norm being minimized are affected by preconditioning.

In *left preconditioning*, the system is multiplied by a nonsingular operator M^{-1}—M being the *preconditioner*—yielding the equivalent *preconditioned system*

$$M^{-1}A\boldsymbol{x} = M^{-1}\boldsymbol{b}. \tag{4.39}$$

In *right preconditioning*, the change of variables $\widehat{\boldsymbol{x}} := M\boldsymbol{x}$ leads to the preconditioned system

$$AM^{-1}\widehat{\boldsymbol{x}} = \boldsymbol{b}. \tag{4.40}$$

Particularly in the Hermitian case, preconditioning is often applied from both sides using a factorization $M = M_1M_2$, resulting the preconditioned system

$$M_1^{-1}AM_2^{-1}\widehat{\boldsymbol{x}} = M_1^{-1}\boldsymbol{b}, \quad \widehat{\boldsymbol{x}} = M_2\boldsymbol{x}. \tag{4.41}$$

Proposition 4.6.1. *When a Krylov subspace method is applied to one of the preconditioned systems* (4.39), (4.40) *or* (4.41), *then the associated sequence of correction spaces is*

$$\mathscr{C}_m = M^{-1}\mathscr{K}_m(AM^{-1}, \boldsymbol{r}_0) = \mathscr{K}_m(M^{-1}A, M^{-1}\boldsymbol{r}_0), \qquad m = 1, 2, \ldots .$$

Proof. We consider the two-sided case: the initial residual of the preconditioned system is $\widehat{\boldsymbol{r}}_0 = M_1^{-1}\boldsymbol{r}_0$. The approximation $\widehat{\boldsymbol{x}}_m = M_2\boldsymbol{x}_m$ thus lies in

$$\begin{aligned} &\widehat{\boldsymbol{x}}_0 + \mathscr{K}_m(M_1^{-1}AM_2^{-1}, M_1^{-1}\boldsymbol{r}_0) \\ &\quad = \widehat{\boldsymbol{x}}_0 + \operatorname{span}\{M_1^{-1}\boldsymbol{r}_0, M_1^{-1}AM^{-1}\boldsymbol{r}_0, \ldots, M_1^{-1}(AM^{-1})^{m-1}\boldsymbol{r}_0\}. \end{aligned} \tag{4.42}$$

Switching back to the original variable $\boldsymbol{x} = M_2^{-1}\widehat{\boldsymbol{x}}$ yields the assertion. □

We now turn to the optimality properties of preconditioned Krylov subspace iterations. Methods based on the Hermitian Lanczos process require that the preconditioned system again be Hermitian, which can be achieved e.g. for $M_1 = M_2^*$, is sometimes referred to as *symmetric preconditioning*. A typical example of this approach is the use of an incomplete Cholesky factorization $M = LL^* \approx A$ as the preconditioner and setting $M_1 = L$ and $M_2 = L^*$. Such a Hermitian splitting, however, is not always available, in particular when the action of the preconditioner is given by some approximate solution algorithm such as a multigrid cycle (see e.g. Hackbush (1985)). Another possibility is to change the inner product in such a way that the preconditioned system becomes selfadjoint with respect to the new inner product. If the latter is given by

$$(\boldsymbol{x}, \boldsymbol{y})_B := (B\boldsymbol{x}, \boldsymbol{y}),$$

with a Hermitian operator B, then A is selfadjoint with respect to the B-inner product if, and only if, it satisfies $BA = (BA)^*$, where the adjoint is understood with respect to the original inner product $(\cdot, \cdot)$. For left preconditioning with $M = M^*$ this leads to $BM^{-1}A = AM^{-1}B$, which is satisfied e.g. by choosing $B = M$. The analogous choice for right preconditioning is $B = M^{-1}$.

Both symmetric preconditioning and changing the inner product will generally cause Krylov subspace algorithms to minimize different quantities than their unpreconditioned counterparts. A notable exception is the conjugate gradient method. Recall that the unpreconditioned variant minimizes $\|\boldsymbol{e}_m\|_A = \|\boldsymbol{r}_m\|_{A^{-1}}$ with the usual notation $\boldsymbol{e}_m = \boldsymbol{x} - \boldsymbol{x}_m$ and $\boldsymbol{r}_m = \boldsymbol{b} - A\boldsymbol{x}_m$. For symmetric preconditioning with preconditioner $M = QQ^*$, the preconditioned matrix becomes $\widehat{A} = Q^{-1}AQ^{-*}$ and the preconditioned quantities are $\widehat{\boldsymbol{e}}_m = Q^*\boldsymbol{e}_m$ and $\widehat{\boldsymbol{r}}_m = Q^{-1}\boldsymbol{r}_m$. Therefore, the symmetrically preconditioned conjugate gradient method minimizes

$$(\widehat{A}\widehat{\boldsymbol{e}}_m, \widehat{\boldsymbol{e}}_m) = (Q^{-1}AQ^{-*}Q^*\boldsymbol{e}_m, Q^*\boldsymbol{e}_m) = (A\boldsymbol{e}_m, \boldsymbol{e}_m)$$

and

$$(\widehat{A}^{-1}\widehat{\boldsymbol{r}}_m, \widehat{\boldsymbol{r}}_m) = (Q^*A^{-1}QQ^{-1}\boldsymbol{r}_m, Q^{-1}\boldsymbol{r}_m) = (A^{-1}\boldsymbol{r}_m, \boldsymbol{r}_m),$$

i.e., the same quantities regardless of which preconditioner is used. The same result is obtained for left preconditioning using the inner product defined by $B = M$, i.e.,

$$(M^{-1}A\boldsymbol{e}_m, \boldsymbol{e}_m)_M = (A\boldsymbol{e}_m, \boldsymbol{e}_m), \qquad ((M^{-1}A)^{-1}\widehat{\boldsymbol{r}}_m, \widehat{\boldsymbol{r}}_m)_M = (A^{-1}\boldsymbol{r}_m, \boldsymbol{r}_m),$$

and similarly for right preconditioning and the choice $B = M^{-1}$.

In contrast, the quantity minimized by MR methods does depend on the preconditioner. For symmetric preconditioning, these algorithms minimize

$$\|\widehat{\boldsymbol{r}}_m\| = (Q^{-1}\boldsymbol{r}_m, Q^{-1}\boldsymbol{r}_m)^{1/2} = (Q^{-*}Q^{-1}\boldsymbol{r}_m\boldsymbol{r}_m)^{1/2} = \|\boldsymbol{r}_m\|_{M^{-1}}$$

and for left preconditioning in the M-inner product, the corresponding quantity is

$$\|\widehat{\boldsymbol{r}}_m\|_M = (MM^{-1}\boldsymbol{r}, M^{-1}\boldsymbol{r})^{1/2} = \|\boldsymbol{r}_m\|_{M^{-1}}.$$

The right preconditioned case with $B = M^{-1}$ leads directly to the minimization of $\|\boldsymbol{r}_m\|_{M^{-1}}$. Thus, Hermitian formulations of the minimal residual method all minimize the preconditioner-dependent quantity $\|\boldsymbol{r}_m\|_{M^{-1}}$.

In the non-Hermitian case, changing the inner product generally brings no advantage for the following reasons: first, if A is non-Hermitian, then an effective preconditioner M—whose inverse should approximate A^{-1}—will generally be non-Hermitian as well, and thus fail to define an inner product. An exception occurs in the real case when the symmetric part of M is positive definite, in which case the inner product defined by M coincides with that defined by its symmetric part. Moreover, finding an inner product in terms of which the preconditioned matrix is selfadjoint is generally not feasible, so that a modified inner product will in general not lead to short recurrences as in the Hermitian case, but can at most affect the quantities being minimized.

For MR methods these quantities depend only on the preconditioning used from the left: if the preconditioned matrix is $\widehat{A} = M_1^{-1}AM_2^{-1}$, then the MR method minimizes the preconditioned residual

$$\|\widehat{\boldsymbol{r}}_m\| = (M_1^{-1}\boldsymbol{r}_m, M_1^{-1}\boldsymbol{r}_m)^{1/2} = \|\boldsymbol{r}_m\|_{(M_1M_1^*)^{-1}}.$$

As a consequence, unless this norm has some special relevance to the specific problem being solved, preconditioning from the right is to be recommended for these algorithms, since the resulting residuals are then minimal with respect to the original norm $\|\cdot\|$.

Methods using non-orthogonal bases such as those based on the non-Hermitian Lanczos process minimize a norm which depends on the basis of the Krylov space generated in the course of the algorithm (cf. (2.68)). Specifically, if the non-Hermitian Lanczos algorithm applied to the preconditioned system yields the basis $V_L = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_L]$ such that $M_1^{-1}AM_2^{-1}V_L = V_LH_L$ with termination index L, then the quantity minimized at step m is

$$\|M_1^{-1}\boldsymbol{r}_m\|_V = \|\widehat{\boldsymbol{r}}_m\|_V = \|V_L\boldsymbol{s}_m\|_V = \|\boldsymbol{s}_m\|,$$

where $\boldsymbol{s}_m$ is the coefficient vector of the preconditioned residual $\widehat{\boldsymbol{r}}_m$ at step m with respect to the basis V_L.

Chapter 5

Truncated and Restarted Krylov Subspace Methods

We have seen in Chapter 4 that, for Krylov subspace MR and OR methods with respect to a *given* inner product norm, full recurrences involving orthogonalization against all previously generated basis vectors cannot be circumvented. For many applications, in particular discretizations of partial differential equations in three space dimensions, this exceeds the capacities of even the most powerful current computing facilities. Therefore, from the outset (cf. Vinsome (1976), Young & Jea (1980)), it has been suggested to run methods requiring full recurrences in either a *truncated* or *restarted* fashion, in both of which only a fixed maximal number m of basis vectors is stored.

In the truncated version each new basis vector is orthogonalized only against the m previously generated basis vectors. This results in a *truncated Arnoldi decomposition* $AV_m = V_{m+1}\widetilde{H}_m$, in which the Hessenberg matrix $\widetilde{H}_m$ is banded with only $m+1$ nonvanishing diagonals, namely those with indices $-1, 0, \ldots, m$ (cf. Section 2.3.3). This permits the construction of $(m+1)$-term update formulas for the approximations and residuals (see, e.g., Saad (1996)). The basis vectors of V_{m+1}, however, are now only locally orthogonal, i.e., orthogonality of each $\boldsymbol{v}_j$ is enforced only against $\boldsymbol{v}_{j-m}, \boldsymbol{v}_{j-m+1}, \ldots, \boldsymbol{v}_{j+m}$. As long as the basis vectors remain linearly independent—and this is so until the termination index L is reached, at which the Krylov space first becomes invariant—the resulting truncated MR and OR methods fall into the category of QMR and QOR methods discussed in Sections 2.3.3 and 4.2.2. In contrast with the (look-ahead) Lanczos process used in Freund and Nachtigal's QMR algorithm (Freund & Nachtigal 1991), this method of basis-generation using a truncated Arnoldi process is unable to detect the invariance of the Krylov space, and the resulting methods are strictly speaking no longer MR or OR methods once the iteration index exceeds L. This is, however, an academic distinction, since Krylov spaces rarely become exactly invariant at index L due to roundoff, and L is usually much larger than the number of iteration steps one is willing to carry out.

Alternatively, *restarting* refers to beginning the entire iteration process anew every m steps of the full iteration, resulting in *restart cycles* consisting of m steps each, in which the final approximation of each cycle serves as the initial approximation of the next. Restarting seems to be the more widely used of these two shortcuts, possibly because it is somewhat easier to implement, and we accordingly devote the most attention to restarted methods in this chapter.

Of course, truncating and restarting can severely impair the convergence behavior of MR and OR methods. As a simple example, consider applying a restarted MR method to the cyclic forward shift in n dimensions: if the initial residual is any unit coordinate vector, the residual norm of the MR approximations remains identically one until the last step. Restarting after any number $m < n$ of steps will yield the same residual vector one started out with, hence no progress is achieved. Although such exact stagnation is rare in practice, it is often observed that the method suffers from near stagnation after the first few restart cycles. The material of this section is an expanded version of Eiermann, Ernst & Schneider (1999).

5.1 Stagnation of Restarted MR Methods

For a given sequence of correction spaces $\{\mathscr{C}_k\}_{k=1,2,\dots,L}$, the MR method succeeds in decreasing the initial residual after the first m steps if and only if $P_{A\mathscr{C}_m}\boldsymbol{r}_0 \neq \mathbf{0}$. Conversely, the initial residual vectors $\boldsymbol{r}_0$ for which the MR method makes no progress during the first m steps, i.e., for which $\|\boldsymbol{r}_m^{\mathrm{MR}}\| = \|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\| = \cdots = \|\boldsymbol{r}_0\|$, are characterized by $\boldsymbol{r}_0 \perp A\mathscr{C}_m$. There are several obvious and well-known equivalent descriptions of this situation, which is commonly known as *stagnation* or *stalling*.

Proposition 5.1.1. *The following statements are equivalent:*

(a) $\|\boldsymbol{r}_0\| = \|\boldsymbol{r}_1^{MR}\| = \cdots = \|\boldsymbol{r}_m^{MR}\|$, *i.e. the MR method stagnates for the first* m *steps.*

(b) $\boldsymbol{r}_0 \perp A\mathscr{C}_m$.

(c) $\sin\measuredangle(\boldsymbol{r}_0, A\mathscr{C}_m) = 1$.

(d) $\sin\measuredangle(\boldsymbol{r}_{k-1}, A\mathscr{C}_k) = 1$ *for all* $k = 1, 2, \dots, m$.

(e) The Hessenberg matrices H_k, $k = 1, 2, \dots, m$, *of* (4.3) *are singular.*

(f) The first row of H_m *is the zero vector.*

The equivalence of (a), (b), (c), (d) and (e) is an immediate consequence of the results of Chapter 2, together with

$$\sin\measuredangle(\boldsymbol{r}_0, A\mathscr{C}_m) = \frac{\|(I - P_{A\mathscr{C}_m})\boldsymbol{r}_0\|}{\|\boldsymbol{r}_0\|} = \prod_{k=1}^{m} \sin\measuredangle(\boldsymbol{r}_{k-1}, A\mathscr{C}_k).$$

That the characterization (f) (which first appears in Strikwerda & Stodder (1995)) is equivalent to (b) follows from the fact that the entries $\eta_{1,k}$ in the first row of H_m are given by $\eta_{1,k} = (A\boldsymbol{c}_k, \boldsymbol{r}_0)/\beta$ (cf. (3.10) and (4.4)). One also could have noticed that any unreduced Hessenberg matrix H_m, whose leading principal submatrices H_k ($k = 1, 2, \dots, k$) are all singular, must have a vanishing first row, hence (e) and (f) are equivalent.

In the context of Krylov subspace methods the statements of Proposition 5.1.1, or rather their negations, are often used to characterize the convergence of the restarted MR method (Eisenstat et al. 1983, Joubert 1994*a*, Saad 1996, Saad 1997*b*, Greenbaum 1997).

5.2 General Augmentation Strategies

The problem of stagnation or near-stagnation discussed in the preceding section results from a Krylov space which extends in the 'wrong directions' with regard to approximating $\boldsymbol{r}_0$, and this calls for some device for enhancing these insufficient approximation properties. The question of how to *augment* a given Krylov space with this goal in mind is the same as that of how to extend any given correction space so as to yield a strong residual reduction. In practice, this task usually arises in the following form: Given a correction space $\mathscr{C} \subset \mathscr{H}$, select vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ such that the augmented correction space $\widetilde{\mathscr{C}} := \mathscr{C} + \operatorname{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k\}$ has better correction properties, or, what is the same, that $\widetilde{\mathscr{W}} = A\mathscr{C} + \operatorname{span}\{A\boldsymbol{a}_1, \ldots, A\boldsymbol{a}_k\}$ contains a better approximation to $\boldsymbol{r}_0$. In view of (2.5) and (3.1), we have $\|\boldsymbol{r}^{\mathrm{MR}}\| = \sin\measuredangle(\boldsymbol{r}_0, A\mathscr{C})\|\boldsymbol{r}_0\|$, hence the correction properties of $\mathscr{C}$ are quantified by the angle $\varphi := \measuredangle(\boldsymbol{r}_0, A\mathscr{C})$ between $\boldsymbol{r}_0$ and $A\mathscr{C}$. $\mathscr{C}$ is optimal, i.e., $\boldsymbol{r}^{\mathrm{MR}} = \boldsymbol{0}$ if and only if $\varphi = 0$. The worst case, in which the optimal correction from $\mathscr{C}$ is the null vector (i.e., $\boldsymbol{r}^{\mathrm{MR}} = \boldsymbol{r}_0$), occurs precisely for $\varphi = \pi/2$, hence $\boldsymbol{r}_0 \perp A\mathscr{C}$, which is one of the characterizations given in Proposition 5.1.1 for stalling. Before turning to some of the methods proposed in the literature, we first address some general questions related to this issue.

It has often been suggested, primarily in the context of Krylov subspace methods, that it is a desirable goal that the correction space $\mathscr{C}$ be either nearly A-invariant or contain a nearly A-invariant subspace, usually spanned by a few approximate eigenvectors of A. Clearly, if a given correction space $\mathscr{C}$ which contains the initial residual $\boldsymbol{r}_0$—as do e.g. all Krylov spaces—is exactly A-invariant, then $\varphi = 0$ and the MR approximation with respect to $\mathscr{C}$ yields the exact solution. If only a subspace of $\mathscr{C}$ is A-invariant or nearly so in the sense that its image under A lies at a small angle to an A-invariant subspace $\mathscr{U}$, Proposition 5.2.1 shows that the MR residual with respect to $\mathscr{C}$ then has a small component in the direction of $\mathscr{U}$.

Proposition 5.2.1. *Given a correction space $\mathscr{C}$, let $\mathscr{U} \subset \mathscr{C}$ denote a subspace such that $\sin\measuredangle(\mathscr{U}, A\mathscr{U}) \leq \varepsilon$. Then the MR residual $\boldsymbol{r}^{MR}$ with respect to $\mathscr{C}$ satisfies $\|P_{\mathscr{U}}\boldsymbol{r}^{MR}\| \leq \varepsilon\|\boldsymbol{r}_0\|$.*

Proof. The assertion follows from $P_{\mathscr{U}}\boldsymbol{r}^{\mathrm{MR}} = P_{\mathscr{U}}(I - P_{A\mathscr{C}})\boldsymbol{r}_0$ and $\|P_{\mathscr{U}}(I - P_{A\mathscr{C}})\| \leq \|P_{\mathscr{U}}(I - P_{A\mathscr{U}})\| = \sin\measuredangle(\mathscr{U}, A\mathscr{U}) \leq \varepsilon$. □

In particular, if $\mathscr{C}$ contains an exactly invariant subspace $\mathscr{U}$, then the MR approximation removes the components of the initial residual in the direction of $\mathscr{U}$ completely. Of course, this may be only of limited use if $\|(I - P_{\mathscr{U}})\boldsymbol{r}_0^{\mathrm{MR}}\|/\|\boldsymbol{r}_0\|$ is large, i.e., if $\mathscr{U}$ does not contain a good approximation of $\boldsymbol{r}_0$. In short, the existence of A-invariant subspaces of $\mathscr{C}$ per se need not be beneficial.

In Lemma 4.3.2 we already proved that if $\mathscr{C} = \mathscr{K}_m(A, \boldsymbol{r}_0)$ is a Krylov subspace, then it cannot contain an A-invariant subspace $\mathscr{U}$ unless $\mathscr{K}_m(A, \boldsymbol{r}_0)$ is itself A-invariant, i.e., $\mathscr{K}_m(A, \boldsymbol{r}_0) = \mathscr{K}_L(A, \boldsymbol{r}_0)$. Obviously, augmenting $\mathscr{K}_m(A, \boldsymbol{r}_0)$ by $\operatorname{span}\{A^m\boldsymbol{r}_0, \ldots, A^{L-1}\boldsymbol{r}_0\}$ leads to the new correction space $\mathscr{K}_L(A, \boldsymbol{r}_0)$ which is A-invariant. We now show that there is no 'faster' way to augment $\mathscr{K}_m(A, \boldsymbol{r}_0)$ to an A-invariant space.

Proposition 5.2.2. *Let $\widetilde{\mathscr{C}}$ be an A-invariant space which contains $\mathscr{K}_m(A, \boldsymbol{r}_0)$. Then $\widetilde{\mathscr{C}}$ also contains $\mathscr{K}_L(A, \boldsymbol{r}_0)$.*

Proof. By $\mathscr{U}_0 := \cap\{\mathscr{U} : \mathscr{U} \text{ is an } A\text{-invariant subspace with } \mathscr{K}_m \subseteq \mathscr{U}\}$ we denote the smallest A-invariant subspace which contains $\mathscr{K}_m$. By definition, $\mathscr{U}_0 \subseteq \mathscr{K}_L$. On the other hand, since $\mathscr{U}_0$ contains $\boldsymbol{r}_0$ and is invariant under A, it must also contain $A^m\boldsymbol{r}_0$ for very $m = 0, 1, \ldots$, i.e., $\mathscr{K}_L \subseteq \mathscr{U}_0$. □

Proposition 5.2.2 should not lead to the conclusion that it is useless to augment a Krylov subspace $\mathscr{C} = \mathscr{K}_m$ by an A-invariant subspace $\mathscr{U}$. After all by Proposition 5.2.1, the MR residual with respect to $\widetilde{\mathscr{C}} = \mathscr{C} + \mathscr{U}$ contains no component in the direction of $\mathscr{U}$.

