Krylov Subspace Methods and their Generalizations for Solving Singular Linear Operator Equations with Applications to Continuous Time Markov Chains

Von der Fakultät für Mathematik und Informatik
der Technischen Universität Bergakademie Freiberg
genehmigte

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.),

vorgelegt

von Dipl.-Math. Olaf Schneider

geboren am 23. 03. 1972 in Weißwasser

Gutachter: Prof. Dr. Michael Eiermann (Freiberg)
Prof. Dr. Marlis Hochbruck (Düsseldorf)
Prof. Dr. Ivo Marek (Prag)

Tag der Verleihung: 9. Oktober 2006

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die zum Gelingen der Arbeit beigetragen haben.

Am Graduiertenkolleg „Räumliche Statistik“ der TU Bergakademie Freiberg hatte ich für drei Jahre die Möglichkeit, meinen Forschungen und Neigungen nachzugehen. Zudem bot sich vielfach Gelegenheit, Neues zu lernen und über den sprichwörtlichen Tellerrand zu schauen. Mein Dank gilt allen Kollegiaten, Professoren und Mitarbeitern für die freundliche Aufnahme und Unterstützung.

Ein weiterer Angelpunkt meiner Freiberger Jahre war das Institut für Angewandte Mathematik II. Ich danke den dortigen Kollegen für die Zusammenarbeit, insbesondere Oliver Ernst, Werner Queck und Volker Sperling.

Am Institut für Wissenschaftliches Rechnen des Forschungszentrums Karlsruhe fand ich schließlich ein neues Wirkungsfeld. Zudem durfte ich Zeit und Ressourcen im Institut für den Abschluss der Arbeit nutzen. Mein Dank dafür gilt dem Institutsleiter, Herrn Mickel, sowie Frank Schmitz und den anderen Kollegen für ihr Verständnis.

Besonderer Dank gebührt Prof. Eiermann, der mich schon während des Studiums für die numerische Mathematik begeisterte und seither – als Betreuer meiner Diplomarbeit und der Dissertation – meine Arbeit begleitet hat. Nicht zuletzt musste er meine unstete Arbeitsweise ertragen und hat mir trotz einer langen Unterbrechnung die Möglichkeit gegeben, die vorliegende Arbeit zu beenden.

Danken möchte ich außerdem Herrn Marcos Köhrmann für Beistand in einer schwierigen Lebensphase.

Contents

1 Introduction

Krylov subspace methods are a standard technique for solving linear equations. The great majority of existing literature is devoted to the study of these methods for a regular coefficient matrix (cf. [22] and the references therein). The analysis of this methods applied to a singular operator is not as well developed as for the nonsingular case. On the other hand, there are applications where linear systems of equations with a singular matrix arise, for example the discretization of partial differential equations with Neumann boundary conditions or discrete inverse problems (cf. [33]). Another important realm of applications is the study of Markov chains with large finite state space, which is also stressed in this thesis.

There are studies of the behavior of widespread used Krylov methods for singular systems, such as [34] for the Conjugate Residual (CR) method, [35] for the Generalized Conjugate Residual (GCR) method, [7, 64] for the Generalized Minimal Residual (GMRES) method, and [29] for the Quasi-Minimal Residual (QMR) method. Other authors have proposed modified versions of established Krylov methods (cf. [61, 62, 63]). Though, the analysis is often short-coming and lacks a consistent presentation of results. Recent advances in unifying the theory of Krylov methods use abstract orthogonal residual (OR) and minimal residual (MR) approximations to describe all kinds of subspace correction methods in a concise manner (cf. [22, 25]). These advances are not yet adopted for the singular case. The main objective of the present work is to close this open issues.

In studying singular systems of linear equations there naturally arise generalizations of the concept of an inverse operator. We shall widely use such generalized inverses. Therefore we start with a short survey about this topic in Section 2. A less common representative of the class of generalized inverses is the subspace inverse, which we introduce in Section 3.2. This section is part of the exposition of abstract MR and OR approximation methods.

Our presentation of the material in Section 3 is greatly inspired by Eiermann and Ernst (cf. [22]). We solely augment their results to cover the singular operator case. To this end, after some introductory remarks about operator equations and subspace correction methods in Section 3.1, we study minimal residual approximations for the solution of operator equations in Section 3.2. The standard implementation of this methods using nested orthonormal bases is described in Section 3.3. The breakdown behavior of the underlying Gram-Schmidt orthonormalization process is studied in in Section 3.4. We show in Section 3.5, how most of the results in [22] about MR and OR methods for solving regular systems may be preserved in the context of a singular operator.

The Drazin inverse, another generalized inverse, plays a prominent role in Section 4 where we specialize the abstract setting to Krylov subspace methods. After the basic definition we point out the polynomial structure of Krylov spaces in Section 4.1. Section 4.2 provides tools to deal with the Drazin inverse and the minimal polynomials of a matrix and a vector. They are used in Section 4.3, where the important connection between Krylov subspaces and the Drazin inverse is established. In Section 4.4 we investigate the question, when a Krylov method yields least squares solution. This leads to new identities involving the Drazin and the Moore-Penrose inverse. We close with some remarks about the stability of Krylov methods and the modifications necessary to obtain Drazin inverse solutions.

As a real world problem involving the solution of singular linear operator equations we present the analysis of continuous-time Markov chains (CTMC) in Section 5. After some preliminaries about the background in stochastics and tools from matrix theory we investigate the transition function of a Markov chain in a purely analytical and matrix algebraic context. The results of this analysis are then interpreted in terms of the stochastic process and used to explain the usefulness of Krylov subspace methods for the analysis of Markov chains.

We conclude with some final remarks on possible topics of further research.

2 Generalized Inverses

In this section we summarize some facts about generalizations of the inverse of a linear operator. This is not only done for introducing the reader into this subject, but also to emphasize common properties and differences between the proper inverse and its various generalizations. To keep things simple we restrict ourselves to the case, when $A \in \mathbb{C}^{m\times n}$ is a matrix. Considerable introductions to this topic can be found in the monographs of Campbell and Meyer [14] or Ben-Israel and Greville [4].

The inverse $X = A^{-1}$ of a regular (and thus square) matrix A fulfills the matrix identities

$$AX = I \qquad \text{and} \tag{2.1}$$
$$XA = I, \tag{2.2}$$

where I is the identity matrix. Both equations characterize the inverse of A. A simple consequence is that a regular matrix commutes with its inverse, i. e.,

$$AX = XA. \tag{2.3}$$

Other obvious identities are

$$AXA = A, \tag{2.4}$$
$$XAX = X, \tag{2.5}$$
$$(AX)^* = AX, \tag{2.6}$$
$$(XA)^* = XA \qquad \text{and} \tag{2.7}$$
$$A^{k+1}X = A^k \tag{2.8}$$

for every $k \geq 0$. Here M^* denotes the adjoint of a matrix (or operator) M with respect to given inner products $(\cdot,\cdot)$ on $\mathcal{H}_1 = \mathbb{C}^n$ and $\mathcal{H}_2 = \mathbb{C}^m$. An obvious way to generalize the term "inverse" is to require a subset of the above identities to be satisfied.

Equations (2.4) to (2.7) are known as the *Penrose conditions*. Together, they define the *Moore-Penrose generalized inverse* or *pseudoinverse* $A^\dagger$ of an arbitrary matrix $A \in \mathbb{C}^{m\times n}$. A generalization of $A^\dagger$ for a linear operator $A : \mathcal{H}_1 \to \mathcal{H}_2$ on arbitrary Hilbert spaces $\mathcal{H}_j$ can be found in [32] (see also [4, Chapter 8]).

Note also, that each of the Penrose conditions (2.4) and (2.5) immediately implies, that AX and XA are projections. This may be regarded as a generalization of equations (2.1) and (2.2) where the identity matrix I is replaced by some other operator which

is the identity on a certain subspace and null elsewhere. Further, the Penrose conditions (2.6) and (2.7) yield that these projections are orthogonal ones. In fact, the Moore-Penrose inverse can be characterized as the uniquely determined operator X for which $AX = P_{\mathcal{R}(A)}$ and $XA = P_{\mathcal{R}(A^*)}$. Here $P_{\mathcal{U}}$ denotes the orthogonal projection onto the subspace $\mathcal{U}$ and $\mathcal{R}(M) := \{M\boldsymbol{v} : \boldsymbol{v} \in \mathcal{H}_1\} \subseteq \mathcal{H}_2$ is the range of the operator (or matrix) $M : \mathcal{H}_1 \to \mathcal{H}_2$.

If it exists, the solution of equations (2.3), (2.4) and (2.5) is called the *group inverse* $A^{\#}$ of the (necessarily square) matrix A. It is called so, because it is the inverse element of A in any multiplicative group of matrices.

Another well known generalized inverse is the *Drazin inverse* of a square matrix A, which is the solution of equations (2.3), (2.5) and (2.8) for $k \geq \text{index}(A)$. The index of a matrix A is related to the Jordan canonical form: It is simply the dimension of the largest Jordan block corresponding to the eigenvalue 0 (and $\text{index}(A) = 0$ if A is nonsingular). Note that the group inverse is just the Drazin inverse of a matrix A with $\text{index}(A) \leq 1$.

Beside its nice algebraic properties, the inverse of a linear operator is a powerful tool in solving operator equations. Consider

$$A\boldsymbol{x} = \boldsymbol{b}, \tag{2.9}$$

where we are looking for solutions $\boldsymbol{x} \in \mathbb{C}^n = \mathcal{H}_1$ for a given right hand side $\boldsymbol{b} \in \mathbb{C}^m = \mathcal{H}_2$. If A is square and regular, the unique solution of this equation is given by $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. It seems natural to generalize this term for singular or even rectangular A. Such a generalization has to overcome two problems: First, (2.9) needs not to be *consistent*, i. e., $\boldsymbol{b} \notin \mathcal{R}(A)$, which means that the equation is not solvable in the usual sense. Second, even if the system is solvable, the solution need not be unique. More precisely, the solution of a consistent linear equation is unique if and only if $\mathcal{N}(A) = \{\boldsymbol{v} : A\boldsymbol{v} = \boldsymbol{0}, \boldsymbol{v} \in \mathcal{H}_1\}$ is the trivial space which contains only the zero vector.

A generalized inverse which supplies a solution of (2.9) if the equation is consistent is called an *equation solving generalized inverse*. The class of all equation solving inverses of A is usually denoted by $A\{1\}$ and A^{-} denotes an arbitrary element of it. $A\{1\}$ is characterized by equation (2.4). If $X \in A\{1\}$ additionally satisfies (2.6) it is called a *least squares generalized inverse*, since for every right hand side $\boldsymbol{b}$ there holds

$$\|\boldsymbol{b} - AX\boldsymbol{b}\| = \min_{\boldsymbol{x} \in \mathcal{H}_1} \|\boldsymbol{b} - A\boldsymbol{x}\|$$

in this case. The norm $\|\cdot\|$ is induced by the inner product $(\cdot, \cdot)$. Similarly, the set of all equation solving inverses, which satisfies

$$\|X\boldsymbol{b}\| = \min_{\boldsymbol{x}, A\boldsymbol{x} = \boldsymbol{b}} \|\boldsymbol{x}\| \quad \text{for all} \quad \boldsymbol{b} \in \mathcal{R}(A)$$

is characterized by equations (2.4) and (2.7) and any such X is called a *minimal norm generalized inverse* of A. A least squares generalized inverse is said to have the *minimal*

norm property if $\|X\boldsymbol{b}\| < \|\boldsymbol{w}\|$ for all least squares solutions $\boldsymbol{w}$ of (2.9). The Moore-Penrose inverse $A^\dagger$ is precisely the only least square generalized inverse with minimal norm property (cf. [14, Theorem 2.1.1]).