We show next that the MR approximation with respect to a Krylov space augmented by an invariant space coincides with the MR approximation with respect to another Krylov subspace associated with a 'smaller' linear system.

Lemma 5.2.3. *Let $\widetilde{\boldsymbol{r}}^{MR}$ denote the MR residual with respect to $\widetilde{\mathscr{C}} = \mathscr{K}_m(A, \boldsymbol{r}_0) + \mathscr{U}$, where $\mathscr{U}$ is an A-invariant subspace. Set further $A_{\mathscr{U}^\perp} := P_{\mathscr{U}^\perp}AP_{\mathscr{U}^\perp}$ and let $\boldsymbol{r}^{MR}$ denote the residual of the MR approximation for $A_{\mathscr{U}^\perp}\boldsymbol{x} = P_{\mathscr{U}^\perp}\boldsymbol{r}_0$ with respect to the correction space $\mathscr{K}_m(A_{\mathscr{U}^\perp}, P_{\mathscr{U}^\perp}\boldsymbol{r}_0)$ and a zero initial approximation. Then there holds*

$$\widetilde{\boldsymbol{r}}^{MR} = \boldsymbol{r}^{MR} \quad \textit{or, equivalently,} \quad P_{\mathscr{U}}\widetilde{\boldsymbol{r}}^{MR} = \boldsymbol{0} \textit{ and } P_{\mathscr{U}^\perp}\widetilde{\boldsymbol{r}}^{MR} = \boldsymbol{r}^{MR}.$$

Proof. As in Section 3.3 we split the computation of $\widetilde{\boldsymbol{r}}^{\mathrm{MR}}$ into two subtasks and write, using that $\mathscr{U}$ is A-invariant,

$$\widetilde{\boldsymbol{r}}^{\mathrm{MR}} = (I - P_{\mathscr{U}})\boldsymbol{r}_0 - P_{\mathscr{Z}}(I - P_{\mathscr{U}})\boldsymbol{r}_0 = (I - P_{\mathscr{Z}})P_{\mathscr{U}^\perp}\boldsymbol{r}_0,$$

where $\mathscr{Z} = (I - P_{\mathscr{U}})A\mathscr{K}_m(A, \boldsymbol{r}_0) = P_{\mathscr{U}^\perp}A\mathscr{K}_m(A, \boldsymbol{r}_0) \subseteq \mathscr{U}^\perp$. This implies $P_{\mathscr{U}}P_{\mathscr{Z}} = O$, and hence $P_{\mathscr{U}}\widetilde{\boldsymbol{r}}^{\mathrm{MR}} = \boldsymbol{0}$ (which we could also have deduced directly from Proposition 5.2.1). By Lemma 4.3.3, $P_{\mathscr{U}^\perp}A\mathscr{K}_m(A, \boldsymbol{r}_0) = A_{\mathscr{U}^\perp}\mathscr{K}_m(A_{\mathscr{U}^\perp}, P_{\mathscr{U}^\perp}\boldsymbol{r}_0)$, and therefore

$$\widetilde{\boldsymbol{r}}^{\mathrm{MR}} = (I - P_{P_{\mathscr{U}^\perp}A\mathscr{K}_m(A,\boldsymbol{r}_0)})P_{\mathscr{U}^\perp}\boldsymbol{r}_0 = (I - P_{A_{\mathscr{U}^\perp}\mathscr{K}_m(A_{\mathscr{U}^\perp}, P_{\mathscr{U}^\perp}\boldsymbol{r}_0)})P_{\mathscr{U}^\perp}\boldsymbol{r}_0,$$

which identifies $\widetilde{\boldsymbol{r}}^{\mathrm{MR}}$ as the residual of the MR approximation for $A_{\mathscr{U}^\perp}\boldsymbol{x} = P_{\mathscr{U}^\perp}\boldsymbol{r}_0$ with respect to the Krylov space $\mathscr{K}_m(A_{\mathscr{U}^\perp}, P_{\mathscr{U}^\perp}\boldsymbol{r}_0)$. □

In this sense, adding the invariant space $\mathscr{U}$ to a Krylov correction space eliminates the space $\mathscr{U}$ from the problem and reduces the original equation (1.1) to one on the smaller space $\mathscr{U}^\perp$.

A different strategy for enriching correction spaces common to many inner-outer iteration schemes is based on the following trivial observation: Suppose that, for a given correction space $\mathscr{C}$ and associated residual space $\mathscr{V} = \mathrm{span}\{\boldsymbol{r}_0\} + A\mathscr{C}$, we are able to solve the equation $A\boldsymbol{c} = \boldsymbol{r}$ for some $\boldsymbol{r} \in \mathscr{V}$. Such an $\boldsymbol{r}$ has a representation $\boldsymbol{r} = \boldsymbol{r}_0 - A\widetilde{\boldsymbol{c}}$ with $\widetilde{\boldsymbol{c}} \in \mathscr{C}$, and therefore, by virtue of

$$A\boldsymbol{c} = \boldsymbol{r} = \boldsymbol{r}_0 - A\widetilde{\boldsymbol{c}} \quad \text{i.e.,} \quad \boldsymbol{r}_0 = A(\boldsymbol{c} + \widetilde{\boldsymbol{c}}),$$

we see that the augmented correction space $\widetilde{\mathscr{C}} = \mathscr{C} + \mathrm{span}\{\boldsymbol{c}\}$ contains the exact correction. In practice, since solving $A\boldsymbol{c} = \boldsymbol{r}$ is generally as difficult as the original problem, one applies an inexpensive approximate solution method to this auxiliary problem, yielding a

vector $\boldsymbol{c}$ satisfying $A\boldsymbol{c} = \boldsymbol{r} + \boldsymbol{h}$ for some $\boldsymbol{h} \neq \boldsymbol{0}$ and consequently, $\|\widetilde{\boldsymbol{r}}^{\mathrm{MR}}\| \leq \|\boldsymbol{h}\|$ for the MR residual with respect to $\widetilde{\mathscr{C}}$.

The FGMRES algorithm of Saad (1993), which we introduced in Section 3.2 as the specialization of the abstract MR algorithm of Section 2.3.2 to solving equations with general correction spaces, was originally introduced as a technique that enlarges the correction space at each step by an approximate solution of such a residual equation. In Saad (1993), this is achieved by selecting the new correction direction $\boldsymbol{c}_{m+1}$ as the result of a preconditioning step applied to the most recent basis vector $\boldsymbol{v}_{m+1}$ of the residual space $\mathscr{V}_{m+1}$, which may be viewed as an approximate solution of the equation $A\boldsymbol{c} = \boldsymbol{v}_{m+1}$. In this regard, any (right) preconditioned Krylov subspace method can be interpreted this way; the original motivation for FGMRES was, however, to allow for a *different* preconditioner in each step, as a result of which the correction space is no longer a Krylov space.

A similar approach is taken in the GMRESR (which stands for GMRES Recursive) method of van der Vorst & Vuik (1994). In each step of GMRESR, the new correction vector $\boldsymbol{c}_{m+1}$ is chosen as the approximate solution of the equation $A\boldsymbol{c} = \boldsymbol{r}_m$ obtained by a fixed number of GMRES steps, where $\boldsymbol{r}_m$ is the residual of the MR approximation using the current correction space $\mathscr{C}_m$. This method was improved upon by de Sturler (1996), who observed that, by enforcing orthogonality of the approximation space of the inner GMRES iteration, one can obtain as a result of the inner GMRES iteration the best approximation of $\boldsymbol{r}_0$ from the sum of the inner and outer approximation spaces as described in Section 3.3.

5.3 Restarted GMRES

The most widely used restarted MR algorithm is GMRES(m), the restarted version of GMRES (Saad & Schultz 1986) using a Krylov space of dimension at most m, which is summarized below as Algorithm 5.3.1. One cycle of GMRES(m) for solving (1.1) with initial residual $\boldsymbol{r}_0$ consists of generating the Krylov space $\mathscr{K}_m(A, \boldsymbol{r}_0)$, forming the MR approximation with respect to the correction space $\mathscr{C} = \mathscr{K}_m(A, \boldsymbol{r}_0)$ according to Algorithm 4.2.3, and then repeating this process using the resulting residual as the initial residual for the next cycle, until a stopping criterion is satisfied.

Algorithm 5.3.1 Restarted GMRES with restart length m (GMRES(m)).

```
1 x_0 given, r_0 := b − A x_0
2 while not converged do
3         perform m steps of GMRES, yielding x_m and r_m
4        x_0 := x_m, r_0 := r_m
5 end
```

In the terminology of Section 3.3, two consecutive cycles of GMRES(m) consist of two MR approximations with respect to the correction spaces

$$\mathscr{C}_1 = \mathscr{K}_m(A, \boldsymbol{r}_0) \qquad \text{and} \qquad \mathscr{C}_2 = \mathscr{K}_m(A, \boldsymbol{r}_m),$$

where $\boldsymbol{r}_m$ denotes the residual of the MR approximation computed in the first cycle. In the second cycle, no orthogonalization of the current residual space $\mathscr{V}_2$ against the approximation space $A\mathscr{C}_1$ of the first cycle is performed, and thus, in general, the approximation produced in the second cycle is no longer the MR approximation with respect to $\mathscr{C}_1 + \mathscr{C}_2$. Besides this unoptimal approximation, it may also happen that the sum of the two correction spaces is not direct. In the extreme case of stagnation there holds $\boldsymbol{r}_m = \boldsymbol{r}_0$ after the first cycle, so that the second cycle constructs the identical Krylov space (as do all subsequent cycles), and no progress is made. The following proposition states a criterion for two consecutive correction spaces to have a direct sum.

Proposition 5.3.1. *For two consecutive cycles of GMRES(m) with initial residual $\boldsymbol{r}_0$, there holds*

$$\mathscr{K}_m(A, \boldsymbol{r}_0) \oplus \mathscr{K}_m(A, \boldsymbol{r}_m) = \mathscr{K}_{2m}(A, \boldsymbol{r}_0) \tag{5.1}$$

if and only if no Galerkin breakdown occurs in the last step of the first cycle.

Proof. By definition,

$$\mathscr{K}_m(A, \boldsymbol{r}_m) = \{q(A)\, p_m^{\mathrm{MR}}(A)\boldsymbol{r}_0 : q \in \mathscr{P}_{m-1}\}$$

where p_m^{MR} denotes the MR polynomial of the last step of the first cycle, and this shows that (5.1) holds if and only if p_m^{MR} has degree m. The representation (4.20) of p_m^{MR} shows that this is equivalent with $v_m(0) \neq 0$, which, by the second of the equivalent characterizations given in Proposition 3.1.1, is the case if and only no Galerkin breakdown occurs in the last step of the first cycle. □

Proposition 5.3.1 shows that, unless a Galerkin breakdown occurs at the end of a restart cycle, restarted GMRES selects its corrections from the same Krylov space as the unrestarted scheme. This is in contrast to truncated Krylov subspace methods, which always generate the full Krylov space $\mathscr{K}_L(A, \boldsymbol{r}_0)$ unless the iteration is stopped before step L.

If stagnation or near-stagnation is encountered in a restart cycle, a new suitable initial residual is necessary to generate more useful search directions. Techniques by which this may be accomplished include the "LSQR switch" of van der Vorst & Vuik (1994) or a method based on a total least squares solution of the projected problem proposed by Simoncini (1999).

Beyond some qualitative statements to be given in Section 6, there are still few quantitative results which describe the convergence behavior of GMRES(m). One of the more common misconceptions regarding GMRES(m) is that a method with larger restart length m_1 applied to the same problem will converge at least as fast as the method with smaller restart length $m_2 < m_1$. A simple counterexample[1] is provided by the 3×3 system

$$A\boldsymbol{x} = \boldsymbol{b}, \quad A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

[1]The author would like thank E. de Sturler for pointing out this phenomenon, reporting a similar observation in the context of a discrete convection-diffusion problem.

with initial guess $\boldsymbol{x}_0 = \mathbf{0}$. Two cycles of GMRES(2) applied to this example result in a residual norm of $\|\boldsymbol{r}_4^{(2)}\| = 4/15 = 0.26\ldots$, whereas four cycles of GMRES(1), which involve the same number of matrix-vector multiplications, yields $\|\boldsymbol{r}_4^{(1)}\| = 0.057\ldots$. The gap between GMRES(1) and GMRES(2) widens further in subsequent iteration steps, e.g., $\|\boldsymbol{r}_{18}^{(1)}\|_2 = 1.6\ldots 10^{-12}$, whereas $\|\boldsymbol{r}_{18}^{(2)}\|_2 = 3.9\ldots 10^{-5}$. Even more surprising in this example is that $\|\boldsymbol{r}_{10}^{(1)}\|_2 < \|\boldsymbol{r}_{20}^{(2)}\|_2$, showing that ten cycles of GMRES(1) have reduced the residual further than ten cycles of GMRES(2). By expanding this example to the analogous matrix for higher dimensions n one can observe that GMRES(m) is ultimately slower for this system than GMRES($m-1$) for $m = 2, \ldots, n-1$.

5.4 Acceleration Techniques

Since restarting usually results in slower convergence (or the loss thereof) much recent work has been devoted to compensating for the loss of information that occurs upon restarting by retaining a judiciously chosen part of previously generated spaces.

We distinguish two fundamental strategies in existing work: The first lies in identifying a subspace $\mathscr{U}$ which slows convergence, approximating this space, and eliminating its influence from the iteration process. We shall refer to such a procedure as *deflation*. Such "problematic" subspaces are often identified as eigenspaces of A associated with eigenvalues of small magnitude, but other spaces may sometimes be better suited. Examples of this approach are the augmentation method introduced in Morgan (1995) and Morgan (1997) and analyzed by Saad (Saad 1997*a*, Saad 1997*b*) and Chapman & Saad (1997). Another device for eliminating $\mathscr{U}$ from the iteration is to introduce a preconditioner which inverts the orthogonal section of A onto $\mathscr{U}$, as proposed by Erhel, Burrage & Pohl (1996), Baglama, Calvetti, Golub & Reichel (1998) and, with certain modifications, by Kharchenko & Yeremin (1995).

Rather than eliminating a certain subspace, a second fundamental strategy consists in maintaining orthogonality to subspaces generated in earlier cycles. The essential orthogonality constraints may be determined by comparing angles between subspaces, and the cost of the algorithm is controlled by maintaining orthogonality only against the most important subspace of a given dimension. Such a strategy, which is closer in character to a truncated MR method, is proposed by de Sturler (1999) and is discussed in Section 5.4.4.

5.4.1 Deflation by Augmentation

The first algorithm which attempts to improve the restarted GMRES method by augmenting the Krylov space is given in Morgan (1995). This approach selects a fixed number of approximate eigenvectors of A to augment the Krylov space of the following cycle, which can be motivated e.g. by the result of Lemma 5.2.3. Since the emphasis of this work is on cases in which the eigenvalues close to the origin limit the convergence rate the most—as is the case e.g. for the so-called model problem of the discrete Laplacian on the unit cube—harmonic Ritz vectors are chosen as the eigenvector approximations, since, as shown by (Morgan 1991), harmonic Ritz values tend to approximate eigenvalues close to zero more accurately than classical Ritz values.

Each step except the first consists of forming the MR approximation with respect to a correction space $\mathscr{C} = \mathscr{C}_1 + \mathscr{C}_2$ with $\mathscr{C}_1 = \mathscr{K}_m(A, \boldsymbol{r}_0)$ and $\mathscr{C}_2 = \text{span}\{\widetilde{\boldsymbol{z}}_1, \ldots, \widetilde{\boldsymbol{z}}_k\}$. The vectors $\widetilde{\boldsymbol{z}}_1, \ldots, \widetilde{\boldsymbol{z}}_k$ are the harmonic Ritz vectors associated with the k harmonic Ritz values $\widetilde{\theta}_1, \ldots, \widetilde{\theta}_k$ of A with respect to the previous correction space which are closest to the origin. Since no eigenvector information is available in the first cycle, the first correction space is chosen as simply $\mathscr{C} = \mathscr{K}_{m+k}(A, \boldsymbol{r}_0)$.

If k harmonic Ritz vectors are used to augment the Krylov space, one can save k matrix-vector multiplications in the implementation of this approach. The key observation is that the augmented space is itself a Krylov space. Morgan (1997) gave a somewhat lengthy and cumbersome proof of this result, so we include a more concise proof below.

Consider the MR approximation with initial residual $\boldsymbol{r}_0$ with respect to the $(m+k)$-dimensional Krylov space $\mathscr{K}_{m+k}(A, \boldsymbol{r}_0)$. As shown in Section 4.3.2, the associated residual vector has the representation

$$\boldsymbol{r}_{m+k}^{\text{MR}} = p_{m+k}^{\text{MR}}(A)\boldsymbol{r}_0$$

where

$$p_{m+k}^{\text{MR}}(\zeta) = \prod_{j=1}^{m+k} \left(1 - \frac{\zeta}{\widetilde{\theta}_j}\right).$$

Denote by q_m the polynomial whose zeros are the harmonic Ritz values $\widetilde{\theta}_{k+1}, \ldots, \widetilde{\theta}_{k+m}$, i.e., those largest in modulus.

Theorem 5.4.1. *The correction space $\mathscr{C}$ of Morgan's method is itself a Krylov space, namely*

$$\mathscr{C} = \mathscr{K}_m(A, \boldsymbol{r}_{m+k}) + \mathit{span}\{\widetilde{\boldsymbol{z}}_1, \ldots, \widetilde{\boldsymbol{z}}_k\} = \mathscr{K}_{m+k}(A, q_m(A)\boldsymbol{r}_0). \tag{5.2}$$

Proof. The rightmost member of (5.2) can be represented as

$$\mathscr{K}_{m+k}(A, q_m(A)\boldsymbol{r}_0) = \{r(A)q_m(A)\boldsymbol{r}_0 : r \in \mathscr{P}_{m+k-1}\}.$$

On the other hand, by (4.23), the harmonic Ritz vectors possess the polynomial representation $\widetilde{\boldsymbol{z}}_j = \widetilde{z}_j(A)\boldsymbol{r}_0$ with

$$\widetilde{z}_j(\zeta) = \frac{p_{m+k}^{\text{MR}}(\zeta)}{\zeta - \widetilde{\theta}_j} = q_m(\zeta) \prod_{\substack{\ell=1 \\ \ell \neq j}}^{k} \left(1 - \frac{\zeta}{\widetilde{\theta}_\ell}\right),$$

whereas $\boldsymbol{r}_{m+k}^{\text{MR}} = p_{m+k}^{\text{MR}}(A)\boldsymbol{r}_0$, with

$$p_{m+k}^{\text{MR}}(\zeta) = q_m(\zeta) \prod_{\ell=1}^{k} \left(1 - \frac{\zeta}{\widetilde{\theta}_\ell}\right).$$

The polynomial representation of vectors in both components of the correction space $\mathscr{C}$ of Morgan's method thus possess the common factor q_m, and therefore $\mathscr{C}$ may be characterized as

$$\mathscr{C} = \{q_m(A)q(A)\boldsymbol{r}_0 : q \in \mathscr{Q}\}$$

where the polynomial space $\mathscr{Q}$ is given by

$$\begin{aligned}\mathscr{Q} :=& \prod_{\ell=1}^{k}\left(1-\frac{\zeta}{\widetilde{\theta}_\ell}\right)\,\mathscr{P}_{m-1} + \operatorname{span}\left\{\prod_{\substack{\ell=1\\ \ell\neq j}}^{k}\left(1-\frac{\zeta}{\widetilde{\theta}_\ell}\right) : j = 1\ldots,k\right\}\\ =& \prod_{\ell=1}^{k}\left(1-\frac{\zeta}{\widetilde{\theta}_\ell}\right)\,\mathscr{P}_{m-1} + \mathscr{P}_{k-1} = \mathscr{P}_{m+k-1},\end{aligned}$$

where the second equality follows from the fact that $\widetilde{\theta}_1,\ldots,\widetilde{\theta}_k$ are distinct. □

Equation (5.2) shows that $\mathscr{C}$ can be generated by first applying the IRA method to $\mathscr{K}_{m+k}(A,\boldsymbol{r}_0)$, using $\widetilde{\theta}_{k+1},\ldots,\widetilde{\theta}_{m+k}$ as shifts, to obtain $\mathscr{K}_k(A,q_m(A)\boldsymbol{r}_0)$. The space $\mathscr{C}$ is then obtained after m further steps of the Arnoldi process. This approach is computationally less expensive in that k fewer matrix-vector multiplications with A are required.

As also noted by Morgan (1997), an analogous method can be used to augment the Krylov space in conjunction with an OR iteration. In this case, however, Ritz values and vectors must be used in place of harmonic Ritz values/vectors, as the Ritz values are the zeros of the OR residual polynomial.

The above proof also shows that

$$\operatorname{span}\{\widetilde{\boldsymbol{z}}_1,\ldots,\widetilde{\boldsymbol{z}}_k,\boldsymbol{r}_{m+k}^{\mathrm{MR}}\} = \mathscr{K}_{k+1}(A,q_m(A)\boldsymbol{r}_0).$$

Rather than use the IRA method, one could also generate the augmented Krylov space by first orthonormalizing the harmonic Ritz vectors, then orthonormalizing $\boldsymbol{r}_{m+k}^{\mathrm{MR}}$ against these, and then generate the remainder of the space by the Arnoldi process. Since these vectors all lie in the previous space $\mathscr{K}_{m+k}(A,\boldsymbol{r}_0)$, of which an orthonormal basis is available from the previous cycle, these orthonormalizations may be carried out in the coordinate space. This modification is proposed by Morgan (1999), who ascribes improved stability properties to this variant compared with using the IRA method.

5.4.2 Deflation by Preconditioning

The next class of methods also attempt to utilize spectral information gained in prior restart cycles to accelerate convergence. Instead of augmenting the Krylov space, the same information is used here to construct a sequence of preconditioners which can be improved as more accurate spectral information becomes available. The first such approach was introduced by Erhel et al. (1996).

To motivate this approach, assume U is an orthonormal basis of an A-invariant subspace $\mathscr{U}$ of dimension k, i.e.,

$$AU = UA_U, \quad A_U \in \mathbb{C}^{k\times k}.$$

Note that A_U is the specific representation of the orthogonal section $A_{\mathscr{U}}$ with respect to the basis U. Denoting by $U_\perp$ an orthonormal basis of the orthogonal complement $\mathscr{U}^\perp$, we can represent the action of A as

$$A\begin{bmatrix}U & U_\perp\end{bmatrix} = \begin{bmatrix}U & U_\perp\end{bmatrix}\begin{bmatrix}A_U & U^*AU_\perp\\ O & U_\perp^*AU_\perp\end{bmatrix}.$$

Under the assumption that k is small, it is feasible solve systems involving A_U directly, and thus to precondition by M defined as

$$M \begin{bmatrix} U & U_\perp \end{bmatrix} = \begin{bmatrix} U & U_\perp \end{bmatrix} \begin{bmatrix} A_U & O \\ O & I_{n-k} \end{bmatrix} \tag{5.3}$$

at each step of the iteration. The resulting right-preconditioned operator is given by

$$AM^{-1} \begin{bmatrix} U & U_\perp \end{bmatrix} = \begin{bmatrix} U & U_\perp \end{bmatrix} \begin{bmatrix} I_k & U^*AU_\perp \\ O & U_\perp^*AU_\perp \end{bmatrix}, \quad \text{i.e.,} \quad AM^{-1} = P_{\mathscr{U}} + AP_{\mathscr{U}^\perp}. \tag{5.4}$$

We now compare this preconditioning scheme with Morgan's method of augmenting the Krylov space $\mathscr{K}_m(A, \boldsymbol{r}_0)$ by the A-invariant subspace $\mathscr{U}$.

Theorem 5.4.2. *Let $\boldsymbol{r}_m^M$ denote the MR residual with respect to the correction space $\mathscr{U} + \mathscr{K}_m(A, \boldsymbol{r}_0)$, where $\mathscr{U}$ is an A-invariant subspace, and let $\boldsymbol{r}_m^E$ denote the MR residual with respect to the correction space $\mathscr{K}_m(AM^{-1}, \boldsymbol{r}_0)$ resulting from preconditioning A from the right by M as defined in* (5.3)*. Then there holds*

$$0 = \|P_{\mathscr{U}} \boldsymbol{r}_m^M\| \leq \|P_{\mathscr{U}} \boldsymbol{r}_m^E\| \quad \text{and} \quad \|P_{\mathscr{U}^\perp} \boldsymbol{r}_m^M\| \leq \|P_{\mathscr{U}^\perp} \boldsymbol{r}_m^E\|, \tag{5.5}$$

and therefore $\|\boldsymbol{r}_m^M\| \leq \|\boldsymbol{r}_m^E\|$. If, in addition, also $\mathscr{U}^\perp$ is A-invariant, then, $P_{\mathscr{U}} \boldsymbol{r}_0 = \boldsymbol{0}$ implies $\boldsymbol{r}_m^E = \boldsymbol{r}_m^M$.

Proof. The left set of inequalities in (5.5) follow from $P_{\mathscr{U}} \boldsymbol{r}_m^M = \boldsymbol{0}$, which is a restatement of the fact that augmenting with an invariant subspace $\mathscr{U}$ eliminates $\mathscr{U}$ from the residual (Lemma 5.2.3).

We next recall that $A_{\mathscr{U}^\perp} = P_{\mathscr{U}^\perp} A P_{\mathscr{U}^\perp}$ is the orthogonal section of A onto $\mathscr{U}^\perp$ (cf. the remark following Lemma 4.3.3). Since $\boldsymbol{r}_m^E = \boldsymbol{r}_0 - AM^{-1}\boldsymbol{c}$, for some $\boldsymbol{c} \in \mathscr{K}_m(AM^{-1}, \boldsymbol{r}_0)$ we obtain using (5.4)

$$P_{\mathscr{U}^\perp} \boldsymbol{r}_m^E = P_{\mathscr{U}^\perp} \boldsymbol{r}_0 - P_{\mathscr{U}^\perp} AM^{-1} \boldsymbol{c} = P_{\mathscr{U}^\perp} \boldsymbol{r}_0 - P_{\mathscr{U}_\perp} A P_{\mathscr{U}^\perp} \boldsymbol{c} = P_{\mathscr{U}^\perp} \boldsymbol{r}_0 - A_{\mathscr{U}^\perp} P_{\mathscr{U}^\perp} \boldsymbol{c}.$$

Moreover, $AM^{-1}\mathscr{U} = \mathscr{U}$ together with Lemma 4.3.3 yield

$$P_{\mathscr{U}^\perp} \boldsymbol{c} \in P_{\mathscr{U}^\perp} \mathscr{K}_m(AM^{-1}, \boldsymbol{r}_0) = \mathscr{K}_m(P_{\mathscr{U}^\perp} AM^{-1}, P_{\mathscr{U}^\perp} \boldsymbol{r}_0) = \mathscr{K}_m(A_{\mathscr{U}^\perp}, P_{\mathscr{U}^\perp} \boldsymbol{r}_0).$$

The last two statements show that $P_{\mathscr{U}^\perp} \boldsymbol{r}_m^E$ is of the form $P_{\mathscr{U}^\perp} \boldsymbol{r}_0 - A_{\mathscr{U}^\perp} \widetilde{\boldsymbol{c}}$ with $\widetilde{\boldsymbol{c}} \in \mathscr{K}_m(A_{\mathscr{U}^\perp}, P_{\mathscr{U}^\perp} \boldsymbol{r}_0)$. On the other hand, by Proposition 5.2.3, there holds

$$\|\boldsymbol{r}_m^M\| = \min_{\boldsymbol{c} \in \mathscr{K}_m(A_{\mathscr{U}^\perp}, P_{\mathscr{U}^\perp} \boldsymbol{r}_0)} \|P_{\mathscr{U}^\perp} \boldsymbol{r}_0 - A_{\mathscr{U}^\perp} \boldsymbol{c}\|,$$

i.e., $\|\boldsymbol{r}_m^M\|$ minimizes all expressions of this form, yielding the right inequality of (5.5).

Next, assuming $A\mathscr{U}^\perp = \mathscr{U}^\perp$, (5.4) implies $AM^{-1}\boldsymbol{r}_0 = A_{\mathscr{U}^\perp} \boldsymbol{r}_0$ for $\boldsymbol{r}_0 \in \mathscr{U}^\perp$, and thus $\mathscr{K}_m(AM^{-1}, \boldsymbol{r}_0) = \mathscr{K}_m(A_{\mathscr{U}^\perp}, P_{\mathscr{U}^\perp} \boldsymbol{r}_0)$, which shows that in this case both methods minimize over the same subspace, hence $\boldsymbol{r}_m^E = \boldsymbol{r}_m^M$. ∎

We note that the assumption $P_{\mathscr{U}}\boldsymbol{r}_0 = \mathbf{0}$ is not restrictive, as this can be enforced by adding the correction $UA_U^{-1}U^*\boldsymbol{r}_0$ to $\boldsymbol{x}_0$ and the preconditioner is built upon the premise that A_U is easily invertible. However, since $P_{\mathscr{U}}\boldsymbol{r}_0 = \mathbf{0}$ by no means implies that $P_{\mathscr{U}}\boldsymbol{r}_m^E = \mathbf{0}$ for $m > 0$, it cannot be guaranteed that $\|\boldsymbol{r}_m^E\| = \|\boldsymbol{r}_m^M\|$ even for such a special choice of initial residual unless $A\mathscr{U}^\perp = \mathscr{U}^\perp$. In the finite-dimensional case, the condition that $\mathscr{U}^\perp$ be invariant whenever $\mathscr{U}$ is invariant—i.e., that all invariant spaces also *reduce* A—is a characterization for A to be normal. Hence, these two approaches are equivalent when A is normal and $\mathscr{U}$ is invariant.

The availability of an (exactly) A-invariant subspace $\mathscr{U}$, on the other hand, is an assumption that can rarely be satisfied in practice. For a non-invariant $\mathscr{U}$ one can nonetheless still define the preconditioner as in (5.3), where now A_U is defined as $A_U := U^*AU$, resulting in

$$AM^{-1}\begin{bmatrix}U & U_\perp\end{bmatrix} = \begin{bmatrix}U & U_\perp\end{bmatrix}\begin{bmatrix}I & U^*AU_\perp \\ U_\perp^*AUA_U^{-1} & U_\perp^*AU_\perp\end{bmatrix},$$

based on the heuristic that $U_\perp^*AUA_U^{-1}$ will be small whenever $\mathscr{U}$ is *nearly* A-invariant. In Erhel et al. (1996) such nearly A-invariant spaces are obtained as the span of selected Ritz or harmonic Ritz vectors determined from Krylov spaces generated during previous cycles.

Baglama et al. (1998) propose a similar algorithm, which preconditions by (5.3) from the left, leading to the preconditioned operator

$$M^{-1}A\begin{bmatrix}U & U_\perp\end{bmatrix} = \begin{bmatrix}U & U_\perp\end{bmatrix}\begin{bmatrix}I & A_U^{-1}U^*AU_\perp \\ O & U_\perp^*AU\ \perp\end{bmatrix},$$
$$\text{or } M^{-1}A = P_{\mathscr{U}} + AP_{\mathscr{U}_\perp} + (A^{-1} - I)P_{\mathscr{U}}AP_{\mathscr{U}_\perp},$$

where we have again assumed that we are in the idealizd case of anwhere $\mathscr{U}$ is exactly A-invariant. The MR correction of the left-preconditioned system is the solution of the minimization problem

$$\|M^{-1}\boldsymbol{r}_m^B\| = \min\{\|M^{-1}(\boldsymbol{r}_0 - AM^{-1}\boldsymbol{c})\| \,:\, \boldsymbol{c} \in \mathscr{K}_m(AM^{-1}, \boldsymbol{r}_0)\}$$

(cf. Section 4.6).

From (5.3), it is evident that,

$$M^{-1} = A^{-1}P_{\mathscr{U}} + P_{\mathscr{U}}^\perp$$

and, consequently, if $A\mathscr{U} = \mathscr{U}$,

$$P_{\mathscr{U}_\perp}M^{-1}\boldsymbol{v} = P_{\mathscr{U}^\perp}\boldsymbol{v}, \quad \text{for all } \boldsymbol{v}.$$

These are the essential ingredients for showing that Proposition 5.4.2 holds in exactly the same way with $\boldsymbol{r}_m^E$ in place of $\boldsymbol{r}_m^B$. The construction of an approximately invariant subspace $\mathscr{U}$ is accomplished by Baglama et al. (1998) by employing the IRA process (cf. Section 4.3.3).

Kharchenko & Yeremin (1995) suggest another adaptive right preconditioner $\widetilde{M}$: After each GMRES cycle the Ritz values and the corresponding left[2] and right Ritz vectors of

[2]Left Ritz vectors are defined by $A^*\tilde{\boldsymbol{z}}_j - \bar{\theta}_j\tilde{\boldsymbol{z}}_j \perp \mathscr{K}_m$ and can be obtained from the left eigenvectors of H_m.

A with respect $\mathscr{K}_m$ are extracted. The aim is to obtain a preconditioner such that the extremal eigenvalues of A, which are approximated by the Ritz values, are translated to one (or at least to a small cluster around one) in the transition from A to $A\widetilde{M}^{-1}$.

The extremal Ritz values are partitioned into, say, k subsets Θ_j of nearby Ritz values. For each Θ_j, a rank-one transformation of the form $I + \boldsymbol{v}_j\tilde{\boldsymbol{v}}_j^*$ is constructed, where $\boldsymbol{v}_j$ and $\tilde{\boldsymbol{v}}_j$ are linear combinations of the associated right and left Ritz vectors. These linear combinations are chosen to translate simultaneously all Ritz values of Θ_j into a small cluster around one, while satisfying certain stability criteria. One preconditioning step now consists of successive multiplication by these rank-one matrices, i.e.,

$$\widetilde{M}^{-1} = (I + \boldsymbol{v}_1\widetilde{\boldsymbol{v}}_1^*)\cdots(I + \boldsymbol{v}_k\widetilde{\boldsymbol{v}}_k^*) = I + V_k\widetilde{V}_k^*, \qquad V_k = \left[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\right], \quad \widetilde{V}_k = \left[\tilde{\boldsymbol{v}}_1 \ldots \tilde{\boldsymbol{v}}_k\right].$$

For the last equation we have made use of the fact that $\tilde{\boldsymbol{v}}_j^*\boldsymbol{v}_i = 0$ for $i \neq j$, since all eigenvalues of H_m have geometric multiplicity one. Note that, if Θ_j has a small diameter and the Ritz values contained in Θ_j are good approximations of eigenvalues of A, then $\boldsymbol{v}_j$ and $\widetilde{\boldsymbol{v}}_j$ are approximate right and left eigenvectors of A. Moreover, the implementation described in Kharchenko & Yeremin (1995) ensures that the diagonal matrix $D := \widetilde{V}_k^*V_k \in \mathbb{C}^{k\times k}$ is nonsingular.

To compare this approach with the preconditioners presented thus far, we choose biorthonormal bases U and $\widetilde{U}$ of $\mathscr{U} := \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ and $\widetilde{\mathscr{U}} := \operatorname{span}\{\widetilde{\boldsymbol{v}}_1, \ldots, \widetilde{\boldsymbol{v}}_k\}$ such that $U^*U = I$, which are given e.g. by

$$U = V_k\widetilde{S}^{-1} \quad \text{and} \quad \widetilde{U} := \widetilde{V}_kD^{-H}\widetilde{S}^H,$$

with $\widetilde{S}$ any matrix that satisfies $V_k^*V_k = \widetilde{S}^H\widetilde{S}$. In this notation the preconditioner $\widetilde{M}$ is given by

$$\widetilde{M}^{-1} = I + U\widetilde{S}D\widetilde{S}^{-1}\widetilde{U}^* = I + US\tilde{U}^*, \qquad S := \widetilde{S}D\widetilde{S}^{-1}.$$

We let $U_\perp$ denote an orthonormal basis of $\mathscr{U}^\perp$ and make the idealizing assumptions that both $\mathscr{U}$ and $\widetilde{\mathscr{U}}$ are invariant with respect to A and A^*, respectively, i.e.,

$$AU = UA_U \quad \text{and} \quad \widetilde{U}^*A = A_U\widetilde{U}^*,$$

and that the eigenvalues corresponding to $\mathscr{U}$ (respectively $\widetilde{\mathscr{U}}$) are translated exactly to 1. Substituting this in the definition of the preconditioner, we obtain using the biorthonormality of U and $\widetilde{U}$,

$$\begin{aligned}
\widetilde{M}^{-1}\begin{bmatrix}U & U_\perp\end{bmatrix} &= \begin{bmatrix}U & U_\perp\end{bmatrix}\begin{bmatrix}I+S & S\widetilde{U}^*U_\perp \\ O & I_{n-k}\end{bmatrix} \quad \text{and} \\
A\widetilde{M}^{-1}\begin{bmatrix}U & U_\perp\end{bmatrix} &= \begin{bmatrix}U & U_\perp\end{bmatrix}\begin{bmatrix}A_U(I+S) & A_US\widetilde{U}^*U_\perp + U^*AU_\perp \\ O & U_\perp^*AU_\perp\end{bmatrix}.
\end{aligned}$$

In addition, our assumptions imply $A_U(I + S) = I$, i.e., $S = A_U^{-1} - I$ and $A_US\widetilde{U}^* = (I - A_U)\widetilde{U}^* = \widetilde{U}^*(I - A)$, resulting in

$$A\widetilde{M}^{-1}\begin{bmatrix}U & U_\perp\end{bmatrix} = \begin{bmatrix}U & U_\perp\end{bmatrix}\begin{bmatrix}I & \widetilde{U}^*(I-A)U_\perp + U^*AU_\perp \\ O & U_\perp^*AU_\perp\end{bmatrix}.$$

This leads to $A\widetilde{M}^{-1} = P_{\mathscr{U}} + P_{\mathscr{U}}^{\widetilde{\mathscr{U}}^\perp}(I-A)P_{\mathscr{U}^\perp} + AP_{\mathscr{U}^\perp}$ as the analogue to (5.4), where $P_{\mathscr{U}}^{\widetilde{\mathscr{U}}^\perp}$ denotes the oblique projection onto $\mathscr{U}$ along $\widetilde{\mathscr{U}}$. Thus, in view of $P_{\mathscr{U}_\perp}A\widetilde{M}^{-1} = A_{\mathscr{U}^\perp}$, the statement made in Theorem 5.4.2 holds also for this preconditioning approach.

5.4.3 Inner Iteration

As mentioned at the end of Section 5.2, one strategy for generating a new correction vector to extend a given correction space $\mathscr{C}$ is to use approximate solutions of the residual equation $A\boldsymbol{c} = \boldsymbol{r}$, where $\boldsymbol{r}$ is the minimum residual with respect to the current correction space $\mathscr{C}$. This strategy is followed in the GMRESR algorithms of van der Vorst & Vuik (1994), which consists of the GCR algorithm for general correction spaces (Algorithm 3.2.2) in which the new correction vector is obtained by applying an approximate solution method to the residual equation (see Algorithm 5.4.1 below). Here $\boldsymbol{c}_{m,k}^{\mathrm{MR}}$ denotes

Algorithm 5.4.1 GMRESR.

```
1  r_0 := b − Ax_0
2  for k = 1, 2, ...
3      perform m steps of GMRES applied to Ac = r_{k−1}
4      ĉ := c_{m,k}^MR, ŵ := Ac_{m,k}^MR
5      for i := 1 to k − 1
6          ĉ := ĉ − (ŵ, w_i)c_i
7          ŵ := ŵ − (ŵ, w_i)w_i
8      end
9      w_k := ŵ/‖ŵ‖, c_k := ĉ/‖ŵ‖
10     y_k^MR = W_k^* r_0
11     x_k^MR := x_0 + C_k y_k^MR = x_{k−1}^MR + (r_{k−1}, w_k)c_k
12     r_k^MR := r_0 − W_k y_k^MR = r_{k−1}^MR − (r_{k−1}, w_k)w_k
13 end
```

the approximate solution of the residual equation $A\boldsymbol{c} = \boldsymbol{r}_{k-1}$ obtained by the m-step inner GMRES iteration in the k-th outer iteration with initial approximation zero. Note that $\widehat{\boldsymbol{w}}$ does not require an extra matrix-vector multiplication as this quantity is constructed in the inner GMRES iteration.

In case of stagnation in the inner GMRES iteration the resulting correction vector $\widehat{\boldsymbol{c}}$ is zero, hence an alternative choice is necessary: van der Vorst & Vuik (1994) suggest using $\widehat{\boldsymbol{c}} := A^*\boldsymbol{r}_{k-1}$, which may be interpreted as the correction direction obtained from one step of a Krylov subspace method applied to the normal equations $A^*A\boldsymbol{c} = A^*\boldsymbol{r}_{k-1}$, resulting in the sequence of Krylov spaces $\mathscr{K}_m(A^*A, A^*\boldsymbol{r}_{k-1})$. One such method is due to Paige & Saunders (1982) and known as LSQR, and for this reason this modification was dubbed the *LSQR switch*.

We note that the same scheme for selecting new correction vectors can be used with FGMRES (Algorithm 3.2.3), although when FGMRES was first proposed (Saad 1993) the residual equation to be approximately solved was $A\boldsymbol{c} = \boldsymbol{c}_{k-1}$, where $\boldsymbol{c}_{k-1}$ is the most recent basis vector of the correction space. In Vuik (1993) FGMRES and GMRESR is compared and an example is given for which FGMRES breaks down before obtaining the exact solution due to a new correction vector which lies in the current correction space. We note that an analogous breakdown is possible for GMRESR.