All generalized inverses introduced so far coincide with A^{-1} if A is regular. But there are useful generalizations which lack this property. An example is the *constrained* or *subspace generalized inverse*, which satisfies the least squares and the minimal norm condition above only on a certain subspace $\mathcal{C} \subset \mathbb{C}^n$ (for details see Section 3.2 and [14, Section 3.6]). Note that the Drazin inverse provides a solution of (2.9) if the right hand side lies in a certain subspace of $\mathcal{R}(A)$ namely the range of $A^{\mathrm{index}(A)}$. In general, each of the above properties of A^{-1} may be restricted to a certain subset of vectors to get a new class of generalized inverses. Again, the subspace inverse may illustrate this. In Section 3.2 we will introduce it as the Moore-Penrose inverse of a certain operator. Alternatively it could be defined as the solution of the following equations:

$$
\begin{aligned}
AXA\boldsymbol{v} &= A\boldsymbol{v} && \text{for all } \boldsymbol{v} \in \mathcal{C} \subseteq \mathcal{H}_1, && (2.4')\\
XAX &= X, && && (2.5\ )\\
(AX\boldsymbol{y}, \boldsymbol{z}) &= (\boldsymbol{y}, AX\boldsymbol{z}) && \text{for all } \boldsymbol{y}, \boldsymbol{z} \in \mathcal{H}_2, && (2.6')\\
(XA\boldsymbol{v}, \boldsymbol{w}) &= (\boldsymbol{v}, XA\boldsymbol{w}) && \text{for all } \boldsymbol{v}, \boldsymbol{w} \in \mathcal{C} \subseteq \mathcal{H}_1, && (2.7')\\
\mathcal{R}(X) &\subseteq \mathcal{C}. &&
\end{aligned}
$$

Generalized inverses of this type are first studied by Minimade and Nakamura [50], where the term *restricted pseudoinverse* is used. Ben-Israel and Greville also discuss *restricted generalized inverses* in [4, Section 2.8].

In fact, almost every method for solving the linear equation (2.9) approximately can be regarded as the action of a certain generalized inverse of A on the right hand side (or the initial residual). In an iterative solution method each step is associated with a specific generalized inverse and the whole iteration can be described by specifying this sequence of operators. We illustrate this in short for the classical stationary iteration methods: Given a splitting $A = M - N$ with M nonsingular, the induced stationary iteration is

$$\boldsymbol{x}_\ell = T\boldsymbol{x}_{\ell-1} + \boldsymbol{c}, \qquad \ell = 1, 2, \ldots$$

with $T = M^{-1}N$ and $c = M^{-1}\boldsymbol{b}$. Using $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$ this can be rewritten as

$$\boldsymbol{x}_\ell = \boldsymbol{x}_0 + \left(I + T + \cdots + T^{\ell-1}\right) M^{-1}\boldsymbol{r}_0.$$

Thus, the ℓth iterate is obtained from the initial guess by adding a suitable correction. This correction may be regarded as an approximate solution of the residual equation. In formulas, setting

$$A_\ell^{\mathrm{X}} := \left(I + T + \cdots + T^{\ell-1}\right) M^{-1} = M^{-1} \sum_{j=0}^{\ell-1} (NM^{-1})^j,$$

the solution of $A\boldsymbol{c} = \boldsymbol{r}_0$ is approximated by $A_\ell^{\mathrm{X}}\boldsymbol{r}_0$, where A_ℓ^{X} is a generalized inverse of A. For a nonsingular matrix A a reasonable iteration matrix T satisfies $\lim_{\ell\to\infty} T^\ell = O$. Since A_ℓ^{X} satisfies the identity $A_\ell^{\mathrm{X}}A = I - T^\ell$ we conclude that (2.1) holds asymptotically for A_ℓ^{X}. In other words, A_ℓ^{X} converges to the proper inverse of A. That an iteration is reasonable means of course, that the iterates converge to the solution of (2.9), which is for nonsingular A equivalent to $\rho(T) < 1$, that is, the spectral radius is less then 1. If A is singular and $\boldsymbol{b} \in \mathcal{R}(A)$, under certain conditions on T (namely that T is semiconvergent, cf. Definition 5.2.9), the stationary iteration converges to a solution of the linear system which can be expressed in terms of the Drazin inverse of $I - T$ (cf. [5, Section 7.6]). Eiermann, Marek and Niethammer derive in [24] similar results for general semiiterative methods.

3 Abstract MR and OR Approximations

In [22] Eiermann and Ernst suggested an abstract approach to iterative projection methods for solving linear equations with an invertible operator. It turns out that many results are independent of the linear operator and thus hold in the singular case too. Other identities hold in a slightly modified form.

Following Eiermann and Ernst, we consider abstract MR and OR approximation problems in a Hilbert space $\mathcal{H}$ with inner product $(\cdot,\cdot)$ and induced norm $\|\cdot\|$. Given a (finite dimensional) subspace $\mathcal{W} \subset \mathcal{H}$ and an element $\boldsymbol{r} \in \mathcal{H}$, the MR approximation $\boldsymbol{h}^{\mathrm{MR}}$ is defined as the best approximation to $\boldsymbol{r}$ from $\mathcal{W}$. Denoting by $\boldsymbol{r}^{\mathrm{MR}}$ the associated approximation error there holds

$$\boldsymbol{h}^{\mathrm{MR}} \in \mathcal{W}, \qquad \boldsymbol{r}^{\mathrm{MR}} := \boldsymbol{r} - \boldsymbol{h}^{\mathrm{MR}} \perp \mathcal{W}.$$

The OR approximation $\boldsymbol{h}^{\mathrm{OR}}$ is obtained by imposing

$$\boldsymbol{h}^{\mathrm{OR}} \in \mathcal{W}, \qquad \boldsymbol{r}^{\mathrm{OR}} := \boldsymbol{r} - \boldsymbol{h}^{\mathrm{OR}} \perp \mathcal{V},$$

where $\mathcal{V}$ is a suitable test space of the same dimension as $\mathcal{W}$. We shall see later in this section, how this approximations may be expressed in terms of orthogonal and oblique projections onto $\mathcal{W}$.

In what follows $A : \mathcal{H} \to \mathcal{H}$ is a bounded linear operator which is not necessarily invertible. All results and conclusions in this section remain true for the more general case where $A : \mathcal{H}_1 \to \mathcal{H}_2$ is an operator between different Hilbert spaces (e. g., a rectangular matrix in finite dimensions). We do without this distinction here both for ease of notation and since the most relevant application of this theory in Krylov subspace methods (cf. Section 4) is covered by our assumptions. We also assume that $\mathcal{H}$ is separable and that the associated scalar field is algebraically closed, or, yet easier, that the scalars are complex numbers.

Some further remarks about the notation: Let $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ a finite set of linear independent vectors in $\mathcal{H}$. For convenience we often identify the basis $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ with its representation as row vector $U_k = \begin{bmatrix}\boldsymbol{u}_1 & \ldots & \boldsymbol{u}_k\end{bmatrix}$ and abbreviate the subspace $\mathcal{U}_k = \operatorname{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ by $\operatorname{span}\{U_k\}$. Similarly, $U_k\boldsymbol{g}$ with $\boldsymbol{g} = \begin{bmatrix}\gamma_1 & \ldots & \gamma_k\end{bmatrix}^\top \in \mathbb{C}^k$ denotes the linear combination $\gamma_1\boldsymbol{u}_1 + \cdots + \gamma_k\boldsymbol{u}_k$ and the notation U_k^* is an abbreviation of the mapping $\mathcal{H} \to \mathbb{C}^k : \cdot \mapsto \big[(\cdot, \boldsymbol{u}_j)\big]_{j=1}^k$. For a linear operator A on $\mathcal{H}$ we denote the *range* by $\mathcal{R}(A) = \{A\boldsymbol{v} \ : \ \boldsymbol{v} \in \mathcal{H}\}$ and the *nullspace* by $\mathcal{N}(A) = \{\boldsymbol{v} \ : \ A\boldsymbol{v} = \boldsymbol{0},\ \boldsymbol{v} \in \mathcal{H}\}$. The adjoint A^* of A is the operator satisfying $(A\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, A^*\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{H}$.

3.1 Operator Equations and Approximation Problems

We consider the linear equation

$$A\boldsymbol{x} = \boldsymbol{b} \tag{3.1}$$

with given right hand side $\boldsymbol{b} \in \mathcal{H}$ and the unknown vector $\boldsymbol{x} \in \mathcal{H}$. Given an initial guess $\boldsymbol{x}_0$ with the corresponding initial residual $\boldsymbol{0} \neq \boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$ and a sequence of nested *correction spaces*

$$\{\boldsymbol{0}\} = \mathcal{C}_0 \subset \mathcal{C}_1 \subset \mathcal{C}_2 \subset \cdots \subset \mathcal{C}_m \subset \mathcal{C}_{m+1} \subset \cdots \subset \mathcal{H}, \quad \dim \mathcal{C}_m = m, \tag{3.2}$$

we investigate iterates of the form $\boldsymbol{x}_m = \boldsymbol{x}_0 + \boldsymbol{c}_m$, $\boldsymbol{c}_m \in \mathcal{C}_m$. That is, we correct the initial approximation of the solution of (3.1) by a *correction* $\boldsymbol{c}_m$ determined from $\mathcal{C}_m$ to obtain a (hopefully better) approximate solution $\boldsymbol{x}_m$. Methods of this kind are known as subspace correction methods and are widely used in applied mathematics.

Setting $\boldsymbol{h}_m := A\boldsymbol{c}_m$, the corresponding residual vector is

$$\boldsymbol{r}_m = \boldsymbol{b} - A\boldsymbol{x}_m = \boldsymbol{r}_0 - A\boldsymbol{c}_m = \boldsymbol{r}_0 - \boldsymbol{h}_m.$$

Thus, $\boldsymbol{h}_m$ can be regarded as a suitable approximation of $\boldsymbol{r}_0$ from the mth *approximation space*

$$\mathcal{W}_m := A\mathcal{C}_m. \tag{3.3}$$

The residual $\boldsymbol{r}_m$ is the approximation error of $\boldsymbol{h}_m$. This approximation problem does not depend on the operator equation (3.1).

A simple but often used result is provided by

Lemma 3.1.1. *If and only if $\boldsymbol{r}_0 \in \mathcal{W}_m$ there exists a correction $\boldsymbol{c}_m \in \mathcal{C}_m$ such that $\boldsymbol{x}_m = \boldsymbol{x}_0 + \boldsymbol{c}_m$ is a solution of* (3.1). *The linear system* (3.1) *then is necessarily consistent.*

Thus, if $\boldsymbol{r}_0 \in \mathcal{W}_m$, there exists a $\boldsymbol{c}_m \in \mathcal{C}_m$ with $A\boldsymbol{c}_m = \boldsymbol{r}_0$ and a reasonable approximation process should return $\boldsymbol{h}_m = \boldsymbol{r}_0$ with the approximation error $\boldsymbol{r}_m = \boldsymbol{0}$.