An improvement of the GMRESR idea was given by de Sturler (1996), who proposed keeping the approximation space of the inner GMRES iteration orthogonal to that of the

outer iteration. The resulting algorithm is called GCRO to reflect the additional inner orthogonalization, and is presented below as Algorithm 5.4.2.

Algorithm 5.4.2 GCRO.

```
1  r_0 := b − Ax_0, W_0 := O
2  for k = 1, 2, ...
3      perform m steps of GMRES applied to (I − W_{k−1}W_{k−1}^*)Ac = r_{k−1}
4      ĉ := c_{m,k}^MR, ŵ = (I − W_{k−1}W_{k−1}^*)Aĉ
5      for i := 1 to k − 1
6          ĉ := ĉ − (ŵ, w_i)c_i
7          ŵ := ŵ − (ŵ, w_i)w_i
8      end
9      x_k := x_{k−1} + c_{m,k}^MR − C_k W_k^* A c_{m,k}^MR
10     r_k := r_{m,k}^MR
11 end
```

As in GMRESR, the quantities $\widehat{c}$ and $\widehat{w}$ are available at the end of the inner iteration as the correction and residual approximation. The inner GMRES iteration of the GCRO algorithm is an implementation of the abstract scheme of MR approximation on the sum $\mathscr{C} = \mathscr{C}_1 \oplus \mathscr{C}_2$ of two correction spaces presented in Section 3.3. In this case, after k outer steps, $\mathscr{C}_1$ is the span of the outer correction vectors $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k$, while $\mathscr{C}_2$ is the inner Krylov space $\mathscr{K}_m((I - W_k W_k^*)A), \boldsymbol{r}_k)$. Since the inner iteration computes the global MR approximation with respect to $\mathscr{C}_1 + \mathscr{C}_2$, the residual of the outer iteration coincides with that of the inner iteration. The quantity on the right in the expression for the outer iterate $\boldsymbol{x}_k$ looks more expensive than it is: the inner correction $\boldsymbol{c}_{m,k}^{\mathrm{MR}}$ has the form $\boldsymbol{c}_{m,k}^{\mathrm{MR}} = V_m \boldsymbol{y}, \boldsymbol{y} \in \mathbb{C}^m$ in terms of the Arnoldi basis V_m of the inner GMRES iteration. Thus the term in question has the form $C_k(W_k^* A V_m)\boldsymbol{y}$, of which the matrix in parentheses is computed as part of the inner orthogonalization process and is therefore available at no additional cost. (See also the discussion at the end of Section 3.3.)

As in GMRESR, other iteration methods besides GMRES are possible as inner iterations, and de Sturler also considers BICGSTAB. The interpretation of the inner approximation as an MR approximation over the inner and outer approximation spaces must then, however, be modified. Moreover, it is also possible to use FGMRES for the outer iteration, as follows immediately from the discussion in Section 3.3.

The discussion in Section 3.3 was based on the assumption that the sum $\mathscr{C}_1 + \mathscr{C}_2$ is *direct*. It is, however, conceivable that the inner Arnoldi process generates an inner correction vector $\boldsymbol{v}_j$ which happens to lie in the outer correction space $\mathscr{C}_1$. In this case $(I - W_k W_k^*)A\boldsymbol{v}_j = \boldsymbol{0}$ and the Arnoldi process breaks down, in general without finding the exact solution. That this is a rare occurrence is shown by de Sturler (1996): one can show that all correction vector (inner and outer) lie in the Krylov space $\mathscr{K}_L(A, \boldsymbol{r}_0)$ with termination index L. Therefore such a breakdown is (assuming exact arithmetic) not possible until the total number of inner iterations km exceeds L. De Sturler also suggests possibilities for continuing the inner and outer iterations after such a breakdown.

5.4.4 Optimal Truncation

The acceleration techniques of Sections 5.4.1 and 5.4.2 are based on restarting an MR iteration once the correction space has reached a given dimension m, and they attempt to compensate for the attendant loss of information by augmenting or preconditioning. The methods discussed in this section are related to the former in that they also attempt to retain information contained in the current correction space—in this case orthogonality constraints—which is deemed most useful for convergence.

In place of restarting, the basic scheme underlying this class of methods is a *truncated* MR iteration, in which, as soon as the correction space has reached a maximal dimension m, only a subset of the most recent m basis vectors of the correction space is retained, or equivalently, one or more of these basis vectors is periodically discarded during the iteration. Such a scheme for selectively discarding subspaces rather than individual basis vectors is proposed by de Sturler (1999) as a truncation scheme for his GCRO algorithm. This selection process, however, does not rely on spectral or invariant subspace information, but rather on angles between subspaces.

To discard a subspace of dimension ℓ, de Sturler's subspace selection scheme compares two approximation spaces $\mathscr{W}_1$ and $\mathscr{W}_2$ associated with correction spaces $\mathscr{C}_1$ and $\mathscr{C}_2$. It assumes the availability of an orthonormal basis $W_m^{(1)} = [\boldsymbol{w}_1^{(1)}, \ldots, \boldsymbol{w}_m^{(1)}]$ of $\mathscr{W}_1$, an *arbitrary* basis $\widehat{W}_k^{(2)} = [\widehat{\boldsymbol{w}}_1^{(2)}, \ldots, \widehat{\boldsymbol{w}}_k^{(2)}]$ of $\mathscr{W}_2$ as well as a factorization

$$(I_k - W_m^{(1)}[W_m^{(1)}]^*)\widehat{W}_k^{(2)} = Z_k R$$

with $Z_k = [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k]$, $Z_k^* Z_k = I_k$ and $R \in \mathbb{C}^{k\times k}$ nonsingular and upper triangular. After computing the singular value decomposition

$$\big([W_m^{(1)}]^*\widehat{W}_k^{(2)}\big)\big(Z_k^*\widehat{W}_k^{(2)}\big)^{-1} = X\Xi\widehat{Y}^H \tag{5.6}$$

the subspace of $\mathscr{W}_1$ to be retained is chosen as that spanned by the vectors $W_m^{(1)}\big[\boldsymbol{x}_1 \cdots \boldsymbol{x}_\ell\big]$, where the vectors $\boldsymbol{x}_j$ are the left singular vectors associated with the ℓ largest singular values. The following proposition relates this choice to the results of Section 3.4.

Theorem 5.4.3. *With the above notation under the assumption $\mathscr{W}_1 \cap \mathscr{W}_2 = \{\mathbf{0}\}$, the singular values appearing in* (5.6) *are the cotangents of the canonical angles between the spaces $\mathscr{W}_1$ and $\mathscr{W}_2$.*

Proof. Let $W_k^{(2)}$ denote an *orthonormal* basis of $\mathscr{W}_2$ such that $\widehat{W}_k^{(2)} = W_k^{(2)}S$ with a nonsingular matrix $S \in \mathbb{C}^{k\times k}$. Then the cosines of the canonical angles between $\mathscr{W}_1$ and $\mathscr{W}_2$ are the singular values of $[W_m^{(1)}]^*W_k^{(2)}$, and we write the associated singular value decomposition as $[W_m^{(1)}]^*W_k^{(2)} = X\Gamma Y^H$ with a diagonal matrix $\Gamma \in \mathbb{R}^{m\times k}$ and the unitary matrices $X \in \mathbb{C}^{m\times m}$ and $Y \in \mathbb{C}^{k\times k}$. From

$$\begin{aligned} Z_k R &= \Big(I_k - W_m^{(1)}[W_m^{(1)}]^*\Big)\widehat{W}_k^{(2)} = \Big(I_k - (W_m^{(1)}X)(W_m^{(1)}X)^*\Big)(W_k^{(2)}Y)Y^H S \\ &= \big[(W_k^{(2)}Y) - (W_m^{(1)}X)\Gamma\big]Y^H S \end{aligned}$$

we obtain $Z_k = \big[(W_k^{(2)}Y) - (W_m^{(1)}X)\Gamma\big]Y^H S R^{-1}$ and therefore, defining the diagonal matrix $\Sigma \in \mathbb{R}^{k\times k}$ by $I_k - \Gamma^H\Gamma = \Sigma^2$, there results

$$I_k = Z_k^* Z_k = (SR^{-1})^H Y \Sigma^2 Y^H (SR^{-1}) = (\Sigma Y^H S R^{-1})^H \Sigma Y^H S R^{-1}$$

which reveals that the $k \times k$ matrix $\Sigma Y^H S R^{-1}$ is also unitary. Note that, in view of $\mathscr{W}_1 \cap \mathscr{W}_2 = \{\mathbf{0}\}$, none of the cosines in Γ are one, hence Σ is nonsingular. Now, inserting

$$\begin{aligned}[W_m^{(1)}]^* \widehat{W}_k^{(2)} &= [W_m^{(1)}]^* W_k^{(2)} S = X \Gamma Y^H S \\ Z_k^* \widehat{W}_k^{(2)} &= (SR^{-1})^H Y \big[(W_k^{(2)} Y)^* - \Gamma^H (W_m^{(1)} X)^* \big] W_k^{(2)} S = (SR^{-1})^H Y \Sigma^2 Y^H S\end{aligned}$$

can express the singular value decomposition (5.6) as

$$\big([W_m^{(1)}]^* \widehat{W}_k^{(2)} \big) \big(Z_k^* \widehat{W}_k^{(2)} \big)^{-1} = X (\Gamma \Sigma^{-1}) (\Sigma Y^H S R^{-1}),$$

which reveals that its singular values are indeed the cotangents of the angles between $\mathscr{W}_1$ and $\mathscr{W}_2$. □

The proof also shows that the left singular vectors of (5.6) coincide with those of $[W_m^{(1)}]^* W_k^{(2)}$, hence the selection scheme discards that subspace of $\mathscr{W}_1$ which lies at the largest canonical angles with $\mathscr{W}_2$. As shown in Section 3.4, this choice yields the greatest possible residual reduction when replacing the approximation space $\mathscr{W}_1 + \mathscr{W}_2$ by $\widetilde{\mathscr{W}_1} + \mathscr{W}_2$ with $\widetilde{\mathscr{W}_1}$ a subspace of $\mathscr{W}_1$ of dimension $\dim \mathscr{W}_1 - k$.

De Sturler applies this scheme to a GMRES cycle of length m in order to determine which directions of the s-dimensional Krylov subspace $\mathscr{K}_s(A, \boldsymbol{r}_0)$, $s < m$, are most important for convergence in the sense that, were one to maintain orthogonality against these directions upon restarting after the first s steps, this would most accelerate the residual reduction. The subspaces to be compared are thus $A\mathscr{K}_s(A, \boldsymbol{r}_0)$ and $A\mathscr{K}_{m-s}(A, \boldsymbol{r}_s)$. The subspace comparison in this case is particularly inexpensive, as both spaces lie in $\mathscr{K}_m(A, \boldsymbol{r}_0)$, for which the Arnoldi process has computed an orthonormal basis. Hence, the angle computations can be performed in the coordinate space with respect to this basis, and therefore involve only small matrices.

This idea is applied to the inner GMRES iteration of GCRO in order to identify a subspace of the inner approximation space as important for convergence, and to select an orthonormal basis of that subspace—rather than just the residual update as in GCRO—to be added to the outer approximation space. Similarly, de Sturler uses this device in order to truncate the outer approximation space by comparing the angles between inner the and outer approximation spaces. For details of this truncated GCRO algorithm, known as GCROT, we refer to (de Sturler 1999).

Chapter 6

Convergence

In this chapter we turn to the question of convergence of Krylov subspace MR and OR methods for solving linear equations. In the finite-dimensional case convergence in a finite number of steps is assured due to the finite termination property of MR and OR methods. The interesting question here is how fast the errors or residuals decrease below a tolerance sufficiently small for a specific application.

We begin with an overview of the standard techniques for deriving residual and error bounds, which are based on eigenvalues, pseudoeigenvalues, or the field of values of A. We then proceed to show how the angles formulation of Chapter 2 allows a simple derivation of the most important bounds which have been derived in the literature, and give an example where a bound on the angles allows us to obtain superlinear convergence. We conclude with a discussion of matrices which generate the same MR and OR approximations and the role of singular values.

6.1 Convergence Bounds Based on Polynomials

In Section 4.3 we saw that the residual of any approximate solution of (1.1) taken from the shifted Krylov space $\boldsymbol{x}_0 + \mathscr{K}_m(A, \boldsymbol{r}_0)$ has the representation

$$\boldsymbol{r} = p(A)\boldsymbol{r}_0, \qquad p \in \mathscr{P}_m, \quad p(0) = 1.$$

This link between residuals and polynomials has inspired the search for bounds on the residual norm which are derived from analytic properties of the associated polynomials as functions defined on the complex plane. The influence of the initial residual is usually suppressed in view of $\|p(A)\boldsymbol{r}_0\| \leq \|p(A)\| \, \|\boldsymbol{r}_0\|$, in which case the issue simplifies to bounding $\|p(A)\|$ for all $p \in \mathscr{P}_m$ normalized by $p(0) = 1$. Since the *error* $\boldsymbol{e} := A^{-1}\boldsymbol{b} - \boldsymbol{x}$ possesses the identical representation $\boldsymbol{e} = p(A)\boldsymbol{e}_0$ in terms of the initial error $\boldsymbol{e}_0$, bounding $\|p(A)\|$ immediately also leads to bounds on the error reduction, although a polynomial which makes $\|p(A)\boldsymbol{r}_0\|$ small obviously need not do the same for $\|p(A)\boldsymbol{e}_0\|$.

In this section we describe the residual bounds which can be derived from this polynomial representation, restricting our considerations once more to MR approximations. Although some of the results which follow can be extended to the general case where A is a linear operator on a possibly infinite-dimensional Hilbert space $\mathscr{H}$, we will assume throughout this section that $\mathscr{H}$ is of finite dimension n, i.e., that A is a nonsingular matrix in $\mathbb{C}^{n\times n}$. This, of course, implies the existence of a finite termination index $L \leq n$.

Nearly all polynomial bounds involve reducing the minimization of $\|p(A)\|$ to a scalar problem. The most popular is to employ the Jordan canonical form of A: if the characteristic polynomial of A is given by $c_A(\zeta) = \prod_{j=1}^{k}(\zeta - \lambda_j)^{n_j}$ and T is any similarity which transforms A to its Jordan canonical form $J_A = T^{-1}AT$, then

$$\|p(A)\| \leq \kappa(T) \max_{1\leq j\leq k} n_j \max_{0\leq \ell\leq n_j-1} \frac{|p^{(\ell)}(\lambda_j)|}{\ell!}, \tag{6.1}$$

where $\kappa(T) := \|T\|\,\|T^{-1}\|$ is the condition number of T, $p^{(\ell)}$ denotes the ℓth derivative of p and the estimate $\|M\| \leq n \max_{i,j} |m_{i,j}|$ has been used. We always assume that the columns of T are chosen to have unit norm. Assuming A is diagonalizable, which we shall do in the sequel without (essential) loss of generality, (6.1) simplifies to

$$\|p(A)\| \leq \kappa(T)\,\|p\|_{\Lambda(A)}, \tag{6.2}$$

where $\Lambda(A)$ is the *spectrum* of A and we shall denote by $\|f\|_\Omega$ the maximum of a continuous function f on a compact set $\Omega \subset \mathbb{C}$.

Another approach makes use of the representation of $p(A)$ as a *contour integral* (see Kato (1980, Section I.5.6))

$$p(A) = \frac{1}{2\pi i}\int_\Gamma p(\zeta)(\zeta I - A)^{-1}\,d\zeta,$$

where Γ is a Jordan curve (or a collection of Jordan curves) which contains $\Lambda(A)$ in its interior. This approach has the advantage of readily generalizing to bounded operators on an infinite dimensional space. Taking norms yields

$$\|p(A)\| \leq \frac{\ell(\Gamma)}{2\pi} \max_{\zeta\in\Gamma} \|(\zeta I - A)^{-1}\|\;\|p\|_\Gamma, \tag{6.3}$$

where $\ell(\Gamma)$ denotes the arc length of Γ.

A third approach for bounding $\|p(A)\|$ based on the Schur decomposition of A may be found in Golub & van Loan (1996, Chapter 11).

A vast simplification occurs when A is a normal, and hence unitarily similar to a diagonal matrix: In this case $\kappa(T) = 1$ and (6.2) further simplifies to

$$\|p(A)\| = \|p\|_{\Lambda(A)}. \tag{6.4}$$

The quantity $\kappa(T)$ is actually a measure of how far A departs from normality—in particular, $\kappa(T) = 1$ if and only if A is normal. In the normal case both the contour integral and the Schur form bounds reduce to (6.4), whereas in the general case they contain measures of the non-normality of A other than $\kappa(T)$.

6.1.1 Bounds Based on Eigenvalue Inclusion Sets

For the MR residual vectors (6.2) together with the minimization property of the MR approximation result in

$$\frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|}{\|\boldsymbol{r}_0\|} \leq \kappa(T)\,\|p_m^{\mathrm{MR}}\|_{\Lambda(A)} \leq \kappa(T) \min_{\substack{p\in\mathscr{P}_m\\ p(0)=1}} \|p\|_{\Lambda(A)}, \qquad m = 1, 2, \ldots, L. \tag{6.5}$$

If A is normal, i.e., if $\kappa(T) = 1$, the right hand side of (6.5) is a standard problem of approximation theory: Among all polynomials p of degree at most m with $p(0) = 1$ determine p_m^* which deviates least from zero on a compact subset Ω of the complex plane (which in our case happens to be the discrete set $\Lambda(A)$), i.e.,

$$\|p_m^*\|_\Omega = \min_{\substack{p \in \mathscr{P}_m \\ p(0)=1}} \|p\|_\Omega. \tag{6.6}$$

This problem has a solution which is uniquely determined if $0 \notin \Omega$ and if the cardinality of Ω exceeds the polynomial degree m. Asymptotically, the numbers $\|p_m^*\|_\Omega$ behave as γ^m as $m \to \infty$, and the number

$$\gamma := \lim_{m \to \infty} \|p_m^*\|_\Omega^{1/m}$$

is called the *asymptotic convergence factor of* Ω. It can be shown that $\gamma < 1$ if $0 \notin \Omega$ and if Ω does not surround the origin. We observe that this asymptotic behavior implies that bounds of this type lead to *linear* convergence rate bounds.

If A is normal it can be shown (Greenbaum & Trefethen 1994, Greenbaum & Gurvits 1994, Joubert 1994*b*) that the bound (6.5) is sharp in the sense that for any m, $1 \le m \le L$, there exists an initial residual $\boldsymbol{r}_0$, which may depend on m, such that equality is attained. In this sense the convergence behavior of the MR method is completely determined by the eigenvalue distribution of A when A is normal.

If A is nonnormal and $\kappa(T)$ is not too large, one can expect that this statement, in essence, will still hold, although in this case (6.5) is no longer sharp, as counterexamples by Faber, Joubert, Knill & Manteuffel (1996) and Toh (1997) demonstrate. Greenbaum & Trefethen (1994) have shown that the (matrix) polynomial approximation problem of minimizing $\|p(A)\|$ subject to $p(0) = 1$ is also uniquely solvable in the general, nonnormal case. An algorithm for computing these optimal polynomials is given by Toh & Trefethen (1998).

An immediate consequence of (6.5) is that

$$\frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|}{\|\boldsymbol{r}_0\|} \le \kappa(T) \min_{\substack{p \in \mathscr{P}_m \\ p(0)=1}} \|p\|_\Omega, \qquad m = 1, 2, \ldots, L \tag{6.7}$$

holds for any compact set $\Omega \subset \mathbb{C}$ which contains the eigenvalues of A. This is the key to most of the known bounds for MR methods with respect to Krylov spaces: Select a compact set Ω, $\Lambda(A) \subseteq \Omega$, and a suitable polynomial $p \in \mathscr{P}_m$ with $p(0) = 1$, then $\kappa(T)\|p\|_\Omega$ (or $\|p\|_\Omega$ in the normal case) is an upper bound for $\|\boldsymbol{r}_m^{\mathrm{MR}}\|/\|\boldsymbol{r}_0\|$. Probably the best known example is the standard estimate for the conjugate gradient method, which is an OR method with respect to $(\cdot,\cdot)$ but an MR method with respect to the inner product $(\cdot,\cdot)_{A^{-1}} := (A^{-1}\cdot,\cdot)$ if A is Hermitian and positive definite. If $\|\cdot\|_A := (A\cdot,\cdot)^{1/2}$ and $\kappa := \lambda_{\max}(A)/\lambda_{\min}(A)$ is the spectral condition number of A, then

$$\frac{\|\boldsymbol{x}_m^{\mathrm{CG}} - A^{-1}\boldsymbol{b}\|_A}{\|\boldsymbol{x}_0 - A^{-1}\boldsymbol{b}\|_A} = \frac{\|\boldsymbol{r}_m^{\mathrm{CG}}\|_{A^{-1}}}{\|\boldsymbol{r}_0\|_{A^{-1}}} \le 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^m.$$

This follows directly from (6.7) if the set Ω is chosen to be the smallest closed interval $[\lambda_{\min}(A), \lambda_{\max}(A)]$ which contains $\Lambda(A)$ and p to be a suitably scaled and transformed

Chebyshev polynomial of degree m. Analogous bounds are available for more complicated spectral inclusion sets $\Omega \subset \mathbb{C}$, but this requires tools from conformal mapping and complex approximation theory in order to bound $\|p_m^*\|_\Omega$ of (6.6) (cf. Eiermann, Niethammer & Varga (1985), Driscoll, Toh & Trefethen (1998)). As a more recent example we quote from Liesen (1998) that, for $\Omega = \Omega(\phi) := \{\zeta = e^{i\theta} \,:\, \phi/2 \le \theta \le 2\pi - \phi/2\}$, there holds

$$\|p_m^*\|_\Omega \le \frac{4}{\gamma^m - 1}, \quad \text{where } \gamma := \frac{1}{\cos(\phi/4)}. \tag{6.8}$$

The determination of suitable spectral inclusion sets is a difficult task in itself. For Krylov subspace methods, the same Krylov space used as the correction space for the linear system may also be used as a Krylov projection method for finding approximate eigenvalues, from which inclusion sets may then be constructed. This is the approach used by so-called *hybrid methods*, in which spectral information is gathered in an initial phase of a Krylov subspace iteration. Once this information is deemed sufficient—after, say, m steps—a spectral inclusion set Ω is constructed and a residual polynomial p^* which solves (6.6) is used to generate further iterates which satisfy $\boldsymbol{r}_{m+j} = (p(A))^j \boldsymbol{r}_m$. For examples of hybrid methods see Nachtigal, Reichel & Trefethen (1992) and Starke & Varga (1993) as well as the references therein.

All bounds for $\|\boldsymbol{r}_m^{\mathrm{MR}}\|$ described above, i.e., (6.5) and (6.7), are based on spectral information available for A. There are, however, matrices which show that, in general, the spectrum has no influence on MR convergence behavior. Indeed, Greenbaum, Strakoš & Ptak (1996) show that for *any* nonincreasing finite sequence of positive real numbers $\rho_0 \ge \rho_1 \ge \cdots \ge \rho_{n-1}$ and *any* choice of (not necessarily distinct) nonzero complex numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$, one can construct a matrix $A \in \mathbb{C}^{n\times n}$ and an initial residual $\boldsymbol{r}_0$ with $\Lambda(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and $\|\boldsymbol{r}_m^{\mathrm{MR}}\| = \rho_m$ $(m = 0, 1, \ldots, n-1)$. We illustrate this result by one of their striking examples: Any matrix A in Frobenius form,

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \cdots & 0 & -\alpha_2 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\alpha_{n-1} \end{bmatrix} \in \mathbb{C}^{n\times n},$$

has $\zeta^n + \alpha_{n-1}\zeta^{n-1} + \cdots + \alpha_1\zeta + \alpha_0$ as its characteristic polynomial, so its eigenvalues can be arbitrarily prescribed. If we choose $\boldsymbol{b}$ and $\boldsymbol{r}_0$ such that $\boldsymbol{r}_0 = \boldsymbol{u}_1$ is the first unit vector, then, for $m = 1, 2, \ldots, n-1$, the approximation space $A\mathscr{K}_m(A, \boldsymbol{r}_0)$ is the span of the unit vectors $\boldsymbol{u}_2, \boldsymbol{u}_3, \ldots, \boldsymbol{u}_m$. The best approximation to $\boldsymbol{r}_0$ from this space is obviously the null vector leading to $\|\boldsymbol{r}_0\| = \|\boldsymbol{r}_1^{\mathrm{MR}}\| = \cdots = \|\boldsymbol{r}_{n-1}^{\mathrm{MR}}\| = 1$ independently of the chosen spectrum. In general it is therefore impossible to predict the convergence behavior of the MR method (and of any other Krylov subspace method) on the basis of the eigenvalue distribution of A alone. Although this fact has been emphasized in several recent papers it is still a widespread but nonetheless incorrect belief that spectral properties of the coefficient matrix (i.e., without any additional assumptions on its departure from normality) determine the speed of convergence of Krylov subspace methods. If A is far from normal, i.e., if $\kappa(T)$ is huge, the bound (6.5), although still valid, becomes useless because it vastly overestimates $\|\boldsymbol{r}_m^{\mathrm{MR}}\|/\|\boldsymbol{r}_0\|$, which we know to be always less than one.

The common interpretation of (6.5) in the sense that two matrices with identical spectra produce a similar MR behavior is incorrect, as the above example vividly illustrates. A systematic investigation of the class of matrices with identical MR behavior (given the same $\boldsymbol{r}_0$) is undertaken by Greenbaum & Strakoš (1994). An example of a problem arising in applications where the spectrum is misleading for the prediction of MR behavior may be found in Ernst (2000).

6.1.2 Bounds Based on Pseudospectra

Pseudospectra were popularized by Trefethen (Trefethen 1990, Trefethen 1992) in the early 1990s as a replacement for the concept of eigenvalues in situations where eigenvalues are sensitive to perturbations, i.e., when eigenvectors are far from orthogonal, which is the case for highly nonnormal matrices or operators. Given a value of the parameter $\epsilon \geq 0$, the set of *pseudoeigenvalues* $\Lambda_\epsilon(A)$ of A is defined by

$$\Lambda_\epsilon(A) = \{\lambda \in \mathbb{C} : \lambda \in \Lambda(A+E), \|E\| \leq \epsilon\},$$

or equivalently,

$$\Lambda_\epsilon(A) = \{\lambda \in \mathbb{C} : \|(\lambda I - A)^{-1}\| \geq \epsilon^{-1}\}.$$

The first definition shows that pseudoeigenvalues are eigenvalues of slight perturbations of A. The second definition characterizes pseudoeigenvalues as all complex numbers λ for which the norm of the resolvent $(\lambda I - A)^{-1}$—which for normal A is the reciprocal distance of λ to $\Lambda(A)$—has at least the magnitude $1/\epsilon$.

Bounds for the residual reduction of MR Krylov subspace methods are obtained by selecting the contour in (6.3) to be the boundary Γ_ϵ of $\Lambda_\epsilon(A)$, along which the resolvent norm has the constant value ϵ^{-1}, so that $\|p(A)\| \leq \frac{\ell(\Gamma_\epsilon)}{2\pi\epsilon} \|p\|_{\Gamma_\epsilon}$. For the MR residual at step m this results in the bound

$$\frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|}{\|\boldsymbol{r}_0\|} \leq \frac{\ell(\Gamma_\epsilon)}{2\pi\epsilon} \|p_m^{\mathrm{MR}}\|_{\Gamma_\epsilon} \leq \frac{\ell(\Gamma_\epsilon)}{2\pi\epsilon} \min_{\substack{p \in \mathscr{P}_m \\ p(0)=1}} \|p\|_{\Gamma_\epsilon}.$$

The information resulting from the pseudospectral bounds indicates that, for nonnormal matrices, in order for a residual polynomial to result in a large residual reduction, it should be small not only on $\Lambda(A)$ but also on the pseudospectral sets $\Lambda_\epsilon(A)$ for suitable values of ϵ. A strategy for choosing ϵ to result in the most useful pseudospectral bound is proposed by Embree (1999).

Some drawbacks of pseudospectral bounds are that pseudospectral information is generally harder to come by than even spectral information, although much effort has recently been devoted to the efficient computation of pseudospectra (see Trefethen (1999) and the references therein). Moreover, there are examples where the pseudospectral bounds result in a large overestimate of the residual reduction (Greenbaum & Strakoš 1994), such as when nonnormality is restricted to a low dimensional reducing subspace.

6.1.3 Bounds Based on the Field of Values

The field of values $W(A)$ of a matrix—also known as the *numerical range* in operator theory—defined in (4.9) has been used by Eiermann to derive bounds on the convergence of iterative methods (Eiermann 1993, Eiermann 1996). $W(A)$ is a convex set in the complex plane which contains $\Lambda(A)$, and for normal matrices it is the convex hull of $\Lambda(A)$. For nonnormal A, however, it may be much larger. A measure of the size of $W(A)$ is the *numerical radius* $\mu(A) := \max\{|\zeta| : \zeta \in W(A)\}$. The numerical radius is related to the norm associated with the underlying inner product by $\frac{1}{2}\|A\| \leq \mu(A) \leq \|A\|$. Further properties of the field of values may be found in the monograph of Horn & Johnson (1991, Chapter 1).

Eigenvalue-based bounds such as (6.2) rely on the fact that the eigenvalues of $p(A)$ are related to those of A via the spectral mapping theorem. An analogous mapping theorem for the field of values, however, fails to hold. Recently, Eiermann (1996) was able to exploit a mapping theorem due to Kato (1965) which holds for convex sets. This result applies to a special sequence of polynomials associated with a set $\Omega \subset \mathbb{C}$ whose complement with respect to the extended plane is simply connected, its *Faber polynomials* (Faber 1903, Faber 1920).

Theorem 6.1.1. *If $\{f_m\}_{m=0}^{\infty}$ denotes the sequence of Faber polynomials of the field of values $W(A)$ of a matrix $A \in \mathbb{C}^{n\times n}$, then the numerical radius $\mu(f_m(A))$ satisfies*

$$\mu(f_m(A)) \leq \|f_m\|_{W(A)}. \tag{6.9}$$

The crucial point of this result is that it relates the field of values of $f_m(A)$ to the size of the polynomial f_m on the set $W(A)$. To use Faber polynomials for estimating the convergence rate of Krylov subspace methods, we turn to the normalized Faber polynomials $\widehat{f}_m(z) := f_m(z)/f_m(0)$, which are admissible as residual polynomials. We must therefore require $f_m(0) \neq 0$, which is assured if we assume $0 \notin W(A)$. The asymptotic behavior of these polynomials as $m \to \infty$ is well-understood: the normalized Faber polynomials $\{\widehat{f}_m\}$ of a convex bounded set $0 \notin \Omega \subset \mathbb{C}$ satisfy $\|\widehat{f}_m\|_\Omega \leq c_m\gamma^m$ with $0 < \gamma < 1$ and $c_m < 2/(1-\gamma^m)$, where γ is the asymptotic convergence factor of Ω. The Faber polynomials thus approximately solve the polynomial approximation problem (6.6) in such a way that the norms of the scaled polynomials have the same m-th root asymptotics as the exact solution sequence $\{p_m^*\}$, which makes them interesting in this context since they provide a tight upper bound.

For the mth MR residual polynomial p_m^{MR} we have

$$\|\boldsymbol{r}_m^{\mathrm{MR}}\| = \|p_m^{\mathrm{MR}}(A)\boldsymbol{r}_0\| \leq \|\widehat{f}_m(A)\boldsymbol{r}_0\| \leq \|\widehat{f}_m(A)\|\|\boldsymbol{r}_0\|.$$

Theorem 6.1.1 together with the asymptotic properties of the normalized Faber polynomials thus yields the bound

$$\frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|_2}{\|\boldsymbol{r}_0\|_2} \leq \|\widehat{f}_m(A)\|_2 \leq 2\mu(\widehat{f}_m(A)) \leq 2\|\widehat{f}_m\|_{W(A)} \leq 2c_m\gamma^m. \tag{6.10}$$

Finally, we note that a residual bound may also be obtained by choosing a closed contour Γ in (6.3) containing $W(A)$ in its interior in the sense that $\delta := \mathrm{dist}(\Gamma, W(A)) > 0$

(Eiermann 1989). Indeed, for any $\zeta \in \Gamma$ and $\omega = (A\boldsymbol{v}, \boldsymbol{v}) \in W(A)$, $\|\boldsymbol{v}\| = 1$, we have by the Cauchy-Schwarz inequality

$$\delta \leq |\zeta - \omega| = |((\zeta I - A)\boldsymbol{v}, \boldsymbol{v})| \leq \|(\zeta I - A)\boldsymbol{v}\|,$$

and consequently $\|(\zeta I - A)^{-1}\| \leq \delta^{-1}$. Proceeding as in the derivation of the pseudospectral bound, we now obtain

$$\frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|}{\|\boldsymbol{r}_0\|} \leq \frac{\ell(\Gamma)}{2\pi\delta} \, \|p_m^{\mathrm{MR}}\|_\Gamma \leq \frac{\ell(\Gamma)}{2\pi\delta} \min_{\substack{p \in \mathscr{P}_m \\ p(0)=1}} \|p\|_\Gamma.$$

6.1.4 Summary

The bounds in the previous sections are all based on polynomial approximation problems on sets in the complex plane, and the differences lie in which sets are used, namely $\Lambda(A)$, $\Lambda_{\epsilon(A)}$ or $W(A)$. In the normal case the eigenvalue bounds tell all there is to know about the MR convergence behavior for given A, whereas for nonnormal A, the bounds derived from pseudospectra and the field of values try to compensate for the nonnormality by using larger sets where the residual polynomial needs to be small. In this way they often lead to tighter bounds than those based on eigenvalues because they do not explicitly contain the condition number of the eigenvector matrix T. An investigation into the relative strengths and weaknesses of these three types of bounds can be found in Embree (1999).

It is an open question whether these is indeed a set in the complex plane on which a given matrix A "lives" in the sense that solving a polynomial approximation problem on this set will lead to sufficiently tight convergence bounds, or whether it is necessary to include geometric information on the mapping properties of A in the general case.

Finally, we note that each of these approaches leads to upper bounds for the *linear* convergence rate, where it is known that Krylov subspace methods often converge faster than linear. Nevanlinna (1993) has proposed a qualitative theory for Krylov subspace methods in the abstract setting of operators on a Banach space, in which the linear phase of convergence is preceded by a *sublinear phase* and followed by a phase of *superlinear* convergence.

6.2 Convergence Bounds Based on Angles

As we saw in Chapter 2, the entire convergence history of an MR or OR iteration is contained in the sequence of angles between the spaces $\mathscr{V}_m$ and $\mathscr{W}_m$, which for Krylov subspace methods correspond to $\mathscr{K}_m$ and $A\mathscr{K}_m$. These angles are thus the invariants associated with the application of a Krylov subspace MR or OR method to the solution of the equation (1.1). As such, they also reflect the influence of the initial residual, which is suppressed in the polynomial bounds of Section 6.1.

In this section we shall apply the results of the abstract theory of Chapter 2 to recover some well-known residual and error bounds for Krylov subspace MR and OR approximations in Sections 6.2.1 and 6.2.2. In essence, this amounts to relating properties of A to the rate of decay of the sines of the angles arising in the iteration. In Section 6.2.3 this is

done for operator equations where the operator is a compact perturbation of the identity to show that, under mild assumptions, Krylov subspace MR and OR methods converge superlinearly. In Section 6.2.4 we derive the class of all Krylov subspace iterations which possess the same invariants and then draw some conclusions regarding which quantities can and cannot have an influence on the convergence rate.

6.2.1 Residual Bounds

By (2.5), the residual norm of the MR approximation of the solution of (1.1) with respect to the initial guess $\boldsymbol{x}_0 \in \mathscr{H}$ and the space $\mathscr{V} \subset \mathscr{H}$ is given by $\|\boldsymbol{r}\| = \|\boldsymbol{r}_0\| \sin \measuredangle(\boldsymbol{r}_0, A\mathscr{V})$, i.e., the factor by which the residual is reduced is given by the sine of the smallest angle formed by $\boldsymbol{r}_0$ and the image of a vector $\boldsymbol{v} \in \mathscr{V}$ under A. When we add the assumption of Section 2.2 that $\boldsymbol{r}_0 \in \mathscr{V}$ we can bound this factor as follows:

$$\begin{aligned}
\frac{\|\boldsymbol{r}\|^2}{\|\boldsymbol{r}_0\|^2} &= 1 - \cos^2 \measuredangle(\boldsymbol{r}_0, A\mathscr{V}) = 1 - \sup_{v \in \mathscr{V}} \frac{|(\boldsymbol{r}_0, A\boldsymbol{v})|^2}{\|\boldsymbol{r}_0\|^2 \, \|A\boldsymbol{v}\|^2} \\
&\leq 1 - \frac{|(\boldsymbol{r}_0, A\boldsymbol{r}_0)|^2}{\|\boldsymbol{r}_0\|^2 \, \|A\boldsymbol{r}_0\|^2} = 1 - \left| \frac{(A\boldsymbol{r}_0, \boldsymbol{r}_0)}{(\boldsymbol{r}_0, \boldsymbol{r}_0)} \right| \left| \frac{(\boldsymbol{r}_0, A\boldsymbol{r}_0)}{(A\boldsymbol{r}_0, A\boldsymbol{r}_0)} \right| \\
&\leq 1 - \inf_{v \in \mathscr{V}} \left| \frac{(A\boldsymbol{v}, \boldsymbol{v})}{(\boldsymbol{v}, \boldsymbol{v})} \right| \inf_{w \in A\mathscr{V}} \left| \frac{(A^{-1}\boldsymbol{w}, \boldsymbol{w})}{(\boldsymbol{w}, \boldsymbol{w})} \right| \\
&= 1 - \inf\{|z| : z \in W(A_{|\mathscr{V}})\} \cdot \inf\{|z| : z \in W(A^{-1}_{|A\mathscr{V}})\} \\
&\leq 1 - \inf\{|z| : z \in W(A)\} \cdot \inf\{|z| : z \in W(A^{-1})\} \\
&=: 1 - \nu(A)\nu(A^{-1}).
\end{aligned}$$

where $W(A)$ (cf. (4.9) and Section 6.1.3) denotes the field of values of A and we have assumed that A is invertible. We have shown

Theorem 6.2.1. *The residual $\boldsymbol{r}$ of the MR approximation to the solution of (1.1) on the subspace $\mathscr{V} \subset \mathscr{H}$ with initial residual $\boldsymbol{r}_0 \in \mathscr{V}$ satisfies*

$$\frac{\|\boldsymbol{r}\|}{\|\boldsymbol{r}_0\|} \leq \sqrt{1 - \nu(A)\nu(A^{-1})} \tag{6.11}$$

with $\nu(A)$ defined above.

Of course, the bound (6.11) only yields a reduction provided $0 \notin W(A)$, which also implies $0 \notin W(A^{-1})$ (cf. Horn & Johnson (1991, p. 66)).