Given an approximation $\boldsymbol{h}_m \in \mathcal{W}_m$ to $\boldsymbol{r}_0$, any solution $\boldsymbol{c}_m$ of the consistent linear equation

$$A|_{\mathcal{C}_m}\boldsymbol{c} = \boldsymbol{h}_m \tag{3.4}$$

defines an iterate $\boldsymbol{x}_m = \boldsymbol{x}_0 + \boldsymbol{c}_m$ with the same residual $\boldsymbol{r}_m = \boldsymbol{r}_0 - \boldsymbol{h}_m$. Two solutions of (3.4) differ only in directions of the subspace $\mathcal{C}_m \cap \mathcal{N}(A)$. Thus, the correction $\boldsymbol{c}_m$ is uniquely determined if and only if

$$\mathcal{C}_m \cap \mathcal{N}(A) = \{\boldsymbol{0}\}, \tag{3.5}$$

i. e., if A is injective on the subspace $\mathcal{C}_m$. Obviously, if A is regular, this holds true for all m.

By (3.2), the approximation spaces form a nested sequence

$$\{\boldsymbol{0}\} = \mathcal{W}_0 \subseteq \mathcal{W}_1 \subseteq \cdots \subseteq \mathcal{W}_{m-1} \subseteq \mathcal{W}_m \subseteq \cdots .$$

If we attempt to apply the results of Eiermann and Ernst to our setting, the main difficulty arise from the fact that $\dim \mathcal{W}_m = m$ is not guaranteed. That is, the subset inclusions above must not be proper ones. Of course, if $\mathcal{W}_{m-1} = \mathcal{W}_m$, the successive approximation problems remain unchanged. If we are only interested in abstract MR and OR methods without reference to any operator equation we may simply renumber the sequence of approximation spaces. If, however, $\mathcal{W}_m$ is defined by (3.3) this is not so easily possible. In Section 3.4 we shall introduce a look-ahead strategy to remedy this deficiency.

3.2 MR Corrections and the Subspace Inverse

A popular choice of the correction $\boldsymbol{c}$ from the correction space $\mathcal{C}$ is to require the *minimal residual* (MR) condition

$$\|\boldsymbol{b} - A\boldsymbol{x}^{\mathrm{MR}}\| = \|\boldsymbol{r} - A\boldsymbol{c}^{\mathrm{MR}}\| = \min_{\boldsymbol{c}\in\mathcal{C}} \|\boldsymbol{r} - A\boldsymbol{c}\|, \tag{3.6}$$

where $\boldsymbol{r} = \boldsymbol{b} - A\boldsymbol{x}_0$ is the residual of the initial guess. The corresponding approximation $\boldsymbol{h}^{\mathrm{MR}} = A\boldsymbol{c}^{\mathrm{MR}}$ is characterized by the orthogonal projection $P_{\mathcal{W}} : \mathcal{H} \to \mathcal{H}$ onto $\mathcal{W} := A\mathcal{C}$:

$$\boldsymbol{h}^{\mathrm{MR}} = P_{\mathcal{W}}\boldsymbol{r}, \qquad \boldsymbol{r}^{\mathrm{MR}} = \boldsymbol{r} - \boldsymbol{h}^{\mathrm{MR}} = (I - P_{\mathcal{W}})\boldsymbol{r} \perp \mathcal{W}. \tag{3.7}$$

The solutions of (3.6) can be described in terms of a special constrained generalized inverse of A (cf. Section 2 and [14, Section 3.6]). Given a subspace $\mathcal{C}$ we call $A^{\mathcal{C}} := (AP_{\mathcal{C}})^{\dagger}$ the *subspace inverse of A with respect to $\mathcal{C}$*. The solution $\boldsymbol{c}^{\mathrm{MR}}$ of (3.6) with minimal norm is

$$\boldsymbol{c}^{\mathrm{MR}}_{\min} = A^{\mathcal{C}}\boldsymbol{r}$$

and all corrections which satisfy (3.6) are given by

$$\boldsymbol{c}^{\mathrm{MR}} = A^{\mathcal{C}}\boldsymbol{r} + \boldsymbol{z} \quad \text{with } \boldsymbol{z} \in \{\, (I - A^{\mathcal{C}}A)\boldsymbol{u} \,:\, \boldsymbol{u} \in \mathcal{C} \,\} \tag{3.8}$$

(cf. [14, Corollary 3.6.1]). We summarize some properties of this generalized inverse

Proposition 3.2.1. *The subspace inverse $A^{\mathcal{C}} = (AP_{\mathcal{C}})^{\dagger}$ of A with respect to $\mathcal{C}$ satisfies*

(i) $\mathcal{R}(A^{\mathcal{C}}) = \mathcal{R}(A^{\mathcal{C}}A) \subseteq \mathcal{C}$,

(ii) $\mathcal{N}(A^{\mathcal{C}}) = (A\mathcal{C})^{\perp} = \mathcal{W}^{\perp}$,

(iii) $A^{\mathcal{C}}AA^{\mathcal{C}} = A^{\mathcal{C}}$,

(iv) $AA^{\mathcal{C}} = P_{\mathcal{W}}$,

(v) $A^{\mathcal{C}}A$ *is a projection onto* $\mathcal{R}(A^{\mathcal{C}})$ *and*

(vi) $\mathcal{N}(A^{\mathcal{C}}A) = \{\boldsymbol{z} \in \mathcal{H} : A\boldsymbol{z} \perp A\mathcal{C}\}$.

Proof. From the definition of the constrained inverse and known properties of the Moore-Penrose inverse (cf. [14, Theorem 1.2.2]) we conclude

$$\mathcal{R}(A^{\mathcal{C}}) = \mathcal{R}((AP_{\mathcal{C}})^{\dagger}) = \mathcal{R}((AP_{\mathcal{C}})^{*}) = \mathcal{R}(P_{\mathcal{C}}A^{*})P_{\mathcal{C}}\mathcal{R}(A^{*}) \subseteq \mathcal{C}$$

and

$$\mathcal{N}(A^{\mathcal{C}}) = \mathcal{N}((AP_{\mathcal{C}})^{\dagger}) = \mathcal{R}(AP_{\mathcal{C}})^{\perp} = (A\mathcal{C})^{\perp} = \mathcal{W}^{\perp}$$

which prove (ii) and the inclusion $\mathcal{R}(A^{\mathcal{C}}) \subseteq \mathcal{C}$ in (i).

This inclusion implies $P_{\mathcal{C}}A^{\mathcal{C}} = A^{\mathcal{C}}$ and thus

$$A^{\mathcal{C}}AA^{\mathcal{C}} = A^{\mathcal{C}}AP_{\mathcal{C}}A^{\mathcal{C}} = (AP_{\mathcal{C}})^{\dagger}(AP_{\mathcal{C}})(AP_{\mathcal{C}})^{\dagger} = (AP_{\mathcal{C}})^{\dagger} = A^{\mathcal{C}}.$$

Similarly, we get

$$AA^{\mathcal{C}} = AP_{\mathcal{C}}A^{\mathcal{C}} = (AP_{\mathcal{C}})(AP_{\mathcal{C}})^{\dagger} = P_{\mathcal{R}(AP_{\mathcal{C}})} = P_{\mathcal{W}}$$

since $\mathcal{R}(AP_{\mathcal{C}}) = A\mathcal{C} = \mathcal{W}$.

As earlier stated, the projection property $(A^{\mathcal{C}}A)^2 = A^{\mathcal{C}}A$ follows immediately from the second Penrose condition (iii). So, demonstrating $\mathcal{R}(A^{\mathcal{C}}A) = \mathcal{R}(A^{\mathcal{C}})$ proves (v) and the remaining part of (i). The inclusion $\mathcal{R}(A^{\mathcal{C}}A) \subseteq \mathcal{R}(A^{\mathcal{C}})$ is trivial. Now, let $\boldsymbol{v} \in \mathcal{R}(A^{\mathcal{C}})$. Hence, there exists a vector $\boldsymbol{u} \in \mathcal{H}$ such that $\boldsymbol{v} = A^{\mathcal{C}}\boldsymbol{u}$. Using $A^{\mathcal{C}}AA^{\mathcal{C}} = A^{\mathcal{C}}$ we compute $A^{\mathcal{C}}A\boldsymbol{v} = A^{\mathcal{C}}AA^{\mathcal{C}}\boldsymbol{u} = A^{\mathcal{C}}\boldsymbol{u} = \boldsymbol{v}$, i. e., $\boldsymbol{v} \in \mathcal{R}(A^{\mathcal{C}}A)$, which implies $\mathcal{R}(A^{\mathcal{C}}) \subseteq \mathcal{R}(A^{\mathcal{C}}A)$.

Assertion (vi) follows easily: A vector $\boldsymbol{z} \in \mathcal{H}$ belongs to the null-space of $A^{\mathcal{C}}A$ if and only if $A\boldsymbol{z} \in \mathcal{N}(A^{\mathcal{C}}) = \mathcal{W}^{\perp}$ (note that this includes all $\boldsymbol{z} \in \mathcal{N}(A)$). □

Corollary 3.2.2. *The set of all solutions of* (3.6) *is*

$$\left\{\boldsymbol{c}^{\mathrm{MR}} = A^{\mathcal{C}}\boldsymbol{r} + \boldsymbol{z} \,|\, \boldsymbol{z} \in \mathcal{N}(A) \cap \mathcal{C}\right\}$$

Proof. We have to show that

$$\left\{\, (I - A^{\mathcal{C}}A)\boldsymbol{u} \,:\, \boldsymbol{u} \in \mathcal{C} \,\right\} = \mathcal{N}(A) \cap \mathcal{C}.$$

Let $\boldsymbol{z} = (I - A^{\mathcal{C}}A)\boldsymbol{u}$ with $\boldsymbol{u} \in \mathcal{C}$. Since $\mathcal{R}(A^{\mathcal{C}}) \subseteq \mathcal{C}$ we have $\boldsymbol{z} \in \mathcal{C}$. A simple computation gives $A\boldsymbol{z} = A(I - A^{\mathcal{C}}A)\boldsymbol{u} = (I - AA^{\mathcal{C}})A\boldsymbol{u} = (I - P_{\mathcal{W}})A\boldsymbol{u} = \boldsymbol{0}$ where the last equation follows from $A\boldsymbol{u} \in \mathcal{W} = A\mathcal{C}$.

Now let $\boldsymbol{z} \in \mathcal{N}(A) \cap \mathcal{C}$ and set $\boldsymbol{u} = (I - A^{\mathcal{C}})\boldsymbol{z}$. Then $\boldsymbol{u} \in \mathcal{C}$ and $(I - A^{\mathcal{C}}A)\boldsymbol{u} = (I - A^{\mathcal{C}}A)(I - A^{\mathcal{C}})\boldsymbol{z} = (I - A^{\mathcal{C}}A - A^{\mathcal{C}} + A^{\mathcal{C}}AA^{\mathcal{C}})\boldsymbol{z} = (I - A^{\mathcal{C}}A - A^{\mathcal{C}} + A^{\mathcal{C}})\boldsymbol{z} = \boldsymbol{z}$ since $A\boldsymbol{z} = \boldsymbol{0}$. □

Note that the proof of the first inclusion is a reformulation of (3.4) and the subsequent considerations.

One may ask, when the inclusion in (i) of Proposition 3.2.1 is in fact an equality. It turns out, that this is characterized by the assumption mentioned in (3.5) which also simplifies some other issues.

Proposition 3.2.3. *With the notation used in this section the following statements are equivalent*

(i) *the least squares problem* (3.6) *has a unique solution,*

(ii) $\mathcal{N}(A) \cap \mathcal{C} = \{\boldsymbol{0}\}$,

(iii) $\mathcal{R}(A^{\mathcal{C}}A) = \mathcal{C}$,

(iv) $\mathcal{N}(A^{\mathcal{C}}A) \cap \mathcal{C} = \{\boldsymbol{0}\}$,

Proof. The equivalence of the first and second item follows form (3.8) and Corollary 3.2.2.