If $\mathscr{V}_m$ is a sequence of Krylov spaces, we obtain

Corollary 6.2.2. *The MR residual with index m satisfies*

$$\frac{\|\boldsymbol{r}_m^{MR}\|}{\|\boldsymbol{r}_0\|} \leq \prod_{j=1}^{m} \sqrt{1 - \nu_j(A)\widetilde{\nu}_j(A^{-1})} \tag{6.12}$$

where the quantities $\nu_j(A)$ and $\widetilde{\nu}_j(A^{-1})$ are defined as

$$\begin{aligned}
\nu_j(A) &:= \inf\{|z| : z \in W(A_{|\mathscr{V}_j \cap \mathscr{W}_{j-1}^{\perp}})\} \qquad \text{and} \\
\widetilde{\nu}_j(A^{-1}) &:= \inf\{|z| : z \in W(A^{-1}_{|A(\mathscr{V}_j \cap \mathscr{W}_{j-1}^{\perp})})\}.
\end{aligned}$$

Proof. From (2.25), the fact that $s_j = \sin \measuredangle(\boldsymbol{r}_{j-1}^{\mathrm{MR}}, A\mathscr{V}_j)$ and $\boldsymbol{r}_{j-1}^{\mathrm{MR}} \in \mathscr{V}_j \cap \mathscr{W}_{j-1}^{\perp}$, we conclude

$$\begin{aligned}
\frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|}{\|\boldsymbol{r}_0\|} &= \prod_{j=1}^{m} s_j = \prod_{j=1}^{m} \sin \measuredangle(\boldsymbol{r}_{j-1}^{\mathrm{MR}}, A\mathscr{V}_j) = \prod_{j=1}^{m} \left(1 - \sup_{\boldsymbol{v} \in \mathscr{V}_j} \frac{|(\boldsymbol{r}_{j-1}^{\mathrm{MR}}, A\boldsymbol{v})|^2}{\|\boldsymbol{r}_{j-1}^{\mathrm{MR}}\|^2 \, \|A\boldsymbol{v}\|^2} \right)^{1/2} \\
&\le \prod_{j=1}^{m} \left(1 - \frac{|(\boldsymbol{r}_{j-1}^{\mathrm{MR}}, A\boldsymbol{r}_{j-1}^{\mathrm{MR}})|^2}{\|\boldsymbol{r}_{j-1}^{\mathrm{MR}}\|^2 \, \|A\boldsymbol{r}_{j-1}^{\mathrm{MR}}\|^2} \right)^{1/2} \\
&= \prod_{j=1}^{m} \left(1 - \left| \frac{(A\boldsymbol{r}_{j-1}^{\mathrm{MR}}, \boldsymbol{r}_{j-1}^{\mathrm{MR}})}{(\boldsymbol{r}_{j-1}^{\mathrm{MR}}, \boldsymbol{r}_{j-1}^{\mathrm{MR}})} \right| \left| \frac{(\boldsymbol{r}_{j-1}^{\mathrm{MR}}, A\boldsymbol{r}_{j-1}^{\mathrm{MR}})}{(A\boldsymbol{r}_{j-1}^{\mathrm{MR}}, A\boldsymbol{r}_{j-1}^{\mathrm{MR}})} \right| \right)^{1/2} \\
&\le \prod_{j=1}^{m} \left(1 - \nu_j(A) \widetilde{\nu}_j(A^{-1}) \right)^{1/2}.
\end{aligned}$$

□

Noting that $\nu_j(A) \ge \nu(A)$ and $\widetilde{\nu}_j(A^{-1}) \ge \nu(A^{-1})$, we obtain the simpler bound given in

Corollary 6.2.3. *The MR residual with index m satisfies*

$$\frac{\|\boldsymbol{r}_m^{MR}\|}{\|\boldsymbol{r}_0\|} \le \left(1 - \nu(A)\nu(A^{-1}) \right)^{m/2}. \tag{6.13}$$

If A is a positive real operator, i.e., if its Hermitian part $H := (A + A^*)/2$ is positive definite, then $\nu(A) = \lambda_{\min}(H)$ and

$$\nu(A^{-1}) = \inf_{\boldsymbol{v} \in \mathscr{H}} \frac{(A^{-1}\boldsymbol{v}, \boldsymbol{v})}{(\boldsymbol{v}, \boldsymbol{v})} = \inf_{\boldsymbol{w} \in \mathscr{H}} \frac{(\boldsymbol{w}, A\boldsymbol{w})}{(\boldsymbol{w}, \boldsymbol{w})} \frac{(\boldsymbol{w}, \boldsymbol{w})}{(A\boldsymbol{w}, A\boldsymbol{w})} \ge \frac{\lambda_{\min}(H)}{\|A\|^2},$$

in view of which Corollary 6.2.3 yields a bound first given by Elman (1982):

$$\frac{\|\boldsymbol{r}_m\|}{\|\boldsymbol{r}_0\|} \le \left(1 - \frac{\lambda_{\min}(H)^2}{\lambda_{\max}(A^T A)} \right)^{m/2}.$$

Remark 6.2.4. If, in the derivation of the residual bound (6.11), one makes the slightly cruder estimate

$$\frac{\|\boldsymbol{r}\|^2}{\|\boldsymbol{r}_0\|^2} \le 1 - \frac{|(\boldsymbol{r}_0, A\boldsymbol{r}_0)|^2}{\|\boldsymbol{r}_0\|^2 \, \|A\boldsymbol{r}_0\|^2} \le 1 - \inf_{\boldsymbol{r}_0 \in \mathscr{H}} \frac{|(\boldsymbol{r}_0, A\boldsymbol{r}_0)|^2}{\|\boldsymbol{r}_0\|^2 \, \|A\boldsymbol{r}_0\|^2} =: \sin^2(\gamma(A)),$$

where $\gamma(A)$ is the largest angle between a nonzero vector $\boldsymbol{v} \in \mathscr{H}$ and its image $A\boldsymbol{v}$. One thus obtains a bound $\sin^m \gamma(A)$ on the residual reduction after m steps. The angle $\gamma(A)$ was introduced by Wielandt (1996) in 1967. See also Gustafson & Rao (1996).

6.2.2 Error Bounds

When solving an equation (1.1) approximately using successive iterates $\boldsymbol{x}_m$, the residual $\boldsymbol{r}_m = \boldsymbol{b} - A\boldsymbol{x}_m$ may be the only computable indication of the progress of the solution process. The quantity of primary interest, however, is usually he error $\boldsymbol{e}_m = \boldsymbol{x} - \boldsymbol{x}_m = A^{-1}\boldsymbol{r}_m$.

For Krylov subspace methods, we have $\boldsymbol{x}_m \in \boldsymbol{x}_0 + \mathscr{K}_m$, so that $\boldsymbol{e}_m = \boldsymbol{e}_0 - \boldsymbol{c}$ for some correction $\boldsymbol{c} \in \mathscr{K}_m$. Of course, the best one could to is to select this correction as the best approximation to $\boldsymbol{e}_0$ from $\mathscr{K}_m = \mathscr{K}_m(A, \boldsymbol{r}_0) = A\mathscr{K}_m(A, \boldsymbol{e}_0)$. This would correspond to computing the MR approximation of $\boldsymbol{e}_0$ with respect to the sequence of approximation spaces $\mathscr{W}_m = A\mathscr{K}_m(A, \boldsymbol{e}_0)$, a process which would require knowledge of the initial error and hence the solution $\boldsymbol{x}$. The relation between residuals and errors, however, allows us to bound the error of the MR approximation with respect to this best possible approximation:

Lemma 6.2.5. *The error $\boldsymbol{e}_m^{MR}$ of the MR approximation satisfies*

$$\|\boldsymbol{e}_m^{MR}\| \le \kappa(A) \inf_{\boldsymbol{v}\in\mathscr{V}_m} \|\boldsymbol{e}_0 - \boldsymbol{v}\| \tag{6.14}$$

where $\kappa(A) = \|A\|\,\|A^{-1}\|$ denotes the condition number of A.

Proof. With $\mathscr{V}_m = \mathscr{K}_m(\boldsymbol{r}_0, A)$ and $\mathscr{W}_m = A\mathscr{V}_m$, there holds

$$\|\boldsymbol{r}_m^{\mathrm{MR}}\| = \min_{\boldsymbol{w}\in\mathscr{W}_m} \|\boldsymbol{r}_0 - \boldsymbol{w}\| = \min_{\boldsymbol{v}\in\mathscr{V}_m} \|A(\boldsymbol{e}_0 - \boldsymbol{v})\| \le \|A\| \min_{\boldsymbol{v}\in\mathscr{V}_m} \|\boldsymbol{e}_0 - \boldsymbol{v}\|,$$

and thus the assertion follows from $\|\boldsymbol{e}_m^{\mathrm{MR}}\| = \|A^{-1}\boldsymbol{r}_m^{\mathrm{MR}}\| \le \|A^{-1}\|\,\|\boldsymbol{r}_m^{\mathrm{MR}}\|$. □

Thus, the error of the MR approximation is within the condition number of A of the error of the best approximation to $\boldsymbol{e}_0$ from the Krylov space. In view of the relation (2.26), this translates to the following bound for the OR error:

$$\|\boldsymbol{e}_m^{\mathrm{OR}}\| \le \frac{\kappa(A)}{c_m} \inf_{\boldsymbol{v}\in\mathscr{V}_m} \|\boldsymbol{e}_0 - \boldsymbol{v}\|. \tag{6.15}$$

However, a stronger bound can be obtained if the field of values of A is bounded away from the origin:

Theorem 6.2.6. *If $\inf\{|z| : z \in W(A)\} \ge \alpha$ for some $\alpha > 0$, then the OR error satisfies*

$$\frac{\|\boldsymbol{e}_m^{OR}\|}{\|\boldsymbol{e}_0\|} \le \frac{\|A\|}{\alpha} \inf_{\boldsymbol{v}\in\mathscr{V}_m} \|\boldsymbol{e}_0 - \boldsymbol{v}\|.$$

Proof. From the characterization of the OR approximation we have $\boldsymbol{e}_m^{\mathrm{OR}} = \boldsymbol{e}_0 - \boldsymbol{v}$ for some $\boldsymbol{v} \in \mathscr{K}_m(\boldsymbol{r}_0, A)$ subject to

$$\boldsymbol{r}_m^{\mathrm{OR}} = A\boldsymbol{e}^{\mathrm{OR}} \perp \mathscr{K}_m(\boldsymbol{r}_0, A) \quad \Leftrightarrow \quad \boldsymbol{e}_m^{\mathrm{OR}} \perp A^*\mathscr{K}_m(\boldsymbol{r}_0, A).$$

This means that the OR error is obtained as the error of an OR approximation of $\boldsymbol{e}_0$ from the space $\mathscr{K}_m(\boldsymbol{r}_0, A)$ orthogonal to $A^*\mathscr{K}_m(\boldsymbol{r}_0, A)$. Thus, with $\mathscr{V}_m = \mathscr{K}_m(\boldsymbol{r}_0, A)$, the results on the norm of an oblique projection give

$$\frac{\|\boldsymbol{e}^{\mathrm{OR}}\|}{\|\boldsymbol{e}_0\|} \le \|I - P_{\mathscr{V}_m}^{A^*\mathscr{V}_m}\| = \frac{1}{\cos\measuredangle(\mathscr{V}_m, A^*\mathscr{V}_m)}.$$

We bound the cosine of the largest canonical angle between $\mathscr{V}_m$ and $A^*\mathscr{V}_m$ by

$$\begin{aligned}\cos^2 \measuredangle(\mathscr{V}_m, A^*\mathscr{V}_m) &= \sup_{\boldsymbol{u}\in\mathscr{V}_m} \sup_{\boldsymbol{v}\in\mathscr{V}_m} \frac{|(\boldsymbol{u}, A^*\boldsymbol{v})|^2}{\|\boldsymbol{u}\|^2\,\|A^*\boldsymbol{v}\|^2} \\ &\geq \sup_{\boldsymbol{u}\in\mathscr{V}_m} \frac{|(\boldsymbol{u}, A^*\boldsymbol{u})|^2}{\|\boldsymbol{u}\|^2\,\|A^*\boldsymbol{u}\|^2} = \sup_{\boldsymbol{u}\in\mathscr{V}_m} \frac{(A\boldsymbol{u}, \boldsymbol{u})}{(\boldsymbol{u}, \boldsymbol{u})} \frac{(A^*\boldsymbol{u}, \boldsymbol{u})}{(A^*\boldsymbol{u}, A^*\boldsymbol{u})} \\ &\geq \inf_{v\in\mathscr{V}_m} \left|\frac{(A\boldsymbol{u}, \boldsymbol{u})}{(\boldsymbol{u}, \boldsymbol{u})}\right| \inf_{v\in\mathscr{V}_m} \left|\frac{(A^*\boldsymbol{u}, \boldsymbol{u})}{(A^*\boldsymbol{u}, A^*\boldsymbol{u})}\right| \geq \frac{\alpha^2}{\|A\|^2}\end{aligned}$$

□

An important class of operators consists of those with a positive definite Hermitian part $H := (A + A^*)/2$. These operators are sometimes referred to as *positive real*. In this case the inner product $(H\cdot, \cdot)$ induced by H defines a norm on $\mathscr{H}$. The next theorem, which is due to Starke (1994), shows that the OR error measured in this norm is optimal up to a factor which depends on the skew-Hermitian part $S := (A - A^*)/2$.

Theorem 6.2.7. *If A is positive real with Hermitian and skew-Hermitian parts H and S, then the OR error satisfies*

$$\|\boldsymbol{e}_m^{OR}\|_H \leq \big(1 + \rho(H^{-1}S)\big) \inf_{\boldsymbol{v}\in\mathscr{V}_m} \|\boldsymbol{e}_0 - \boldsymbol{v}\|_H,$$

where $\rho(H^{-1}S)$ denotes the spectral radius of $H^{-1}S$.

Proof. Since $\boldsymbol{r}_m^{\mathrm{OR}} = A\boldsymbol{e}_m^{\mathrm{OR}} \perp \mathscr{V}_m$, we have, noting $(H\boldsymbol{v}, \boldsymbol{v}) = \mathrm{Re}(A\boldsymbol{v}, \boldsymbol{v})$ and $(S\boldsymbol{v}, \boldsymbol{v}) = \mathrm{Im}(A\boldsymbol{v}, \boldsymbol{v})$ for $\boldsymbol{v} \in \mathscr{V}_m$,

$$\begin{aligned}\|\boldsymbol{e}_m^{\mathrm{OR}}\|_H^2 &= (H\boldsymbol{e}_m^{\mathrm{OR}}, \boldsymbol{e}_m^{\mathrm{OR}}) \\ &\leq |(A\boldsymbol{e}_m^{\mathrm{OR}}, \boldsymbol{e}_m^{\mathrm{OR}})| = |(A\boldsymbol{e}_m^{\mathrm{OR}}, \boldsymbol{e}_0)| = |(A\boldsymbol{e}_m^{\mathrm{OR}}, \boldsymbol{e}_0 - \boldsymbol{v})|\end{aligned}$$

for arbitrary $\boldsymbol{v} \in \mathscr{V}_m$, and therefore

$$\|\boldsymbol{e}_m^{\mathrm{OR}}\|_H^2 \leq |(H\boldsymbol{e}_m^{\mathrm{OR}}, \boldsymbol{e}_0 - \boldsymbol{v}) + (S\boldsymbol{e}_m^{\mathrm{OR}}, \boldsymbol{e}_o - \boldsymbol{v})|.$$

The first term is bounded by $\|\boldsymbol{e}_0 - \boldsymbol{v}\|_H \|\boldsymbol{e}_m^{\mathrm{OR}}\|_H$, and for the second term we obtain

$$\begin{aligned}|(S\boldsymbol{e}_m^{\mathrm{OR}}, \boldsymbol{e}_o - \boldsymbol{v})| &= |(H^{1/2}H^{-1/2}SH^{-1/2}H^{1/2}\boldsymbol{e}_m^{\mathrm{OR}}, \boldsymbol{e}_0 - \boldsymbol{v})| \\ &\leq \|H^{-1/2}SH^{-1/2}H^{1/2}\boldsymbol{e}_m^{\mathrm{OR}}\| \|\boldsymbol{e}_0 - \boldsymbol{v}\|_H \\ &\leq \|H^{-1/2}SH^{-1/2}\| \|\boldsymbol{e}_m^{\mathrm{OR}}\|_H \|\boldsymbol{e}_0 - \boldsymbol{v}\|_H \\ &= \rho(H^{-1}S) \|\boldsymbol{e}_m^{\mathrm{OR}}\|_H \|\boldsymbol{e}_0 - \boldsymbol{v}\|_H,\end{aligned}$$

which, together with the bound for the first term, yields the assertion. □

An immediate consequence of Theorem 6.2.7 is that, for a Hermitian positive definite operator A, the OR method, which simplifies to the well-known conjugate gradient method in this case, yields the best approximation in the A-norm.

Remark 6.2.8. In view of the remark preceding Lemma 6.2.5 that the best approximation of the initial error $\boldsymbol{e}_0$ from the Krylov space $\mathscr{K}_m(A, \boldsymbol{r}_0) = A\mathscr{K}_m(A, \boldsymbol{e}_0)$ has the same structure as the best approximation of $\boldsymbol{r}_0$ from $A\mathscr{K}(A, \boldsymbol{r}_0)$ with $\boldsymbol{r}_0$ replaced by $\boldsymbol{e}_0$, the infimum in (6.14) and (6.15) may be bounded in an analogous manner to the MR residual in Theorem 6.2.1 and Corollaries 6.2.2 and 6.2.3.

6.2.3 An Application: Compact Operators

Many applications such as the solution of elliptic boundary value problems by the integral equation method require the solution of second-kind Fredholm equations, i.e., operator equations (1.1) in which A has the form $A = \lambda I + K$ with $\lambda \neq 0$ and $K : \mathscr{H} \to \mathscr{H}$ a compact operator. The development of fast multiplication algorithms (cf. Greengard & Rokhlin (1987), Hackbusch & Nowak (1989)) has made Krylov subspace methods attractive as solution algorithms for discretizations of these problems, since they require only applications of the (discrete) operator to vectors. Moreover, as shown by Moret (1997) for GMRES and by Winther (1980) for CG, Krylov subspace methods converge q-superlinearly for operator equations involving compact perturbations of (multiples of) the identity.

The simple reason for this is that, for these operators, the sines s_m of the angles between the Krylov space $\mathscr{V}_m = \mathscr{K}_m$ and $\mathscr{W}_m = A\mathscr{K}_m$ converge to zero. To show this, we recall a basic result on compact operators and orthonormal systems:

Theorem 6.2.9. *Let $K : \mathscr{H} \to \mathscr{H}$ be a compact linear operator and $\{\boldsymbol{v}_m\}_{m\geq 1} \subset \mathscr{H}$ be an orthonormal system. Then*

$$\lim_{m\to\infty} (K\boldsymbol{v}_m, \boldsymbol{v}_{m+1}) = 0.$$

Proof. See e.g. Ringrose (1971). □

Corollary 6.2.10. *Let $A = \lambda I + K$ with $\lambda \neq 0$ and $K : \mathscr{H} \to \mathscr{H}$ compact, let $\{\boldsymbol{v}_j\}_{j\geq 1}$ denote the orthonormal system generated by the Arnoldi process applied to A and $\boldsymbol{r}_0 \in \mathscr{H}$, and let $\{\boldsymbol{w}_j\}_{j\geq 1}$ be an orthonormal system such that $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ is an orthonormal basis of $\mathscr{W}_m = A\mathscr{V}_m$. Then the sines s_m of largest canonical angle between $\mathscr{V}_m$ and $\mathscr{W}_m$ form a null sequence.*

Proof. Lemmas 2.2.4 and 4.2.3 combined with (2.43) yield

$$|s_m| = |(\boldsymbol{v}_{m+1}, \boldsymbol{w}_m)| \leq (\boldsymbol{v}_{m+1}, A\boldsymbol{v}_m)\, \|A^{-1}\| = (\boldsymbol{v}_{m+1}, K\boldsymbol{v}_m)\, \|A^{-1}\| \to 0,$$

since A^{-1} is bounded whenever $\lambda \neq 0$. □

In particular, since $s_m \to 0$ implies that $|s_m| \neq 1$ for m sufficiently large, this means that the OR approximation is always defined except for possibly a finite number of indices. Moreover, as $|s_m|$ is bounded away from one, c_m is accordingly bounded away from zero, hence the relation (2.26) also implies the q-superlinear convergence of the OR approximation. We summarize this result in the following theorem.

Theorem 6.2.11. *Given $K : \mathscr{H} \to \mathscr{H}$ compact, $0 \neq \lambda \in \mathbb{C}$ and $\boldsymbol{b} \in \mathscr{H}$, let $\boldsymbol{x}_0 \in \mathscr{H}$ be an initial guess at the solution of (1.1) with $A = \lambda I + K$. Then the OR approximation with respect to the spaces $\mathscr{V}_m = \mathscr{K}_m(\boldsymbol{r}_0, A)$ and $\mathscr{W}_m = A\mathscr{V}_m$ with initial residual vector $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$ exists for all sufficiently large indices. Moreover, the sequence of MR and OR approximations converge q-superlinearly.*

We remark that the rate of superlinear convergence may be quantified in terms of the rate of decay of the singular values of K (Moret 1997). We also note that, in view of (2.70), this result applies to all MR/OR pairs of Krylov subspace methods including QMR/BCG given a bound on the conditioning of the basis of the Krylov space being used. For bases generated by the look-ahead Lanczos method, such bounds are guaranteed e.g. by the implementation given by Freund et al. (1993).

6.2.4 Parameters Determining the Rate of Convergence

As stated in the introduction to Section 6.2, the invariants of the MR and OR processes are the angles between the spaces $\mathscr{K}_m$ and $A\mathscr{K}_m$. Furthermore, Theorem 2.3.5 and the ensuing discussion revealed that the sines and cosines of these angles appear explicitly as the parameters of the Givens rotations in the recursive construction of the QR-factorizations (2.49) of the Hessenberg matrices $\widetilde{H}_m$ relating the bases W_m and V_{m+1} of the approximation space and the residual space in (2.39), which was the abstract counterpart of the Arnoldi decomposition

$$AV_m = V_{m+1}\widetilde{H}_m = V_{m+1}Q_m \begin{bmatrix} R_m \\ \mathbf{0} \end{bmatrix}, \qquad m = 1, \ldots, L. \tag{6.16}$$

As a consequence, all convergence information up to step m is contained in the unitary matrix Q_m.

MR-equivalent matrices

From relation (6.16), we conclude that there exists a family of matrices $\widetilde{A}$ which produce the identical MR/OR residual sequence for the given initial residual $\boldsymbol{r}_0$.

Proposition 6.2.12. *For $A \in \mathbb{C}^{n\times n}$ and $\boldsymbol{r}_0 \in \mathbb{C}^n$, denote by L the termination index* (4.5)*, and by*

$$AV_L = V_L Q_{L-1} R_L = \widehat{V}_L R_L \tag{6.17}$$

the Arnoldi decomposition (6.16) *in the last step $m = L$. Let further $\widetilde{R}_L$ be any nonsingular upper triangular matrix and define $\widetilde{A} \in \mathbb{C}^{n\times n}$ by*

$$\widetilde{A}V_L = V_L Q_{L-1}\widetilde{R}_L = \widehat{V}_L\widetilde{R}_L. \tag{6.18}$$

Then A and $\widetilde{A}$ are "MR-equivalent" in the sense that the MR method produces identical residual vectors $\boldsymbol{r}_m^{MR}$ for both systems $A\boldsymbol{x} = \boldsymbol{b}$ and $\widetilde{A}\boldsymbol{x} = \widetilde{\boldsymbol{b}}$ provided the starting vectors $\boldsymbol{x}_0$ and $\widetilde{\boldsymbol{x}}_0$ are chosen such that $\boldsymbol{b} - A\boldsymbol{x}_0 = \widetilde{\boldsymbol{b}} - \widetilde{A}\widetilde{\boldsymbol{x}}_0$.