Since $A^{\mathcal{C}}A$ is a projection, each vector $\boldsymbol{v} \in \mathcal{H}$ can be uniquely decomposed as $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{z}$ with $\boldsymbol{w} \in \mathcal{R}(A^{\mathcal{C}}A)$ and $\boldsymbol{z} \in \mathcal{N}(A^{\mathcal{C}}A)$. Let $\boldsymbol{c} \in \mathcal{C}$ and denote by $\boldsymbol{c} = \boldsymbol{d} + \boldsymbol{z}$ this decomposition. That is $\boldsymbol{d} \in \mathcal{R}(A^{\mathcal{C}}A)$ and, since $\boldsymbol{c}$ and $\boldsymbol{d}$ belong to $\mathcal{C}$, we get $\boldsymbol{z} = \boldsymbol{c} - \boldsymbol{d} \in \mathcal{N}(A^{\mathcal{C}}A) \cap \mathcal{C}$. Thus we have the direct decomposition

$$\mathcal{R}(A^{\mathcal{C}}A) \oplus \left[\mathcal{N}(A^{\mathcal{C}}A) \cap \mathcal{C}\right] = \mathcal{C}.$$

This shows the equivalence of (iii) and (iv).

We finish the proof by demonstrating the equivalence of (ii) and (iv). The direction from left to right follows from equation (vi) in Proposition 3.2.1: Assume $\boldsymbol{z} \in \mathcal{N}(A^{\mathcal{C}}A) \cap \mathcal{C}$ and $\boldsymbol{z} \neq \boldsymbol{0}$. Then (ii) implies $A\boldsymbol{z} \neq \boldsymbol{0}$ and since $A\boldsymbol{z} \in \mathcal{W}$ we have a contradiction to $A\boldsymbol{z} \perp \mathcal{W}$. To prove the other direction assume $\boldsymbol{z} \in \mathcal{C}$ with $A\boldsymbol{z} = \boldsymbol{0}$. Then there holds also $A^{\mathcal{C}}A\boldsymbol{z} = \boldsymbol{0}$ which by (iv) implies that $\boldsymbol{z}$ is zero. □

In the case characterized by Proposition 3.2.3, the subspace inverse coincides with the generalized inverse $A^{(2)}_{T,S}$ described in [5, Theorem 2.12]. It is characterized there as the unique operator (if it exists) satisfying the second Penrose condition and having prescribed range and nullspace.

For later use we state

Lemma 3.2.4. *Let* $\boldsymbol{c}^{\mathrm{MR}}$ *denotes an arbitrary solution of* (3.6). *Assume an orthonormal basis* $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_\ell\}$ *of* $\mathcal{N}(A) \cap \mathcal{C}$ *is given and let* $Z = \begin{bmatrix}\boldsymbol{z}_1 & \ldots & \boldsymbol{z}_\ell\end{bmatrix}$. *Then the subspace inverse solution (i. e., the solution of* (3.6) *which has minimal norm) is given by*

$$\boldsymbol{c}^{\mathrm{MR}}_{\mathrm{min}} = \boldsymbol{c}^{\mathrm{MR}} - Z\,(Z^*\boldsymbol{c}^{\mathrm{MR}}).$$

If $\mathcal{N}(A) \cap \mathcal{C}$ *is one-dimensional and* $\boldsymbol{z} \in \mathcal{N}(A) \cap \mathcal{C}$ *is an arbitrary nontrivial vector contained in it, the subspace inverse solution satisfies*

$$\boldsymbol{c}_{\min}^{\mathrm{MR}} = \boldsymbol{c}^{\mathrm{MR}} - \frac{(\boldsymbol{z}, \boldsymbol{c}_{L-1}^{\mathrm{MR}})}{(\boldsymbol{z}, \boldsymbol{z})}\boldsymbol{z}.$$

Proof. A simple consequence of Corollary 3.2.2 is that given an arbitrary solution $\boldsymbol{c}^{\mathrm{MR}}$ the set of all solutions of (3.6) is

$$\left\{ \boldsymbol{c}^{\mathrm{MR}} - \boldsymbol{z} \,|\, \boldsymbol{z} \in \mathcal{N}(A) \cap \mathcal{C} \right\}.$$

We are looking for $\boldsymbol{c}_{\min}^{\mathrm{MR}}$ with

$$\|\boldsymbol{c}_{\min}^{\mathrm{MR}}\| = \min_{\boldsymbol{z} \in \mathcal{N}(A) \cap \mathcal{C}} \|\boldsymbol{c}^{\mathrm{MR}} - \boldsymbol{z}\|.$$

Writing $\boldsymbol{z} = Z\boldsymbol{g}$ with $\boldsymbol{g} \in \mathbb{C}^{\ell}$ this transforms into the least squares problem

$$\min_{\boldsymbol{g} \in \mathbb{C}^{\ell}} \|\boldsymbol{c}^{\mathrm{MR}} - Z\boldsymbol{g}\|$$

which is solved by $\boldsymbol{g} = Z^{\dagger}\boldsymbol{c}^{\mathrm{MR}} = Z^{*}\boldsymbol{c}^{\mathrm{MR}}$ (cf. [14] and the subsequent remark). In the one dimensional case the orthonormalization of the basis of $\mathcal{N}(A) \cap \mathcal{C}$ can be done implicitly in the formula for $\boldsymbol{c}_{\min}^{\mathrm{MR}}$. □

Remark 3.2.5. Regarding the original linear equation (3.1) and the corresponding least squares problem (3.6) it is more appropriate to ask for a solution $\boldsymbol{x}_{\min}^{\mathrm{MR}}$ with minimal norm property. Fortunately, the answer can be easily derived from the results above.

Note that $Z\,Z^{*} = P_{\mathcal{Z}}$ is the orthogonal projection onto $\mathcal{Z} := \mathcal{C} \cap \mathcal{N}(A)$. Thus, the main result of Lemma 3.2.4 can be rewritten as

$$\boldsymbol{c}_{\min}^{\mathrm{MR}} = (I - P_{\mathcal{Z}})\boldsymbol{c}^{\mathrm{MR}},$$

which suggests an alternative proof using Pythagoras' theorem. The squared norm of an arbitrary least squares solution $\boldsymbol{c}^{\mathrm{MR}} - \boldsymbol{z}$ with $\boldsymbol{z} \in \mathcal{Z}$ can be rewritten as

$$\|\boldsymbol{c}^{\mathrm{MR}} - \boldsymbol{z}\|^2 = \|P_{\mathcal{Z}}\boldsymbol{c}^{\mathrm{MR}} - \boldsymbol{z}\|^2 + \|(I - P_{\mathcal{Z}})\boldsymbol{c}^{\mathrm{MR}}\|^2$$

and thus is minimized for $\boldsymbol{z} = P_{\mathcal{Z}}\boldsymbol{c}^{\mathrm{MR}}$.

Since every solution of (3.6) is of the form $\boldsymbol{x}_0 + \boldsymbol{c}^{\mathrm{MR}} = \boldsymbol{x}_0 + A^{\mathcal{C}}\boldsymbol{r} - \boldsymbol{z}$ with $\boldsymbol{z} \in \mathcal{Z}$ we get

$$\boldsymbol{x}_{\min}^{\mathrm{MR}} = (I - P_{\mathcal{Z}})(\boldsymbol{x}_0 + A^{\mathcal{C}}\boldsymbol{r}) = (I - P_{\mathcal{Z}})\boldsymbol{x}^{\mathrm{MR}}, \tag{3.9}$$

where $\boldsymbol{x}^{\mathrm{MR}}$ denotes an arbitrary solution of (3.6). The reformulation

$$\boldsymbol{x}_{\min}^{\mathrm{MR}} = (I - P_{\mathcal{Z}})A^{\mathcal{C}}\boldsymbol{b} + (I - P_{\mathcal{Z}})(I - A^{\mathcal{C}}A)\boldsymbol{x}_0$$

shows that the MR solution with minimal norm is additively composed of one part only dependent on the right hand side and another part determined by the initial guess.

Finally, we investigate the question of when the MR solution supplies a least square solution or even a pseudoinverse solution of (3.1). The proof of the two following lemmata is straight forward.

Lemma 3.2.6. *With the notation used in this section the following statements are equivalent*

(i) $\boldsymbol{x}^{\mathrm{MR}}$ *is a least squares solution of* (3.1),

(ii) $\boldsymbol{x}^{\mathrm{MR}} = A^{\dagger}\boldsymbol{b} + \boldsymbol{z}$ *with* $\boldsymbol{z} \in \mathcal{N}(A)$,

(iii) $\boldsymbol{c}^{\mathrm{MR}} = A^{\dagger}\boldsymbol{r} + \tilde{\boldsymbol{z}}$ *with* $\tilde{\boldsymbol{z}} \in \mathcal{N}(A)$,

(iv) $\boldsymbol{h}^{\mathrm{MR}} = P_{\mathcal{R}(A)}\boldsymbol{r}$,

(v) $\boldsymbol{r}^{\mathrm{MR}} = P_{\mathcal{N}(A^*)}\boldsymbol{r} = P_{\mathcal{N}(A^*)}\boldsymbol{b}$,

(vi) $P_{\mathcal{R}(A)}\boldsymbol{r} \in \mathcal{W}$.

Lemma 3.2.7. *With the notation used in this section the following statements are equivalent*

(i) $\boldsymbol{x}_{\min}^{\mathrm{MR}}$ *is a least squares solution of* (3.1) *with minimal norm property,*

(ii) $\boldsymbol{x}_{\min}^{\mathrm{MR}} = A^{\dagger}\boldsymbol{b}$,

(iii) $\boldsymbol{c}^{\mathrm{MR}} = A^{\dagger}\boldsymbol{r} - P_{\mathcal{N}(A)}\boldsymbol{x}_0$,

(iv) $A^{\dagger}\boldsymbol{b} \in \boldsymbol{x}_0 + \mathcal{C}$.

3.3 Corrections from Nested Subspaces

The standard implementation to compute MR corrections from nested correction spaces as introduced in (3.2) relies on nested orthonormal bases of the *residual spaces*

$$\mathcal{V}_{m+1} := \operatorname{span}\{\boldsymbol{r}_0\} + \mathcal{W}_m. \tag{3.10}$$

The basis of $\mathcal{V}_{m+1}$ is generated inductively by orthonormalizing $A\boldsymbol{c}_m$ against a (previously constructed) orthonormal basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ of $\mathcal{V}_m$ using the (modified) Gram-Schmidt algorithm. Here $\boldsymbol{c}_m$ is an arbitrary vector from $\mathcal{C}_m \setminus \mathcal{C}_{m-1}$. If $P_{\mathcal{V}_m}$ denotes the orthogonal projection onto $\mathcal{V}_m$, a compact presentation of this procedure reads

$$\begin{aligned} \boldsymbol{v}_1 &= \boldsymbol{r}_0/\beta, \text{ where } \beta := \|\boldsymbol{r}_0\|, \\ \boldsymbol{v}_{m+1} &= \frac{(I - P_{\mathcal{V}_m})A\boldsymbol{c}_m}{\|(I - P_{\mathcal{V}_m})A\boldsymbol{c}_m\|} \qquad (m = 1, 2, \ldots). \end{aligned} \tag{3.11}$$

The algorithm terminates in step m if (and only if)

$$A\boldsymbol{c}_m \in \mathcal{V}_m. \tag{3.12}$$

If such an index m exists we define

$$L := \min\{m \,:\, A\boldsymbol{c}_m \in \mathcal{V}_m\} \tag{3.13}$$

and set $L := \infty$ otherwise. Note that in a finite dimensional space $\mathcal{H}$, which is the most relevant case for practical applications, we always have $L < \infty$.

With $C_m := \begin{bmatrix}\boldsymbol{c}_1 & \boldsymbol{c}_2 & \dots & \boldsymbol{c}_m\end{bmatrix}$ and $V_{m+1} := \begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \dots & \boldsymbol{v}_{m+1}\end{bmatrix}$, the first m orthonormalization steps establish the following Arnoldi-type decomposition of A,

$$AC_m = V_{m+1}\tilde{H}_m = V_m H_m + \begin{bmatrix}\boldsymbol{0} & \dots & \boldsymbol{0} & \eta_{m+1,m}\boldsymbol{v}_{m+1}\end{bmatrix}, \tag{3.14}$$

where $\tilde{H}_m = \left[\eta_{j,k}\right]_{j,k=1}^{m+1,m} \in \mathbb{C}^{(m+1)\times m}$ is an upper Hessenberg matrix and $H_m := \begin{bmatrix}I_m & \boldsymbol{0}\end{bmatrix}\tilde{H}_m \in \mathbb{C}^{m\times m}$ is the square matrix obtained by deleting the last row of $\tilde{H}_m$. The nonzero entries of $\tilde{H}_m$ are given by $\eta_{j,k} = (A\boldsymbol{c}_k, \boldsymbol{v}_j)$, $1 \le k \le j \le m$, and $\eta_{k+1,k} = \|(I - P_{\mathcal{V}_k})A\boldsymbol{c}_k\| \ge 0$, with equality holding if and only if $k = L$. In other words, $\tilde{H}_m$ is an unreduced upper Hessenberg matrix (and thus of full rank m) as long as $m < L$.

For notational convenience we set $\boldsymbol{v}_{L+1} = \boldsymbol{0}$, so (3.14) holds true for $m = L$. Note also, that the Hessenberg matrix $\tilde{H}_L$ has a zero last row.

With respect to the orthonormal basis V_{m+1} of $\mathcal{V}_{m+1}$, the vector $\boldsymbol{r}_0 = \beta\boldsymbol{v}_1 = V_{m+1}\beta\boldsymbol{u}_1^{(m+1)}$ has the coordinates $\beta\boldsymbol{u}_1^{(m+1)}$ (where $\boldsymbol{u}_1^{(m+1)}$ denotes the first unit vector of $\mathbb{C}^{m+1}$). Consequently, writing $\boldsymbol{c} \in \mathcal{C}_m$ as $\boldsymbol{c} = C_m\boldsymbol{y}$ with a coordinate vector $\boldsymbol{y} \in \mathbb{C}^m$, the corresponding residual is given by

$$\|\boldsymbol{b} - A\boldsymbol{x}\| = \|\boldsymbol{r}_0 - AC_m\boldsymbol{y}\| = \|V_{m+1}(\beta\boldsymbol{u}_1^{(m+1)} - \tilde{H}_m\boldsymbol{y})\| = \|\beta\boldsymbol{u}_1^{(m+1)} - \tilde{H}_m\boldsymbol{y}\|_2$$

($\|\cdot\|_2$ denotes the Euclidean norm in $\mathbb{C}^{m+1}$). For $\boldsymbol{c}_m^{\mathrm{MR}} = C_m\boldsymbol{y}_m^{\mathrm{MR}}$, the least squares problem (3.6) reduces therefore to

$$\|\beta\boldsymbol{u}_1^{(m+1)} - \tilde{H}_m\boldsymbol{y}_m^{\mathrm{MR}}\|_2 = \min_{\boldsymbol{y}\in\mathbb{C}^m}\|\beta\boldsymbol{u}_1^{(m+1)} - \tilde{H}_m\boldsymbol{y}\|_2. \tag{3.15}$$

This problem has a unique solution if and only if $\tilde{H}_m$ has full rank m, which is certainly true as long as $m < L$. On the other hand we know from Proposition 3.2.3, that the least squares problem has a unique solution if and only if $\mathcal{C}$ contains no directions in the nullspace of A. Thus, for $m < L$, we have $\mathcal{C}_m \cap \mathcal{N}(A) = \{\boldsymbol{0}\}$ and

$$\boldsymbol{c}_m^{\mathrm{MR}} = C_m\boldsymbol{y}_m^{\mathrm{MR}} = \beta C_m\tilde{H}_m^{\dagger}\boldsymbol{u}_1^{(m+1)} = A^{\mathcal{C}_m}\boldsymbol{r}_0 \tag{3.16}$$

is the (uniquely determined) MR correction. We should mention here that, if $\tilde{H}_m$ has full rank m, then the Moore-Penrose pseudoinverse is given by $\tilde{H}_m^{\dagger} = (\tilde{H}_m^{\mathrm{H}}\tilde{H}_m)^{-1}\tilde{H}_m^{\mathrm{H}}$ (cf. [14, Theorem 1.3.2]).

Usually the solution of (3.15) is computed using an inductively generated QR decomposition of $\tilde{H}_m$. Assume we have already constructed

$$Q_{m-1}\tilde{H}_{m-1} = \begin{bmatrix} R_{m-1} \\ \boldsymbol{0}^\top \end{bmatrix},$$

with $Q_{m-1} \in \mathbb{C}^{m\times m}$ unitary ($Q_{m-1}^{\mathrm{H}} Q_{m-1} = I_m$) and $R_{m-1} \in \mathbb{C}^{(m-1)\times(m-1)}$ upper triangular (and nonsingular since $\tilde{H}_{m-1}$ has full rank), then

$$\begin{bmatrix} Q_{m-1} & \boldsymbol{0} \\ \boldsymbol{0}^\top & 1 \end{bmatrix} \tilde{H}_m = \begin{bmatrix} Q_{m-1} & \boldsymbol{0} \\ \boldsymbol{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \tilde{H}_{m-1} & \boldsymbol{h}_m \\ \boldsymbol{0}^\top & \eta_{m+1,m} \end{bmatrix} = \begin{bmatrix} R_{m-1} & \boldsymbol{t} \\ \boldsymbol{0}^\top & \tau \\ \boldsymbol{0}^\top & \eta_{m+1,m} \end{bmatrix}. \tag{3.17}$$

We define a Givens rotation

$$G_m := \begin{bmatrix} I_{k-1} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & c_m & s_m e^{-i\phi_m} \\ \boldsymbol{0} & -s_m e^{i\phi_m} & c_m \end{bmatrix} \qquad (c_m, s_m \geq 0, c_m^2 + s_m^2 = 1, \phi_m \in \mathbb{R}) \tag{3.18}$$

to accomplish

$$G_m \begin{bmatrix} Q_{m-1} & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{bmatrix} \tilde{H}_m = \begin{bmatrix} R_{m-1} & \boldsymbol{t} \\ \boldsymbol{0}^\top & \rho \\ \boldsymbol{0}^\top & 0 \end{bmatrix} \tag{3.19}$$

and set

$$Q_m := G_m \begin{bmatrix} Q_{m-1} & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{bmatrix} \qquad \text{and} \qquad R_m := \begin{bmatrix} R_{m-1} & \boldsymbol{t} \\ \boldsymbol{0}^\top & \rho \end{bmatrix}. \tag{3.20}$$

We can verify by a simple calculation that the appropriate parameters of the Givens rotation are

$$\begin{aligned} c_m &:= \frac{|\tau|}{\sqrt{|\tau|^2 + \eta_{m+1,m}^2}}, \quad s_m := \frac{\eta_{m+1,m}}{\sqrt{|\tau|^2 + \eta_{m+1,m}^2}}, \\ \phi_m &:= \arg(\eta_{m+1,m}) - \arg(\tau) = -\arg(\tau), \end{aligned} \tag{3.21}$$

which results in

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

Inserting this approach recursively we conclude with

$$Q_m = G_m \begin{bmatrix} G_{m-1} & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{bmatrix} \begin{bmatrix} G_{m-2} & O \\ O & I_2 \end{bmatrix} \cdots \begin{bmatrix} G_1 & O \\ O & I_{m-1} \end{bmatrix},$$

and

$$Q_m \tilde{H}_m = \begin{bmatrix} R_m \\ \boldsymbol{0} \end{bmatrix}. \tag{3.22}$$

Since ρ is nonzero as long as $\eta_{m+1,m} \neq 0$ the triangular matrix R_m is nonsingular if $m < L$. Even if $\eta_{m+1,m} = 0$, i. e., $m = L$, we may obtain a nonsingular R_L, namely if and only if $\tau \neq 0$ (cf. (3.28)). If $\rho = 0$, Equation (3.22) holds true for c_m and s_m chosen arbitrarily subject to $c_m^2 + s_m^2 = 1$.

Using this QR decomposition we can rewrite (3.15) as

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

where $[\boldsymbol{q}_m^\top, q_{m+1,1}^{(m)}]^\top = Q_m \boldsymbol{u}_1^{(m+1)}$ ($\boldsymbol{q}_m \in \mathbb{C}^m$) denotes the first column of Q_m. The unique solution of the above least-squares problem is $\boldsymbol{y}_m^{\mathrm{MR}} = \beta R_m^{-1} \boldsymbol{q}_m$ and the associated least-squares error is given by $\|\boldsymbol{r}_m^{\mathrm{MR}}\| = \beta |q_{m+1,1}^{(m)}|$.

More generally, (3.22) implies

$$\tilde{H}_m = Q_m^{\mathrm{H}} \begin{bmatrix} R_m \\ \boldsymbol{0}^\top \end{bmatrix} \qquad \text{and} \qquad \tilde{H}_m^\dagger = \begin{bmatrix} R_m^{-1} & \boldsymbol{0} \end{bmatrix} Q_m \tag{3.23}$$

for $m < L$. As another consequence the Arnoldi-type decomposition (3.14) can be rewritten as

$$AC_m = V_{m+1} \tilde{H}_m = V_{m+1} Q_m^{\mathrm{H}} \begin{bmatrix} I_m \\ \boldsymbol{0} \end{bmatrix} R_m =: \hat{V}_m R_m \tag{3.24}$$

for $m < L$ (a similar relations hold true for $m = L$, cf. (3.28) and (3.32)). The basis $\hat{V}_m$ is often referred as Paige-Saunders basis (cf. [52]) and forms an orthonormal basis of $\mathcal{W}_m$.

3.4 The Breakdown Behavior

If the orthonormalization process (3.11) breaks down in step L, the Arnoldi-type decomposition reduces to

$$AC_L = V_L H_L \tag{3.25}$$

and the least squares problem (3.6) is equivalent to

$$\|\beta \boldsymbol{u}_1^{(L)} - H_L \boldsymbol{y}_L^{\mathrm{MR}}\|_2 = \min_{\boldsymbol{y}\in\mathbb{C}^m} \|\beta \boldsymbol{u}_1^{(L)} - H_L \boldsymbol{y}\|_2, \tag{3.26}$$

where $\boldsymbol{y}_L^{\mathrm{MR}}$ is the coordinate vector of the MR correction $\boldsymbol{c}_L^{\mathrm{MR}} = V_L \boldsymbol{y}_L^{\mathrm{MR}}$. We distinguish two cases:

Definition 3.4.1. In the termination step L of (3.11) a *regular breakdown*[1] occurs if the linear system

$$H_L \boldsymbol{y} = \beta \boldsymbol{u}_1^{(L)}$$

has a unique solution. Otherwise, the termination is called a *singular breakdown*[2].