Proof. The assertion follows immediately since both matrices produce the same sequence of approximation spaces. □

Remark 6.2.13. We note that, by the same reasoning, A and $\widetilde{A}$ are also *OR-equivalent*.

The family of matrices (6.18) is actually a complete parameterization of all matrices which are MR-equivalent to A with regard to $\boldsymbol{r}_0$.

Proposition 6.2.14. *If* (6.17) *denotes the Arnoldi factorization in the last step $m = L$ of the MR process, then* (6.18) *contains all matrices which are MR-equivalent to A with respect to $\boldsymbol{r}_0$ if $\widetilde{R}$ is allowed to vary over all nonsingular upper triangular $L \times L$ matrices.*

Proof. $\widetilde{A}$ is MR-equivalent to A with respect to $\boldsymbol{r}_0$ if and only if both generate the same sequence of approximation spaces, i.e.,

$$A\mathscr{K}_m(A, \boldsymbol{r}_0) = \widetilde{A}\mathscr{K}_m(\widetilde{A}, \boldsymbol{r}_0), \qquad m = 1, \ldots L.$$

In view of

$$\mathscr{K}_m(A, \boldsymbol{r}_0) = \operatorname{span}\{\boldsymbol{r}_0\} + A\mathscr{K}_{m-1}(A, \boldsymbol{r}_0) = \operatorname{span}\{\boldsymbol{r}_0\} + \widetilde{A}\mathscr{K}_{m-1}(\widetilde{A}, \boldsymbol{r}_0) = \mathscr{K}_m(\widetilde{A}, \boldsymbol{r}_0),$$

which holds for $m = 2, \ldots, L$ (the assertion is trivial for $m = 1$), this implies that both matrices also generate the same sequence of Krylov spaces, which possess the same ascending orthonormal basis V_L. Hence all possible factorizations of the type given in (6.18) result from varying the nonsingular upper triangular matrix $\widetilde{R}$. □

A particularly simple representative of this family of MR-equivalent matrices is obtained by setting $\widetilde{R}$ to be the identity, resulting in the decomposition $\widetilde{A}V_L = V_L Q_{L-1}$. If $\mathscr{V}_L = \mathscr{K}_L$ is a proper subspace of $\mathscr{H}$, $\widetilde{A}$ may be made unique by extending it to be the identity on the complementary space. This means that, with $\boldsymbol{r}_0$ as initial vector, the isometry Q_{L-1} generates the same sequence of Krylov spaces, approximation spaces and associated angles as obtained with A, hence the identical sequence of MR—and, when these exist, OR—residual vectors result. Since Q_{L-1} is in particular a normal matrix, its MR-behavior is determined by its eigenvalues, which lie on the unit circle, and application of (6.5) yields

$$\frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|}{\|\boldsymbol{r}_0\|} \leq \min_{\substack{p \in \mathscr{P}_m \\ p(0)=1}} \|p\|_{\Lambda(Q_{L-1})}. \tag{6.19}$$

The class of MR-equivalent matrices was first considered by Greenbaum & Strakoš (1994), who give a slightly different parameterization: If W_L denotes the ascending basis of the sequence of approximation spaces $\mathscr{W}_m = A\mathscr{K}_m$, then an analogous Arnoldi decomposition $AW_L = W_L H_L^{(1)}$ exists for W_L, and it is shown there that the MR-equivalent matrices are given by

$$\widetilde{A}W_L = W_L \widetilde{R} H_L^{(1)} \tag{6.20}$$

with any nonsingular upper triangular $\widetilde{R} \in \mathbb{C}^{L \times L}$. To relate this characterization to (6.18), note that the (essentially unique) ascending orthonormal basis of the approximation space is given by $W_L = V_L Q_{L-1}$, resulting in

$$AW_L = AV_L Q_{L-1} = V_L H_L Q_{L-1} = V_L Q_{L-1} R_L Q_{L-1} = W_L H_m^{(1)},$$

which shows that $H_L^{(1)} = R_L Q_{L-1}$ may be obtained from H_L by one step of the QR algorithm. Finally, Greenbaum and Strakoš give yet another parameterization

$$\widetilde{A}W_L = W_L \widehat{R} \widehat{H}_L, \tag{6.21}$$

where $\widehat{R}$ again varies over all nonsingular upper triangular matrices and the Hessenberg matrix $\widehat{H}_L$ has the particularly simple structure

$$\widehat{H}_L = \begin{bmatrix} \boldsymbol{u}_2 & \ldots & \boldsymbol{u}_L & \boldsymbol{z} \end{bmatrix}, \qquad \boldsymbol{z} = \frac{-1}{(\boldsymbol{r}_0, \boldsymbol{w}_L)} \begin{bmatrix} -1 \\ (\boldsymbol{r}_0, \boldsymbol{w}_1) \\ \vdots \\ (\boldsymbol{r}_0, \boldsymbol{w}_{L-1}) \end{bmatrix},$$

where $W_L = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_L]$. This follows from the fact that $[\boldsymbol{r}_0, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{L-1}]\widehat{H}_L = W_L$ and that, for any MR-equivalent matrix $\widetilde{A}$, $\widetilde{A}[\boldsymbol{r}_0, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{L-1}] = W_L\widehat{R}$ for some nonsingular upper triangular matrix $\widehat{R}$. By selecting $\widetilde{R}$ and $\widehat{R}$ as the inverse of the R-factor in an RQ-factorization of the Hessenberg matrices in (6.20) and (6.21), respectively, Greenbaum and Strakoš also obtain unitary matrices which are MR-equivalent with A. As noted there, it is not surprising that any MR behavior may be observed for a unitary matrix, since the two extremes, namely stagnation until the last step (cyclic shift) and convergence in one step (the identity) are both obtained for unitary matrices.

Recent work along these lines by Liesen (1998) attempts to use MR-equivalent unitary matrices in order to "extrapolate" the convergence rate of an MR iteration at later stages from information available in earlier phases of the iteration. Using a unitary matrix Q_m which is MR-equivalent to the given system up to step m, Liesen uses bounds of the type (6.19) to determine the current linear convergence rate, which is then viewed as an indicator of the convergence rate to be expected during the remainder of the iteration. The spectral inclusion set in this case is the smallest segment of the unit circle containing the eigenvalues of Q_m, the asymptotic convergence rate of which is given by (6.8).

Its is, however, clear that this approach works only under additional assumptions since, in general, Q_m need not contain any information about the progress of the MR iteration beyond step m. It is not difficult to show that a given initial sequence of residual norms $\|\boldsymbol{r}_0\| \geq \cdots \geq \|\boldsymbol{r}_m^{\mathrm{MR}}\|$ can be complemented e.g., in such a way that $\|\boldsymbol{r}_m^{\mathrm{MR}}\| = \|\boldsymbol{r}_{m+1}^{\mathrm{MR}}\| = \ldots = \|\boldsymbol{r}_{n-1}^{\mathrm{MR}}\|$, i.e., that the iteration stagnates from step $m+1$ until step $n-1$.

Finally, we mention another recent approach which seeks more easily analyzable operators to bound the MR convergence rate for a given system: Huhtanen & Nevanlinna (1998) propose decomposing A into the sum of a normal operator and one of minimal rank, and argue that often the normal component determines the MR convergence behavior.

Singular Values

To conclude this section, we turn to the role that singular values play in the context of the convergence of the MR method. It has often been suggested in the literature that analyzing the singular values of either A or its projections H_m and $\widetilde{H}_m$ may reveal information on the convergence of GMRES. As shown above in the case of eigenvalues, we now show that singular value information alone says nothing about convergence.

Proposition 6.2.15. *To a given linear system of equations* (1.1)*, there is a linear system with coefficient matrix $\widetilde{A}$ for which GMRES exhibits the identical convergence history, and for which $\widetilde{A}$ has arbitrarily prescribed singular values.*

Proof. Denote by L the termination index of A with respect to the given system as defined in (4.14). For any Krylov subspace method applied to the this system, only the L singular values of $A_{\mathscr{K}_L(A,\boldsymbol{r}_0)}$ are noticeable, hence A can be defined arbitrarily on $\mathscr{K}_L(A,\boldsymbol{r}_0)^\perp$. To prescribe the singular values of $A_{\mathscr{K}_L(A,\boldsymbol{r}_0)}$, denote by $\widetilde{R}_L$ an upper triangular $L \times L$ matrix $\widetilde{R}_L$ possessing L arbitrary singular values. With $\widehat{V}_L$ denoting the Paige-Saunders basis (cf. Proposition 2.3.6), the matrix $\widetilde{A}$ by $\widetilde{A}\widehat{V}_L = \widehat{V}_L\widetilde{R}_L$ clearly possesses the same set of singular values and, by Proposition 6.2.12, $\widetilde{A}$ is "MR-equivalent" to A. □

By the same technique, we can prescribe, for fixed m, the singular values of $\widetilde{H}_m$ since, in view of (2.49) these coincide with those of R_m.

For the singular values of the square Hessenberg matrices H_m there is a slight complication. From (2.53) and (2.55) we see that a QR-factorization of H_m is given by

$$H_m = Q_{m-1}^H \begin{bmatrix} R_{m-1} & \boldsymbol{r} \\ \mathbf{0} & \tau \end{bmatrix}$$

The mth plane rotation is determined so that the vector $[\tau \ \eta_{m+1,m}]^\top$ is rotated to the vector $[r_{m,m} \ 0]^\top$, where $r_{m,m}$ is the entry in the (m,m)-position of R_m. This implies that $\tau = c_m r_{m,m}$. Thus, if we prescribe R_m to be a diagonal matrix, then the singular values of H_m are given by $|r_{1,1}|, \ldots, |r_{m-1,m-1}|, c_m|r_{m,m}|$ and thus can be selected arbitrarily. The only exception occurs when $c_m = 0$, i.e., H_m is singular and clearly only $m-1$ singular values can be chosen freely.

We remark that the same proof shows that neither the singular values of A nor those of $\widetilde{H}_m$ or H_m have any influence on the convergence behavior of an OR method.

6.3 Convergence Criteria for Restarted MR Methods

Since practical limitations often require that MR methods be restarted (cf. Chapter 5), his brings up the question of what can be said about the convergence of restarted MR methods. Unfortunately, there are as of yet few satisfactory results.

We mention two convergence characterizations, which follow immediately from (b) and (c) in Proposition 5.1.1, respectively. We denote by MR(m) the MR iteration with cycle length m when not referring to a specific algorithm, such as GMRES(m).

Theorem 6.3.1. *The MR(m) iteration converges for any initial residual $\boldsymbol{r}_0$ (or equivalently, for any right-hand side $\boldsymbol{b}$ and any starting vector $\boldsymbol{x}_0$) if and only if*

$$\mathbf{0} \notin W_m(A) := \{\, [(A\boldsymbol{v}, \boldsymbol{v}), (A^2\boldsymbol{v}, \boldsymbol{v}), \ldots, (A^m\boldsymbol{v}, \boldsymbol{v})]^\top \,:\, \|\boldsymbol{v}\| = 1\} \subset \mathbb{C}^m.$$

An equivalent condition is

$$\mu_m(A) := \max_{\|\boldsymbol{v}\|=1} \sin \measuredangle(\boldsymbol{v}, A\mathscr{K}_m(A, \boldsymbol{v})) < 1.$$

The first characterization is due to Faber et al. (1996), who used the set $W_m(A) \subset \mathbb{C}^m$ to construct a matrix A for which

$$\max_{\|\boldsymbol{v}\|=1} \min_{p \in \mathscr{P}_m,\, p(0)=1} \|p(A)\boldsymbol{v}\| < \min_{p \in \mathscr{P}_m,\, p(0)=1} \max_{\|\boldsymbol{v}\|=1} \|p(A)\boldsymbol{v}\|,$$

which shows that there exist linear systems (1.1) for which the least effective reduction of the initial residual by m steps of MR over all possible initial vectors is strictly less than that of the best polynomial iterative method using a fixed polynomial of degree m.

The set $W_m(A) \subset \mathbb{C}^m$ is sometimes referred to as the *generalized field of values* of A because, for $m = 1$, it reduces to the ordinary field of values of A,

$$W_1(A) = W(A) = \{(A\boldsymbol{v}, \boldsymbol{v}) \,:\, \|\boldsymbol{v}\| = 1\} \subset \mathbb{C}.$$

Many properties of the field of values $W(A)$ of A are known, the most prominent among them being that $W(A)$ is convex for any bounded linear operator A. For $m > 1$, $W_m(A)$ is much less understood (Faber et al. (1996) show, e.g., that $W_m(A)$ is convex for nonsingular normal matrices as well as for upper triangular Toeplitz matrices, but also a matrix $A \in \mathbb{R}^{4\times 4}$ is constructed for which $W_3(A)$ fails to be convex).

Theorem 6.3.1 leads to an upper bound for the reduction of the residual by, say, j cycles of MR(m):

$$\|\boldsymbol{r}_{jm}\| \leq \mu_m(A)\|\boldsymbol{r}_{(j-1)m}\| \leq \cdots \leq \mu_m(A)^j\|\boldsymbol{r}_0\|. \tag{6.22}$$

However, this estimate has no practical value since, particularly for $m > 1$, there is no known way of computing or bounding $\mu_k(A)$ with acceptable computational cost. The distance between 0 and $W_m(A)$ gives rise to another upper bound for the norm of the residuals generated by MR(m), (cf. (Faber et al. 1996, Theorem 2.9)) but, as with (6.22), its practical relevance is questionable as well.

In the case of $m = 1$, we obtain directly from Theorem 6.2.1

$$\mu_1(A)^2 = 1 - \min\{|\omega| \,:\, \omega \in W(A)\} \cdot \inf\{|\omega| \,:\, \omega \in W(A^{-1})\}.$$

Similar manipulations are the basis of many (if not all) existing residual bounds for MR(1), as can be found in e.g. (Eisenstat et al. 1983, Joubert 1994*a*, Saad 1996, Saad 1997*b*, Greenbaum 1997).

We finally mention that Theorem 6.3.1 immediately implies a sufficient condition for the convergence of MR(m) which is sometimes easier to check:

Corollary 6.3.2. *If at least one of the m fields of values $W_1(A)$, $W_1(A^2), \ldots, W_1(A^m)$ of A, A^2, $\ldots$, A^m, respectively, does not contain the origin, then MR(m) converges to the solution of* (1.1) *for any right-hand side $\boldsymbol{b}$ and any starting vector $\boldsymbol{x}_0$.*

Since the (ordinary) field of values is always convex, it does not contain the origin if and only if it is contained in a half-plane $\Upsilon \subset \mathbb{C}$ with $0 \notin \Upsilon$. The assumption of Corollary 6.3.2 is therefore known as the *half-plane condition*. Easy consequences of it are for instance that MR(1) converges for any positive definite (negative definite) selfadjoint linear operator A, but also for those which possess a positive definite selfadjoint part $(A + A^*)/2$. In addition, MR(2) converges for any nonsingular operator A which is either selfadjoint or skew-selfadjoint (because $W_1(A^2)$ is then contained in the right or left half-plane, respectively). Further results along these lines are contained, e.g., in Strikwerda & Stodder (1995).

Theorem 6.3.1 shows that, in the case of $m = 1$, the half-plane condition (which reduces then to $0 \notin W(A)$) is not only sufficient but also necessary for the convergence of MR(1) (for any initial residual vector). The following example shows this is no longer true if $m > 1$. The field of values of the matrix

$$A = \begin{bmatrix} 1 & \gamma & 0 \\ 0 & 1 & \gamma \\ 0 & 0 & 1 \end{bmatrix}, \qquad \gamma > 0,$$

is the closed disk of radius $\gamma/\sqrt{2}$ centered at one, i.e., MR(1) is guaranteed to converge if and only if $\gamma < \sqrt{2}$. It is easy to see that the smallest eigenvalue $\theta(\gamma)$ of the symmetric

matrix $(A^2 + A^{2\top})/2$ is a monotonically decreasing function of γ with a zero at $\sqrt{6}/3 = 0.81\ldots$, i.e., $0 \notin W(A^2)$ if and only if $\gamma < \sqrt{6}/3$. Corollary 6.3.2 therefore ensures the convergence of MR(2) for $\gamma < \sqrt{2} = 1.41\ldots$. Using a technique of Strikwerda and Stodder (Strikwerda & Stodder 1995, Theorem 4.1) to determine the smallest value of γ for which null vector is contained in $W_2(A)$ for this particular example, one can, however, show that MR(2) converges (for any $\boldsymbol{r}_0$) if and only if $\gamma < 2\sqrt{6}/3 = 1.63\ldots$.

Bibliography

Arnoldi, Walter E. (1951), 'The principle of minimized iterations in the solution of the matrix eigenvalue problem', *Quart. Appl. Math.* **9**, 17–29.

Ashby, Steven F. & Martin H. Gutknecht (1993), A matrix analysis of conjugate gradient algorithms, *in* M. N.et al., ed., 'Parallel processing for scientific computing. Proceedings of the 9th PCG symposium, held at Keio University, Tokyo, Japan, on March 2nd, 1993', Vol. 9 of *Adv. Numer. Methods Large Sparse Sets Linear Equations*, pp. 32–47.

Aubin, Jean-Pierre (1972), *Approximation of Elliptic Boundary-Value Problems*, Wiley-Interscience, New York.

Baglama, James, Daniela Calvetti, Gene H. Golub & Lothar Reichel (1998), 'Adaptively preconditioned GMRES algorithms', *SIAM J. Sci. Comput.* **20**(1), 243–269.

Björck, Åke & Gene H. Golub (1973), 'Numerical methods for computing angles between linear subspaces', *Math. Comp.* **27**, 579–594.

Brenner, Susanne C. & L. Ridgway Scott (1994), *The Mathematical Theory of Finite Element Methods*, number 15 *in* 'Texts in Applied Mathematics', Springer-Verlag, New York.

Brezzi, Franco & Michel Fortin (1991), *Mixed and hybrid finite element methods*, number 15 *in* 'Springer series in computational mathematics', Springer-Verlag.

Brown, P. N. (1991), 'A theoretical comparison of the Arnoldi and GMRES algorithms', *SIAM J. Sci. Stat. Comput.* **12**, 58–77.

Chan, Tony F., Estratios Gallopoulos, Valeria Simoncini, Tedd Szeto & C. H. Tong (1994), 'A quasi-minimal residual variant of the Bi-CGSTAB algorithm for nonsymmetric systems', *SIAM J. Sci. Comput.* **15**, 338–347.

Chapman, Andrew & Yousef Saad (1997), 'Deflated and augmented Krylov subspace techniques', *Numer. Linear Algebra Appl.* **4**(1), 43–66.

Chatelin, Francoise (1993), *Eigenvalues of Matrices*, John Wiley & Sons.

Concus, Paul & Gene H. Golub (1976), A generalized conjugate gradient method for nonsymmetric systems of linear equations, *in* R.Glowinski & J. L.Lions, eds, 'Computing Methods in Applied Sciences and Engineering', Springer-Verlag, New York, pp. 56–65.

Concus, Paul, Gene H. Golub & Dianne P. O'Leary (1976), A generalized conjugate gradient method for solution of elliptic partial differential equations, *in* J. R.Bunch & D. J.Rose, eds, 'Proc. Symp. on Sparse Matrix Computations', Argonne National Lab.(Sept. 1975), Academic Press, New York.

Cullum, Jane & Anne Greenbaum (1996), 'Relations between Galerkin and norm-minimizing iterative methods for solving linear systems.', *SIAM J. Matrix Anal. Appl.* **17**(2), 223–247.

Davis, Chandler (1963), 'The rotation of eigenvectors by a perturbation', *J. Math. Anal. Appl.* **6**, 159–173.

Davis, Chandler (1965), 'The rotation of eigenvectors by a perturbation. II', *J. Math. Anal. Appl.* **11**, 20–27.

Davis, Chandler & William M. Kahan (1970), 'The rotation of eigenvectors by a perturbation. III', *SIAM J. Numer. Anal.* **7**(1), 1–46.

de Sturler, Eric (1996), 'Nested Krylov methods based on GCR', *J. Comp. Appl. Math.* **67**, 15–41.

de Sturler, Eric (1999), 'Truncation strategies for optimal Krylov subspace methods', *SIAM J. Numer. Anal.* **36**(3).

Driscoll, Tobin A., Kim-Chung Toh & Lloyd N. Trefethen (1998), 'From potential theory to matrix iterations in six steps', *SIAM Rev.* **40**(3), 547–578.

Eiermann, Michael (1989), 'Semiiterative Verfahren für nichtsymmetrische lineare Gleichungssysteme', Habilitationsschrift, Universität Karlsruhe.