The situation in case of a regular breakdown is the same as if A is nonsingular:

Proposition 3.4.2. *If the Gram-Schmidt procedure* (3.11) *breaks down regularly in some step L, then the linear system* (3.1) *is necessarily consistent and the (uniquely determined) MR correction*

$$\boldsymbol{c}_L^{\mathrm{MR}} = A^{\mathfrak{C}_L} \boldsymbol{r}_0 = \beta C_L H_L^{-1} \boldsymbol{u}_1^{(L)} \tag{3.27}$$

leads to a solution $\boldsymbol{x}_L^{\mathrm{MR}} = \boldsymbol{x}_0 + \boldsymbol{c}_L^{\mathrm{MR}}$ of (3.1), *i. e., $A\boldsymbol{x}_L^{\mathrm{MR}} = \boldsymbol{b}$.*

Proof. Obviously, if H_L is regular, the solution of the least squares problem (3.26) is $\boldsymbol{y}_L^{\mathrm{MR}} = \beta H_L^{-1} \boldsymbol{u}_1^{(L)}$. The correction $\boldsymbol{c}_L^{\mathrm{MR}} = C_L \boldsymbol{y}_L^{\mathrm{MR}}$ satisfies $\boldsymbol{r}_0 - A\boldsymbol{c}_L^{\mathrm{MR}} = \boldsymbol{r}_L^{\mathrm{MR}} = \boldsymbol{0}$, i. e. $A\boldsymbol{x}_L^{\mathrm{MR}} = A\boldsymbol{x}_0 + A\boldsymbol{c}_L^{\mathrm{MR}} = A\boldsymbol{x}_0 + \boldsymbol{r}_0 = \boldsymbol{b}$. □

The coordinate vector $\boldsymbol{y}_L^{\mathrm{MR}}$ can also be expressed in terms of the implicitly computed QR factorization of H_L. Recall that

$$Q_{L-1} H_L = Q_{L-1} \begin{bmatrix} \tilde{H}_{L-1} & \boldsymbol{h}_L \end{bmatrix} = \begin{bmatrix} R_{L-1} & \boldsymbol{t} \\ \boldsymbol{0}^\top & \tau \end{bmatrix} = R_L \tag{3.28}$$

where $\tau \neq 0$. Since $\eta_{L+1,L} = 0$ we obtain $c_L = 1$ and $s_L = 0$ (cf. (3.21)). Thus we have $H_L^{-1} = R_L^{-1} Q_{L-1}$ and

$$\boldsymbol{y}_L^{\mathrm{MR}} = \beta R_L^{-1} Q_{L-1} \boldsymbol{u}_1^{(L)}.$$

Note further, that our proof above shows $\boldsymbol{r}_0 = A\boldsymbol{c}_L^{\mathrm{MR}} \in \mathcal{W}_L$ (see also Lemma 3.1.1). This characterizes the regular breakdown:

Proposition 3.4.3. *The following statements are equivalent:*

(i) (3.11) *breaks down regularly in step m,*

(ii) $\boldsymbol{r}_0 \in \mathcal{W}_m = \mathrm{span}\{A\boldsymbol{c}_1, \ldots, A\boldsymbol{c}_m\}$,

(iii) $\boldsymbol{x}_m^{\mathrm{MR}}$ *is a solution of* (3.1),

(iv) $\mathcal{W}_m = \mathcal{V}_m$,

[1] In the context of Krylov subspace methods (cf. Section 4) Brown and Walker use in [7] the terms *breakdown through degeneracy of the Krylov space* and

[2] *breakdown through rank deficiency of the least-squares problem.*

Proof. We have already seen in Proposition 3.4.2, that (i) implies (ii) and (iii). In fact, by Lemma 3.1.1, (ii) and (iii) are equivalent.

Suppose we have a regular breakdown in step $m = L$, that is $AC_L = V_L H_L$ where H_L is nonsingular. Then

$$\mathcal{W}_L = \operatorname{span}\{AC_L\} = \operatorname{span}\{V_L\} = \mathcal{V}_L,$$

which proves (i)⇒(iv).

By the definition $\mathcal{V}_m = \operatorname{span}\{\boldsymbol{r}_0\} + \mathcal{W}_{m-1}$ we have, that $\mathcal{V}_m = \mathcal{W}_m$ implies $\boldsymbol{r}_0 \in \mathcal{W}_m$, that is (iv)⇒(ii).

Now, suppose $\boldsymbol{r}_0 \in \mathcal{W}_m$. Since $\mathcal{W}_m = A\mathcal{C}_m$ there exists an $\boldsymbol{f} \in \mathbb{C}^m$ such that $\boldsymbol{r}_0 = AC_m\boldsymbol{f}$. Using the Arnoldi-type decomposition (3.14) we get

$$\boldsymbol{r}_0 = AC_m\boldsymbol{f} = V_{m+1}\tilde{H}_{m+1}\boldsymbol{f}.$$

On the other hand we know that $\boldsymbol{r}_0 = \beta V_{m+1}\boldsymbol{u}_1^{(m+1)}$. Thus we have

$$\tilde{H}_{m+1}\boldsymbol{f} = \beta\boldsymbol{u}_1^{(m+1)}. \tag{3.29}$$

Due to the upper Hessenberg structure of $\tilde{H}_{m+1}$ this is only possible if $\tilde{H}_{m+1}$ is not unreduced, that is $\eta_{m+1,m} = 0$ and $m = L$. Then (3.29) reduces to

$$H_L\boldsymbol{f} = \begin{bmatrix}\tilde{H}_{L-1} & \boldsymbol{h}_L\end{bmatrix}\boldsymbol{f} = \beta\boldsymbol{u}_1^L.$$

The existence of a vector $\boldsymbol{f}$ satisfying the above equation implies that $\boldsymbol{u}_1^L$ lies in the column span of H_L. Since $\tilde{H}_{L-1}$ is an unreduced upper Hessenberg matrix, the first $L-1$ columns of H_L are linearly independent. Assuming that H_L is singular is therefore equivalent to stating $\boldsymbol{h}_L \in \operatorname{span}\{\tilde{H}_{L-1}\}$. But this would imply that $\boldsymbol{u}_1^L$ lies already in the column span of $\tilde{H}_{L-1}$, which is impossible since the columns of the triangular matrix $\begin{bmatrix}\boldsymbol{u}_1^L & \tilde{H}_{L-1}\end{bmatrix}$ forms a linear independent set of vectors in $\mathbb{C}^L$. Thus we conclude that H_L is regular and have now demonstrated (ii)⇒(i).

Taking all together we have (i)⇒(iv)⇒(ii)⇒(i) and (ii)⇔(iii). □

We now turn to the singular breakdown.

Proposition 3.4.4. *The following statements are equivalent:*

(i) (3.11) *breaks down singularly in step m;*

(ii) *the constrained least squares problem*

$$\|\boldsymbol{r}_0 - A\boldsymbol{c}^{\mathrm{MR}}\| = \min_{\boldsymbol{c}\in\mathcal{C}_m}\|\boldsymbol{r} - A\boldsymbol{c}\| = \min_{\boldsymbol{y}\in\mathbb{C}^m}\|\beta\boldsymbol{u}_1^{(m+1)} - \tilde{H}_m\boldsymbol{y}\|_2 \tag{3.30}$$

has no unique solution,

(iii) $\operatorname{rank}(\tilde{H}_m) < m$,

(iv) $\tilde{H}_m^{\mathrm{H}}\tilde{H}_m$ *is singular,*[3]

(v) R_m *is singular,*

(vi) $\mathcal{C}_m \cap \mathcal{N}(A) \neq \{\boldsymbol{0}\}$,

(vii) $A\boldsymbol{c}_m \in \mathcal{W}_{m-1}$,

(viii) $\mathcal{W}_m = \mathcal{W}_{m-1}$,

(ix) $\dim \mathcal{W}_m < \dim \mathcal{C}_m = m$,

Proof. We have already discussed, that the least squares problem in (ii) has a unique solution if and only if $\tilde{H}_m$ has full rank m, that is (iii). It is easy to see, that (iii) is also equivalent to (iv) and (v) (cf. (3.22)).

(iii)⇔(i): As previously noted, $\operatorname{rank}(\tilde{H}_m) < m$ is only possible if $\eta_{m+1,m} = 0$ and $m = L$. Therefore the Arnoldi-type decomposition reduces to (3.25) and $\operatorname{rank}(H_L) = \operatorname{rank}(\tilde{H}_L)$ shows, that H_L is singular if and only if $\operatorname{rank}(\tilde{H}_L) < L$.

(i)⇒(vii): Let H_L be singular. Since $H_L = \left[\tilde{H}_{L-1}\boldsymbol{h}_L\right]$ and $\tilde{H}_{L-1}$ has full rank $L-1$ the last column of H_L must be a linear combination of the columns of $\tilde{H}_{L-1}$, i. e., there exists a vector $\boldsymbol{g} \in \mathbb{C}^{L-1}$ such that $\boldsymbol{h}_L = \tilde{H}_{L-1}\boldsymbol{g}$. The last column of (3.25) can be rewritten as

$$A\boldsymbol{c}_L = V_L\boldsymbol{h}_L = V_L\tilde{H}_{L-1}\boldsymbol{g} = AC_{L-1}\boldsymbol{g} \in A\mathcal{C}_{L-1} = \mathcal{W}_{L-1}.$$

(vii)⇒(vi): We have $A\boldsymbol{c}_m \in \operatorname{span}\{A\boldsymbol{c}_1, \ldots, A\boldsymbol{c}_{m-1}\}$. Thus, there exists $\gamma_1, \ldots, \gamma_{m-1}$ such that $A\boldsymbol{c}_m = \gamma_1 A\boldsymbol{c}_1 + \cdots + \gamma_{m-1}A\boldsymbol{c}_{m-1}$. Hence $\tilde{\boldsymbol{c}} := \gamma_1\boldsymbol{c}_1 + \cdots + \gamma_{m-1}\boldsymbol{c}_{m-1} - \boldsymbol{c}_m \in \mathcal{C}_m \cap \mathcal{N}(A)$ and $\tilde{\boldsymbol{c}} \neq \boldsymbol{0}$ since by definition $\boldsymbol{c}_m \notin \mathcal{C}_{m-1}$.

(vi)⇒(i): Suppose $\boldsymbol{0} \neq \boldsymbol{z} \in \mathcal{C}_m \cap \mathcal{N}(A)$, that is $\boldsymbol{z} = C_m\boldsymbol{f}$ with a nonzero coordinate vector $\boldsymbol{f} \in \mathbb{C}^m$. Using (3.14) we get

$$\boldsymbol{0} = A\boldsymbol{z} = AC_m\boldsymbol{f} = V_{m+1}\tilde{H}_m\boldsymbol{f}.$$

So, either $\tilde{H}_m\boldsymbol{f} = \boldsymbol{0}$, which implies (iii) and further (i), or the rows of V_{m+1} are linearly dependent, which by construction implies $\boldsymbol{v}_{m+1} = \boldsymbol{0}$ and $h_{m+1,m} = 0$. But then $m = L$ and the equation above reduces to $\boldsymbol{0} = V_L H_L\boldsymbol{f}$. Since the basis V_L is linearly independent, we conclude $H_L\boldsymbol{f} = \boldsymbol{0}$, i. e., H_L is singular.

Now, we finish the proof by noting that (vii) is equivalent to each of the subsequent conditions: For (viii) this can be seen directly from the expanded formulation $\operatorname{span}\{A\boldsymbol{c}_1, \ldots A\boldsymbol{c}_{m-1}, A\boldsymbol{c}_m\} = \operatorname{span}\{A\boldsymbol{c}_1, \ldots A\boldsymbol{c}_{m-1}\}$. Note further, that we have by definition $\dim \mathcal{C}_m = m$. Together with the nested structure of the correction spaces (cf. (3.2)) and $\mathcal{W}_m = A\mathcal{C}_m$ this yields the remaining equivalence. □

[3] In context of Krylov subspace methods this condition is derived by Smoch in [64], who examines the determinant of $\tilde{H}_m^{\mathrm{H}}\tilde{H}_m$.

Note that the equivalence of (ii) and (vi) could be also derived directly from Proposition 3.2.3.

Note further, that the vector $\boldsymbol{g}$ above can be characterized as the unique solution of $\tilde{H}_{L-1}\boldsymbol{y} = \boldsymbol{h}_L$. Any nontrivial vector in $\mathcal{N}(H_L)$ can be expressed in terms of this solution by

$$\tilde{\boldsymbol{g}} = \begin{bmatrix} \boldsymbol{g} \\ -1 \end{bmatrix} = \begin{bmatrix} \tilde{H}_{L-1}^{\dagger}\boldsymbol{h}_L \\ -1 \end{bmatrix}. \tag{3.31}$$

From the Hessenberg structure, there follows $\operatorname{rank}(H_L) < L$ and $\operatorname{rank}\left(\begin{bmatrix} \boldsymbol{u}_1^{(L)} & H_L \end{bmatrix}\right) = L$ which shows that $\boldsymbol{u}_1^{(L)} \notin \mathcal{R}(H_L)$ and $\operatorname{rank}(H_L) = L-1$. Thus, the nullspace of H_L (just as $\mathcal{N}(A) \cap \mathcal{C}_L$) is one-dimensional. We get $\mathcal{N}(H_L) = \operatorname{span}\{\tilde{\boldsymbol{g}}\}$ and $\mathcal{N}(A) \cap \mathcal{C}_L = \operatorname{span}\{C_L\tilde{\boldsymbol{g}}\}$. Applying the previous Givens rotations to H_L results in (cf. (3.17))

$$Q_{L-1}H_L = Q_{L-1}\begin{bmatrix} \tilde{H}_{L-1} & \boldsymbol{h}_L \end{bmatrix} = \begin{bmatrix} R_{L-1} & \boldsymbol{t} \\ \boldsymbol{0}^{\top} & \tau \end{bmatrix} = \begin{bmatrix} R_{L-1} & \boldsymbol{t} \\ \boldsymbol{0}^{\top} & 0 \end{bmatrix}, \tag{3.32}$$

where τ has to be zero since H_L is singular. For this reason, c_L and s_L in (3.21) could be chosen arbitrarily. This transforms the least squares problem (3.26) into

$$\min_{\boldsymbol{y}_L \in \mathbb{C}^L} \|\beta\boldsymbol{u}_1^{(L)} - H_L\boldsymbol{y}_L\|_2 = \min_{\substack{\boldsymbol{y}_{L-1}\in\mathbb{C}^{L-1} \\ \gamma\in\mathbb{C}}} \left\| \beta\, Q_{L-1}\boldsymbol{u}_1^{(L)} - \begin{bmatrix} R_{L-1} & \boldsymbol{t} \\ \boldsymbol{0}^{\top} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{y}_{L-1} \\ \gamma \end{bmatrix} \right\|_2 =$$

$$\min_{\substack{\boldsymbol{y}_{L-1}\in\mathbb{C}^{L-1} \\ \gamma\in\mathbb{C}}} \left\| \begin{bmatrix} \beta\boldsymbol{q}_{L-1} - R_{L-1}\boldsymbol{y}_{L-1} - \gamma\boldsymbol{t} \\ \beta\, q_{L,1}^{(L-1)} \end{bmatrix} \right\|_2,$$

where $\begin{bmatrix} \boldsymbol{q}_{L-1}^{\top} & q_{L,1}^{(L-1)} \end{bmatrix}^{\top} = Q_{L-1}\boldsymbol{u}_1^{(L)}$ denotes the first column of Q_{L-1}. Since the upper block can be forced to be zero by setting $\boldsymbol{y}_{L-1} := R_{L-1}^{-1}(\beta\boldsymbol{q}_{L-1} - \gamma\boldsymbol{t})$ the minimum is independent from γ. Note further, that $R_{L-1}^{-1}\boldsymbol{t} = \tilde{H}_{L-1}^{\dagger}\boldsymbol{h}_L = \boldsymbol{g}$ (cf. (3.23)) is the unique solution of $\tilde{H}_{L-1}\boldsymbol{y} = \boldsymbol{h}_L$ and $\beta R_{L-1}^{-1}\boldsymbol{q}_{L-1} = \boldsymbol{y}_{L-1}^{\mathrm{MR}}$ is the coordinate vector of the previous MR correction. In other words: All solutions of the above least-squares problem are in

$$\left\{ \begin{bmatrix} \boldsymbol{y}_{L-1}^{\mathrm{MR}} - \gamma\boldsymbol{g} \\ \gamma \end{bmatrix} \;:\; \gamma \in \mathbb{C} \right\} = \left\{ \begin{bmatrix} \boldsymbol{y}_{L-1}^{\mathrm{MR}} \\ 0 \end{bmatrix} - \gamma\tilde{\boldsymbol{g}} \;:\; \gamma \in \mathbb{C} \right\} \tag{3.33}$$

and the associated least-squares error is given by $\|\boldsymbol{r}_L^{\mathrm{MR}}\| = \beta|q_{L,1}^{(L-1)}| = \|\boldsymbol{r}_{L-1}^{\mathrm{MR}}\|$.

We summarize the consequences of these observations.

Proposition 3.4.5. *If the process* (3.11) *breaks down singularly in the Lth step, there holds:*

The linear system $H_L\boldsymbol{y} = \beta\boldsymbol{u}_1^{(L)}$ is inconsistent and

$$\boldsymbol{c}_{L-1}^{\mathrm{MR}} = C_{L-1}\boldsymbol{y}_{L-1}^{\mathrm{MR}} \;\in\; A^{\mathcal{C}_L}\boldsymbol{r}_0 + \mathcal{N}(A) \cap \mathcal{C}_L, \tag{3.34}$$

i. e., the (uniquely determined) $(L-1)$*th MR correction is a possible MR correction in the terminating step. Thus, the MR approximation process makes no progress, i. e.* $\|\boldsymbol{r}_L^{\mathrm{MR}}\| = \|\boldsymbol{r}_{L-1}^{\mathrm{MR}}\|$*. All MR corrections are in the set*

$$\left\{\boldsymbol{c}_{L-1}^{\mathrm{MR}} - \gamma\,(C_{L-1}\boldsymbol{g} - \boldsymbol{c}_L) \;:\; \gamma \in \mathbb{C}\right\} \tag{3.35}$$

and the correction with minimal norm (i. e., the subspace inverse correction) can be computed as

$$\begin{aligned} A^{\mathcal{C}_L}\boldsymbol{r}_0 = \boldsymbol{c}_{L-1}^{\mathrm{MR}} - \frac{(\boldsymbol{z}, \boldsymbol{c}_{L-1}^{\mathrm{MR}})}{(\boldsymbol{z},\boldsymbol{z})}\boldsymbol{z} \qquad & \textit{with}\;\; \boldsymbol{z} = C_{L-1}\boldsymbol{g} - \boldsymbol{c}_L \\ & \textit{and}\;\; \boldsymbol{g} = \tilde{H}_{L-1}^{\dagger}\boldsymbol{h}_L = R_{L-1}^{-1}\boldsymbol{t} \end{aligned} \tag{3.36}$$

where R_{L-1} *and* $\boldsymbol{t}$ *are defined as in* (3.32)*.*

Proof. Premultiplying (3.33) with C_L together with $\mathcal{N}(A) \cap \mathcal{C}_L = \operatorname{span}\{C_L\tilde{\boldsymbol{g}}\}$ result in equations (3.34) and (3.35). The remaining assertion follows using Lemma 3.2.4 with $\boldsymbol{z} = C_L\tilde{\boldsymbol{g}} = C_{L-1}\boldsymbol{g} - \boldsymbol{c}_L \in \mathcal{N}(A) \cap \mathcal{C}_L$. □

The equivalences in Proposition 3.4.4 are often formulated in its negated form, for example: The constrained least squares problem (3.30) has a unique solution (or, equivalently,

$$\begin{aligned} \mathcal{C}_m \cap \mathcal{N}(A) &= \{\boldsymbol{0}\}, && \text{or} \\ \operatorname{rank}(\tilde{H}_m) &= m, && \text{or} \\ \dim \mathcal{W}_m &= m, && \text{or} \end{aligned}$$

$\tilde{H}_m^{\top}\tilde{H}_m$ is regular, or R_m is regular) if and only if no singular breakdown occur in step m. If A is nonsingular this conditions are always satisfied, as we had already stated regarding (3.5). Thus

Corollary 3.4.6. *If* A *is nonsingular, no singular breakdown occurs.*

In other words, a singular breakdown is only possible (but not necessary) for a singular operator A. In practical computations with presence of round-off errors it is often a concern to detect near singularity. Brown and Walker suggest in [7] to monitor the condition number of R_m. A similar approach is, to observe the last diagonal entry of R_m, specifically, if ρ in (3.20) becomes close to zero (see also [64]).

The two following propositions provide several equivalent definitions for the breakdown index L. Each is a direct consequence of one of the breakdown characterizations in Propositions 3.4.3 and 3.4.4.

Proposition 3.4.7. *If* (3.11) *breaks down regularly, then the breakdown index L can be characterized as*

$$\begin{aligned} L &= \min\{m \,:\, A\boldsymbol{c}_m \in \mathcal{V}_m\} \\ &= \min\{m \,:\, \boldsymbol{r}_0 \in \mathcal{W}_m\} \\ &= \min\{m \,:\, \mathcal{W}_m = \mathcal{V}_m\} \\ &= \min\{m \,:\, A\boldsymbol{x}_m^{\mathrm{MR}} = \boldsymbol{b}\}. \end{aligned}$$

Proposition 3.4.8. *If* (3.11) *breaks down singularly, then the breakdown index L can be characterized as*

$$\begin{aligned} L &= \min\{m \,:\, A\boldsymbol{c}_m \in \mathcal{V}_m\} \\ &= \min\{m \,:\, A\boldsymbol{c}_m \in \mathcal{W}_{m-1}\} \\ &= \min\{m \,:\, \mathcal{W}_m = \mathcal{W}_{m-1}\} \\ &= \min\{m \,:\, \dim \mathcal{C}_m > \dim \mathcal{W}_m\} \\ &= \min\{m \,:\, \dim \mathcal{W}_m < m\} \\ &= \min\{m \,:\, \mathcal{C}_m \cap \mathcal{N}(A) \neq \{\boldsymbol{0}\}\} \\ &= \min\{m \,:\, \operatorname{rank}(\tilde{H}_m) < m\} \\ &= \min\{m \,:\, \det(\tilde{H}_m^{\mathrm{H}}\tilde{H}_m) = 0\} \\ &= \min\{m \,:\, r_{m,m} = 0\}, \end{aligned}$$

where $r_{m,m} = \rho$ is the last diagonal entry of R_m (cf. (3.20)*).*

Remark 3.4.9. In a certain sense, singular breakdowns can be avoided – at least if we consider the sequence of correction spaces (3.2) as properly chosen or explicitly given a priori (in contrast to, e. g., the implicitly constructed Krylov spaces – cf. Section 4).