Eiermann, Michael (1993), 'Fields of values and iterative methods', *Linear Algebra Appl.* **180**, 167–197.

Eiermann, Michael (1996), 'Field of values and iterative methods'. in preparation.

Eiermann, Michael & Oliver G. Ernst (1998), 'Geometric aspects in the theory of Krylov subspace methods'. to appear in Acta Numerica 2001.

Eiermann, Michael, Oliver G. Ernst & Olaf Schneider (1999), 'Analysis of acceleration strategies for restarted minimal residual methods'. to appear in J. Comp. Appl. Math.

Eiermann, Michael, Wilhelm Niethammer & Richard S. Varga (1985), 'A study of semiiterative methods for nonsymmetric systems of linear equations', *Numer. Math.* **47**, 505–533.

Eisenstat, Stanley C., Howard C. Elman & Martin H. Schultz (1983), 'Variational iterative methods for nonsymmetric systems of linear equations', *SIAM J. Sci. Comput.* **20**, 345–357.

Elman, Howard C. (1982), Iterative Methods for Sparse, Nonsymmetric Systems of Linear Equations, PhD thesis, Yale University, Department of Computer Science.

Elman, Howard C., Oliver G. Ernst & Dianne P. O'Leary (1999), A multigrid method enhanced by Krylov subspace iteration for discrete Helmholtz equations, Technical Report 99-36, University of Maryland Institute for Advanced Computer Studies.

Embree, Mark (1999), How Descriptive are GMRES Convergence bounds, Technical Report 99/08, Oxford University Computing Laboratory.
URL: *http://web.comlab.ox.ac.uk/oucl/work/mark.embree/*

Erhel, Jocelyne, Kevin Burrage & Bert Pohl (1996), 'Restarted GMRES preconditioned by deflation', *J. Comp. Appl. Math.* **69**, 303–318.

Ernst, Oliver G. (2000), 'Krylov subspace methods for stabilized discretizations of the convection-diffusion equation', to appear in SIAM J. Matrix Anal. Appl.

Faber, Georg (1903), 'Über polynomische Entwickelungen', *Math. Ann.* **57**, 389–408.

Faber, Georg (1920), 'Über Tschebyscheffsche Polynome', *J. Reine Angew. Math.* **150**, 79–106.

Faber, Vance & Thomas A. Manteuffel (1984), 'Necessary and sufficient conditions for the existence of a conjugate gradient method', *SIAM J. Numer. Anal.* **21**, 352–362.

Faber, Vance & Thomas A. Manteuffel (1987), 'Orthogonal error methods', *SIAM J. Numer. Anal.* **24**, 170–187.

Faber, Vance, Wayne Joubert, Emanuel Knill & Thomas A. Manteuffel (1996), 'Minimal residual method stronger than polynomial preconditioning', *SIAM J. Matrix Anal. Appl.* **17**(4), 707–729.

Fischer, Bernd (1996), *Polynomial Based Iteration Methods for Symmetric Linear Systems*, Wiley-Teubner, Leipzig.

Freund, Roland W. (1992*a*), Quasi-kernel polynomials and convergence results for quasi-minimal residual iterations, *in* D.Braess & L. L.Schumaker, eds, 'Numerical Methods of Approximation Theory', Birkhäuser, Basel.

Freund, Roland W. (1992*b*), 'Quasi-kernel polynomials and their use in non-Hermitian matrix iterations', *J. Comput. Appl. Math.* **43**(1–2), 135–158.

Freund, Roland W. (1993), 'A transpose-free quasi-minimal residual algorithm for non-Hermitian linear systems', *SIAM J. Sci. Comput.* **14**(2), 470–482.

Freund, Roland W. (1994), Transpose-free quasi-minimal residual methods for non-Hermitian linear systems, *in* G.Golub, A.Greenbaum & M.Luskin, eds, 'Recent Advances in Iterative Methods', Vol. 60 of *Mathematics and its Applications*, IMA, Springer-Verlag, New York, pp. 69–94.

Freund, Roland W., M.H Gutknecht & Noel M. Nachtigal (1993), 'An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices', *SIAM J. Sci. Comput.* **14**(1).

Freund, Roland W. & Noel M. Nachtigal (1991), 'QMR: a quasi-minimal residual method for non-Hermitian linear systems', *Numer. Math.* **60**, 315–339.

Golub, Gene H. & Charles F. van Loan (1989), *Matrix Computations*, Johns Hopkins Univ. Press.

Golub, Gene H. & Charles F. van Loan (1996), *Matrix Computations*, third edn, Johns Hopkins University Press, Baltimore, MD.

Greenbaum, Anne (1997), *Iterative Methods for Solving Linear Systems*, Vol. 17 of *Frontiers in Applied Mathematics*, SIAM, Philadelphia, PA.

Greenbaum, Anne & L. Gurvits (1994), 'Max-min properties of matrix factor norms', *SIAM J. Sci. Comput.* **15**, 348–358.

Greenbaum, Anne & Lloyd N. Trefethen (1994), 'GMRES/CR and Arnoldi/Lanczos as matrix approximation problems', *SIAM J. Sci. Comput.* **15**, 359–368.

Greenbaum, Anne, Miroslav Rozložník & Zděnek Strakoš (1997), 'Numerical behaviour of the modified Gram-Schmidt GMRES implementation', *BIT* **37**(3), 706–719.

Greenbaum, Anne & Zdeněk Strakoš (1994), Matrices that generate the same Krylov residual spaces, *in* G.Golub, A.Greenbaum & M.Luskin, eds, 'Recent Advances in Iterative Methods', Springer, New York, pp. 95–118.

Greenbaum, Anne, Zdeněk Strakoš & Vlastimil Ptak (1996), 'Any nonincreasing convergence curve is possible for GMRES', *SIAM J. Matrix Anal. Appl.* **17**(3), 465–469.

Greengard, L. & V. Rokhlin (1987), 'A fast algorithm for particle simulations', *J. Comput. Phys.* **73**(2), 325–348.

Greville, T. N. E. (1974), 'Solutions of the matrix equation $XAX = X$ and relations between oblique and orthogonal projectors', *SIAM J. Appl. Math.* **26**(4).

Gustafson, Karl E. & Duggirala K. M. Rao (1996), *Numerical Range: the Field of Values of Operators and Matrices*, Springer-Verlag, New York.

Gutknecht, Martin H. (1992), 'A completed theory of the unsymmetric Lanczos process and related algorithms. I', *SIAM J. Matrix Anal. Appl.* **13**, 594–639.

Gutknecht, Martin H. (1993*a*), 'Changing the norm in conjugate gradient type algorithms', *SIAM J. Numer. Anal.* **30**, 40–56.

Gutknecht, Martin H. (1993*b*), 'Variants of BiCGStab for matrices with complex spectrum', *SIAM J. Sci. Stat. Comput.* **14**(5), 1020–1033.

Gutknecht, Martin H. (1994), 'A completed theory of the unsymmetric Lanczos process and related algorithms. II', *SIAM J. Matrix Anal. Appl.* **15**, 15–58.

Gutknecht, Martin H. (1997), 'Lanczos-type solvers for nonsymmetric linears systems of equations', *Acta Numerica* **6**, 271–397.

Hackbusch, W. & Z.P. Nowak (1989), 'On the fast matrix multiplication in the boundary element method by panel clustering', *Numer. Math.* **54**, 463–491.

Hackbush, Wolfgang (1985), *Multi-Grid Methods and Applications*, Springer-Verlag, Berlin.

Hestenes, Magnus R. & Eduard Stiefel (1952), 'Methods of conjugate gradients for solving linear systems', *J. Res. Nat. Bur. Standards* **49**, 409–436.

Hochbruck, Marlis & Christian Lubich (1998), 'Error analysis of Krylov methods in a nutshell', *SIAM J. Sci. Comput.* **19**, 695–701.

Horn, Roger A. & Charles R Johnson (1985), *Matrix Analysis*, Cambridge University Press, Cambridge, UK.

Horn, Roger A. & Charles R. Johnson (1991), *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK.

Householder, Alston S. (1964), *The Theory of Matrices in Numerical Analysis*, Dover, New York.

Huhtanen, Marko & Olavi Nevanlinna (1998), Minimal decompositions and iterative methods, Technical Report Preprint 3, Centro Int. Math., Portugal. to appear in Numer. Math.

Jordan, Camille (1875), 'Essai sur la géométrie à n dimensions', *Bull. Soc. Math. France* **3**, 103–174.

Joubert, Wayne D. (1994*a*), 'On the convergence behavior of the restarted GMRES algorithm for solving nonsymmetric linear systems', *Num. Lin. Alg. Appl.* **1**, 427–447.

Joubert, Wayne D. (1994*b*), 'A robust GMRES-based adaptive polynomial preconditioning algorithm for nonsymmetric linear systems', *SIAM J. Sci. Comput.* **15**, 427–439.

Joubert, Wayne D. & David M. Young (1987), 'Necessary and sufficient conditions for the existence of generalized conjugate-gradient iterations', *Lin. Alg. Appl.* **88/89**, 449–485.

Kato, Tosio (1960), 'Estimation of iterated matrices, with application to the von Neumann condition', *Numer. Math.* **2**, 22–29.

Kato, Tosio (1965), 'Some mapping theorems for the numerical range', *Proc. Japan Acad.* **41**, 652–655.

Kato, Tosio (1980), *Perturbation Theory for Linear Operators*, Classics in Mathematics, Springer-Verlag, Berlin. Originally published as Vol. 132 of the series *Grundlehren der mathematischen Wissenschaften*.

Kharchenko, Serge A. & Alex Yu. Yeremin (1995), 'Eigenvalue translation based preconditioners for the GMRES(k) method', *Numer. Lin. Alg. Appl,* **2**(1), 51–77.

Kress, Rainer (1989), *Linear Integral Equations*, number 82 *in* 'Applied Mathematical Sciences', Springer-Verlag, Berlin.

Krylov, Aleksei N. (1931), 'On the numerical solution of the equation by which the frequency of small oscillations is determined in technical problems', *Isz. Akad. Nauk SSSR Ser. Fiz.-Math.* **4**, 491–539.

Lanczos, Cornelius (1950), 'An iteration method for the solution of the eigenvalue problem of linear differential and integral operators', *J. Res. Nat. Bur. Standards* **45**, 255–282.

Lanczos, Cornelius (1952), 'Solution of systems of linear equations by minimized iterations', *J. Res. Nat. Bur. Standards* **49**, 33–53.

Liesen, Jörg (1998), Contruction and Analysis of Polynomial Iterative Methods for Non-Hermitian Systems of Linear Equations, PhD thesis, University of Bielefeld.

Manteuffel, Thomas A. (1977), 'The Tchebychev iteration for nonsymmetric linear systems', *Numer. Math.* **28**, 307–327.

Manteuffel, Thomas A. & James S. Otto (1994), 'On the roots of the orthogonal polynomials and residual polynomials associated with a conjugate gradient method', *Num. Lin. Alg. Appl.* **1**(5), 449–475.

Moret, Igor (1997), 'A note on the superlinear convergence of GMRES', *SIAM J. Numer. Anal.* **34**(2), 513–516.

Morgan, Ronald B. (1991), 'Computing interior eigenvalues of large matrices', *Lin. Alg. Appl.* **154–156**, 289–309.

Morgan, Ronald B. (1995), 'A restarted GMRES method augmented with eigenvectors', *SIAM J. Matrix Anal. Appl.* **16**(4), 1154–1171.

Morgan, Ronald B. (1997), 'Implicitly restarted GMRES and Arnoldi methods for nonsymmetric systems of equations', *SIAM J. Matrix Anal. Appl.* . to appear.
URL: *http://www.Baylor.edu/ Ronald_Morgan/reports.html*

Morgan, Ronald B. (1999), 'GMRES with deflated restarting'. manuscript.
URL: *http://www.Baylor.edu/ Ronald_Morgan/reports.html*

Nachtigal, Noel M., Lothar Reichel & Lloyd N. Trefethen (1992), 'A hybrid GMRES algorithm for nonsymmetric linear systems', *SIAM J. Matrix Anal. Appl.* **13**, 796–825.

Nevanlinna, Olavi (1993), *Convergence of Iterations for Linear Equations*, Birkhäuser, Basel.

Paige, Christopher C., Beresford N. Parlett & Henk A. van der Vorst (1995), 'Approximate solutions and eigenvalue bounds from Krylov subspaces', *Num. Lin. Alg. Appl.* **2**(2), 115–133.

Paige, Christopher C. & M. Wei (1994), 'History and generality of the CS decomposition', *Lin. Alg. Appl.* **208/209**, 303–326.

Paige, Christopher C. & Michael A. Saunders (1975), 'Solution of sparse indefinite systems of linear equations', *SIAM J. Numer. Anal.* **12**, 617–629.

Paige, Christpher C. & Michael A. Saunders (1982), 'LSQR: an algorithm for sparse linear systems and sparse least squares', *ACM Trans. Math. Soft.* **8**, 43–71.

Reid, John K. (1971), Onthe method of conjugate gradients for the solution of large sparse systems of linear equations, *in* 'Large Sparse Sets of Linear Equations', Academic Press, New York, pp. 231–254.

Ressel, Klaus J. & Martin H. Gutknecht (1996), QMR-smoothing for Lanczos-type product methods based on three-term recurrences, Technical Report TR-96-18, Swiss Center for Scientific Computing, Zuerich.

Riesz, Frigyes & Béla Sz.-Nagy (1955), *Functional Analysis*, Frederick Ungar, New York. (engl. transl. of Leçons d'Analyse Fonctionelle).

Ringrose, John R. (1971), *Compact non-self-adjoint operators*, Van Nostrand Reinhold Company, London.

Rozložník, Miroslav & Zděnek Strakoš (1996), Variants of the residual minimizing Krylov space methods, *in* I.Marek, ed., 'Proceedings of the XI. School 'Software and Algorithms of Numerical Mathematics', Zelezna Ruda, University of West Bohemia', pp. 208–225.

Rutishauser, Heinz (1959), Theory of gradient methods, *in* 'Refined Iterative Methods for Computation of the Solution and the Eigenvalues of Self-Adjoint Boundary Value Problems', number 8 *in* 'Mitteilungen aus dem Institut für angewandte Mathematik', Birkhäuser Verkag, Basel/Stuttgart, pp. 24–49.

Saad, Youcef & Martin H. Schultz (1985), 'Conjugate gradient-like algorithms for solving nonsymmetric linear systems', *Math. Comp.* **44**, 417–424.

Saad, Yousef (1981), 'Krylov subspace methods for solving large unsymmetric linear systems', *Math. Comp.* **37**, 105–126.

Saad, Yousef (1982), 'The Lanczos biorthogonalization algorithm and other oblique projection methods for solving large unsymmetric systems', *SIAM J. Numer. Anal.* **19**, 470–484.

Saad, Yousef (1993), 'A flexible inner-outer preconditioned GMRES algorithm', *SIAM J. Sci. Comput.* **14**, 461–469.

Saad, Yousef (1996), *Iterative Methods for Sparse Linear Systems*, PWS Publishing, Boston.

Saad, Yousef (1997*a*), 'Analysis of augmented Krylov subspace methods', *SIAM J. Matrix Anal. Appl.* **18**(2), 435–449.

Saad, Yousef (1997*b*), Further analysis of minimum residual iterations, Technical Report UMSI-97-14, University of Minnesota Supercomputer Institute.

Saad, Yousef & Martin H. Schultz (1986), 'GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems', *SIAM J. Sci. Comput.* **7**, 856–869.

Schönauer, Willi (1987), *Scientific Computing on Vector Computers*, Elsevier, Amsterdam.

Shortley, George H. (1953), 'Use of Tschebyscheff-polynomial operators in the numerical solution of boundary-value problems', *J. Appl. Phys.* **24**, 392–396.

Simoncini, Valeria (1999), 'A new variant of restarted GMRES', *Num. Lin. Alg. Appl.* **6**, 61–77.

Sonneveld, Peter (1989), 'CGS, a fast Lanczos-type solver for nonsymmetric linear systems.', *SIAM J. Sci. Stat. Comput.* **10**(1), 36–52.

Sorensen, Danny C. (1992), 'Implicit application of polynomial filters in a k-step Arnoldi method', *SIAM J. Matrix Anal. Appl.* **13**, 357–385.

Starke, Gerhard (1994), 'Iterative methods and decomposition-based preconditioners for nonsymmetric elliptic boundary value problems', Habilitationsschrift, Universität Karlsruhe.

Starke, Gerhard & Richard S. Varga (1993), 'A hybrid Arnoldi-Faber method for nonsymmetric systems of linear equations', *Numer. Math.* **64**, 312–240.

Stewart, G. W. (1977), 'On the perturbation of pseudo-inverses, projections, and linear least squares problems', *SIAM Rev.* **19**, 643–662.

Stewart, G. W. (1998), *Matrix Algorithms, Vol. I Basic Decompositions*, SIAM, Philadelphia, PA.

Stiefel, Eduard L. (1955), 'Relaxationsmethoden bester Strategie zur Lösung linearer Gleichungssysteme', *Comm. Math. Helv.* **29**, 157–179.

Stiefel, Eduard L. (1958), 'Kernel polynomials in linear algebra and their numerical applications', *J. Res. Nat. Bur. Standards, Appl. Math. Ser.* **49**, 1–22.

Strikwerda, John C. & Suzan C. Stodder (1995), Convergence results for GMRES(m), Technical Report 1280, Dept. Comp. Science, Univ. of Wisconsin.

Toh, Kim-Chuan (1997), 'GMRES vs. ideal GMRES', *SIAM J. Matrix Anal. Appl.* **18**, 30–36.

Toh, Kim-Chuan & Lloyd N. Trefethen (1998), 'The Chebyshev polynomials of a matrix', *SIAM J. Matrix Anal. Appl.* **20**, 400–419.

Trefethen, Lloyd N. (1990), Approximation theory and numerical linear algebra, *in* J.Mason & M.Cox, eds, 'Algorithms for Approximation II', Chapman and Hall, London, pp. 336–360.

Trefethen, Lloyd N. (1992), Pseudospectra of matrices, *in* D. F.Griffiths & G. A.Watson, eds, 'Numerical Analysis 1991', Longman 1992.

Trefethen, Lloyd N. (1999), 'Computation of pseudospectra', *Acta Numerica* pp. 247–295.

van der Vorst, Henk A. (1992), 'BICGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems', *SIAM J. Sci. Stat. Comput.* **13**, 631–644.

van der Vorst, Henk A. & C. Vuik (1994), 'GMRESR: a family of nested GMRES methods', *Num. Lin. Alg. Appl.* **1**, 369–386.

Varga, Richard S. (1999), *Matrix Iterative Analysis*, second edn, Springer-Verlag, Berlin/Heidelberg.

Vinsome, P. K. W. (1976), Orthomin: an iterative method for solving sparse sets of simultaneous linear equations, *in* 'Proc. Fourth Symposium of Numerical Simulation of Reservoir Performance of the Society of Petroleum Engineers of the AIME, Los Angeles, 19–20 February'. SPE Paper 5729.

Vuik, C. (1993), 'Further experiences with GMRESR', *Supercomputer* **55**, 13–27.

Walker, Homer F. (1995), 'Residual smoothing and peak/plateau behavior in Krylov subspace methods', *Appl. Numer. Math.* **19**, 279–286.

Walker, Homer & Lu Zhou (1994), 'A simpler GMRES', *Numer. Linear Algebra Appl.* **1**(6), 571–581.

Wedin, Per Åke (1983), On angles between subspaces of a finite dimensional inner product space, *in* B.Kågström & A.Ruhe, eds, 'Matrix Pencils', number 273 *in* 'Lecture Notes in Mathematics', Springer-Verlag, New York, pp. 263–285.

Weiss, Rüdiger (1990), Convergence behavior of generalized conjugate gradient methods, PhD thesis, Universität Karlsruhe, Karlsruhe, Germany.

Weiss, Rüdiger (1994), 'Properties of generalized conjugate gradient methods', *Numer. Linear Algebra Appl.* **1**(1), 45–63.

Widlund, Olof (1978), 'A Lanczos method for a class of nonsymmetric systems of linear equations', *SIAM J. Numer. Anal.* **15**, 801–812.

Wielandt, Helmut (1996), Topics in the analytic theory of matrices, *in* B.Huppert & H.Schneider, eds, 'Helmut Wielandt, Mathematical Works', Vol. 2, Walter de Gruyter.

Winther, Ragnar (1980), 'Some superlinear convergence results for the conjugate gradient method', *SIAM J. Numer. Anal.* **17**, 14–17.

Xu, Jinchao (1992), 'A new class of iterative methods for nonselfadjoint or indefinite problems', *SIAM J. Numer. Anal.* **29**(2), 303–319.

Young, David M. & Kang C. Jea (1980), 'Generalized conjugate gradient acceleration of nonsymmetrizable iterative methods', *Lin. Alg. Appl.* **34**, 159–194.

Zhou, Lu & Homer F. Walker (1994), 'Residual smoothing techniques for iterative methods', *SIAM J. Sci. Comput.* **15**, 297–312.