Suppose we have

$$\begin{aligned} \mathcal{C}_1 \subset \cdots \subset \mathcal{C}_{L-1} \subset \mathcal{C}_L \subset \mathcal{C}_{L+1} \subset \cdots \\ \mathcal{W}_1 \subset \cdots \subset \mathcal{W}_{L-1} = \mathcal{W}_L \subset \mathcal{W}_{L+1} = A\mathcal{C}_{L+1} \subset \cdots . \end{aligned}$$

In the Lth step, the correction space is enlarged into an unwanted direction, namely in a direction contained in the nullspace of A. But there are useful directions in the subsequent correction spaces (or at least we hope so). So we want to throw away the misleading direction $\boldsymbol{c}_L \in \mathcal{C}_L \setminus \mathcal{C}_{L-1}$ and continue the Gram-Schmidt process with $\boldsymbol{c}_{L+1} \in \mathcal{C}_{L+1} \setminus \mathcal{C}_L$. The only interesting information, which should be saved, is the direction of $\mathcal{N}(A) \cap \mathcal{C}_L$. This can be done by computing and storing $\boldsymbol{z}$ as described in (3.36).

Throwing away $\boldsymbol{c}_L$ actually means to substitute it by an arbitrary vector $\boldsymbol{c}_{L+1}$ from $\mathcal{C}_{L+1} \setminus \mathcal{C}_L$. Since all relevant information on $\boldsymbol{c}_L$ is stored in $\boldsymbol{z}$ we can cancel

it from all subsequent correction spaces, i. e., we replace $\mathcal{C}_m$ with $m > L+1$ by $\mathcal{C}'_{m'} = \operatorname{span}\{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_{L-1}, \boldsymbol{c}_{L+1}, \ldots, \boldsymbol{c}_m\}$ where $m' = m-1$. In other words, we choose another sequence of correction spaces.

This strategy to avoid a singular breakdown can be applied more than once and by orthonormalizing the vectors $\boldsymbol{z}$ we obtain a basis $Z_\ell = \begin{bmatrix}\boldsymbol{z}_1 & \ldots & \boldsymbol{z}_\ell\end{bmatrix}$ of the intersection of $\mathcal{N}(A)$ and the original correction spaces. Using this basis, it is also possible to compute the subspace inverse solutions with respect to the original correction spaces, i. e., to minimize the norm of the correction by adding a suitable vector from $\operatorname{span}\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_\ell\}$ (cf. Remark 3.2.5). Example 3.4.10 illustrates this technique.

Of course, if $\mathcal{H}$ is finite dimensional, the spaces $\mathcal{C}_m$ form a strictly ascending sequence only up to a finite index and will stagnate henceforth. It may happen that also the termination at this index is a singular breakdown. Indeed, Proposition 3.4.2 implies that if the linear system (3.1) is inconsistent, any termination is a singular breakdown. Thus, the look-ahead strategy does not work in any situation.

Example 3.4.10. The sequence of correction spaces is defined by the nested bases

$$\boldsymbol{c}_1 = \begin{bmatrix}1\\0\\0\\\vdots\end{bmatrix}, \boldsymbol{c}_2 = \begin{bmatrix}1\\1\\0\\\vdots\end{bmatrix}, \boldsymbol{c}_3 = \begin{bmatrix}1\\1\\1\\\vdots\end{bmatrix}, \ldots .$$

We restrict ourselves to three dimensions. Consider

$$A = \begin{bmatrix}1 & 0 & 0\\1 & 0 & -1\\0 & 0 & 0\end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix}1\\0\\0\end{bmatrix} \quad \text{and} \quad \boldsymbol{x}_0 = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Thus we get $\boldsymbol{v}_1 = \boldsymbol{r}_0 = \boldsymbol{b}$ and orthonormalizing $A\boldsymbol{c}_1 = \begin{bmatrix}1 & 1 & 0\end{bmatrix}^\top$ against $\boldsymbol{v}_1$ yields in $\boldsymbol{v}_2 = \begin{bmatrix}0 & 1 & 0\end{bmatrix}$. In the second step the Gram-Schmidt procedure breaks down singularly with the Arnoldi-type decomposition

$$A\begin{bmatrix}1 & 1\\0 & 1\\0 & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\\0 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}, \quad \text{which results in } \boldsymbol{g} = \begin{bmatrix}1\end{bmatrix} \text{ and } \boldsymbol{z} = \begin{bmatrix}0\\1\\0\end{bmatrix}.$$

At this point we cancel $\boldsymbol{c}_2$ and the last column of the Hessenberg matrix and continue with orthonormalizing $A\boldsymbol{c}_3$ against $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. This leads to the decomposition

$$A\begin{bmatrix}1 & 1\\0 & 1\\0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\\0 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\1 & 0\end{bmatrix},$$

i. e., we have a regular breakdown in this step. The corresponding correction

$$\boldsymbol{c}^{\mathrm{MR}} = \begin{bmatrix}1 & 1\\0 & 1\\0 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}^\dagger\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}1 & 1\\0 & 1\\0 & 1\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

solves the system $A\boldsymbol{x} = \boldsymbol{b}$. Since $\operatorname{span}\{\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3\} = \mathcal{H}$, the solution in the correction space, which has minimal norm, coincides with the pseudoinverse solution and can be computed by solving the minimization problem

$$\min_{\zeta \in \mathbb{C}} \left\| \boldsymbol{c}^{\mathrm{MR}} + \zeta \boldsymbol{z} \right\|,$$

which is given by

$$\zeta = -\frac{(\boldsymbol{z}, \boldsymbol{c}^{\mathrm{MR}})}{(\boldsymbol{z}, \boldsymbol{z})} = -1 \quad \text{and} \quad \boldsymbol{c}^{\mathrm{MR}} + \zeta \boldsymbol{z} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Regarding Remark 3.4.9 and the subsequent example we should give a reformulation of the algorithm stated in (3.11). We assume a given correction space $\tilde{\mathcal{C}}$, which covers all possible correction spaces and may contain also directions in $\mathcal{N}(A)$. Algorithm 3.4.11 constructs a decomposition $\tilde{\mathcal{C}}_{m+\ell} = \mathcal{C}_m \oplus \mathcal{Z}_\ell \subseteq \tilde{\mathcal{C}}$ with $\mathcal{C}_m \cap \mathcal{N}(A) = \{\boldsymbol{0}\}$ and $\mathcal{Z}_\ell \subseteq \tilde{\mathcal{C}} \cap \mathcal{N}(A)$ and computes the corresponding subspace inverse solution

$$\boldsymbol{x}_{\min}^{\mathrm{MR}} = (I - P_{\mathcal{Z}_\ell})(\boldsymbol{x}_0 + A^{\mathcal{C}_m} \boldsymbol{r}_0) = (I - P_{\mathcal{Z}_\ell})(A^{\mathcal{C}_m} \boldsymbol{b} + (I - A^{\mathcal{C}_m} A)\boldsymbol{x}_0).$$

Moreover there holds $\dim \mathcal{C}_m = m$ and $\dim \mathcal{Z}_\ell = \ell$. All relevant quantities are defined directly in the algorithm or by referencing to previous stated equations.

If we refer to $\mathcal{C}_m$, $\mathcal{Z}_\ell$ and derived quantities in the next section we assume, that they are constructed by Algorithm 3.4.11. The breakdown index L denotes the next breakdown, i. e., $L \geq m$. Moreover there holds $\mathcal{C}_m \cap \mathcal{N}(A) = \{\boldsymbol{0}\}$ for $m < L$ and the results of this section apply for $m = L$. After a singular breakdown we may proceed with a look-ahead as long as there are unused directions in the global correction space.

Algorithm 3.4.11. General MR method for a possibly singular operator and an arbitrary correction space

Given: $A, \boldsymbol{b}, \boldsymbol{x}_0, \tilde{\mathcal{C}}$
$m := 0$, $\ell := 0$, $\boldsymbol{r}_0 := \boldsymbol{b} - A\boldsymbol{x}_0$, $\beta := \|\boldsymbol{r}_0\|$, $\boldsymbol{v}_1 := \boldsymbol{r}_0/\beta$, $V_1 := \begin{bmatrix}\boldsymbol{v}_1\end{bmatrix}$
Assume empty matrices Z_0, C_0, Q_0 and $\tilde{H}_0$
while $\dim \tilde{\mathcal{C}} > m + \ell$
    $\mathcal{C}_m = \operatorname{span}\{C_m\}, \quad \mathcal{Z}_\ell = \operatorname{span}\{Z_\ell\}, \quad \mathcal{V}_{m+1} = \operatorname{span}\{V_{m+1}\}$
    Choose $\tilde{\boldsymbol{c}} \in \tilde{\mathcal{C}} \setminus \mathcal{C}_m \oplus \mathcal{Z}_\ell$
    $\tilde{\boldsymbol{h}} := \left[(A\tilde{\boldsymbol{c}}, \boldsymbol{v}_j)\right]_{j=1}^{m+1}$
    $\tilde{\boldsymbol{v}} := (I - P_{\mathcal{V}_{m+1}})A\tilde{\boldsymbol{c}}, \qquad \tilde{\eta} := \|\tilde{\boldsymbol{v}}\|$
    $\begin{bmatrix}\boldsymbol{t} \\ \tau\end{bmatrix} := Q_m\tilde{\boldsymbol{h}} \quad (\tau := (A\tilde{\boldsymbol{c}}, \boldsymbol{v}_1)$ for $m = 0)$
    if $\tilde{\eta} \neq 0$
        $m := m + 1$
        Update the Arnoldi-type decomposition:
        $C_m := \begin{bmatrix}C_{m-1} & \tilde{\boldsymbol{c}}\end{bmatrix}, \quad V_{m+1} := \begin{bmatrix}V_m & \tilde{\boldsymbol{v}}\end{bmatrix}, \quad \tilde{H}_m := \begin{bmatrix}\tilde{H}_{m-1} & \tilde{\boldsymbol{h}} \\ \boldsymbol{0}^\top & \tilde{\eta}\end{bmatrix}$
        Compute Q_m and R_m according to (3.20), (3.18) and (3.21).
        Compute $\boldsymbol{c}_m^{\mathrm{MR}}$ using (3.16) and (3.23).
    else
        if $\tau = 0$ (singular breakdown, try look-ahead)
            $\boldsymbol{g} := {R_m}^{-1}\boldsymbol{t}$
            $\boldsymbol{z} := C_m\boldsymbol{g} - \tilde{\boldsymbol{c}} \quad (\boldsymbol{z} := \tilde{\boldsymbol{c}}$ for $m = 0)$
            $\tilde{\boldsymbol{z}} = (I - P_{\mathcal{Z}_\ell})\boldsymbol{z}/\|(I - P_{\mathcal{Z}_\ell})\boldsymbol{z}\|$
            $Z_{\ell+1} := \begin{bmatrix}Z_\ell & \tilde{\boldsymbol{z}}\end{bmatrix}$
            $\ell := \ell + 1$
        else (regular breakdown)
            $m := m + 1$
            $C_m := \begin{bmatrix}C_{m-1} & \tilde{\boldsymbol{c}}\end{bmatrix}, \quad H_m := \begin{bmatrix}\tilde{H}_{m-1} & \tilde{\boldsymbol{h}}\end{bmatrix}, \quad R_m := \begin{bmatrix}R_{m-1} & \boldsymbol{t} \\ \boldsymbol{0}^\top & \tau\end{bmatrix}$
            $\boldsymbol{c}_m^{\mathrm{MR}} := \beta C_m {R_m}^{-1} Q_{m-1}\boldsymbol{u}_1^{(m)}$
            break
        end
    end
    Compute the subspace inverse solution with respect to $\mathcal{C}_m \oplus \mathcal{Z}_\ell$ by:
    $\boldsymbol{x}_{\min}^{\mathrm{MR}} := (I - P_{\mathcal{Z}_\ell})(\boldsymbol{x}_0 + \boldsymbol{c}_m^{\mathrm{MR}})$ where $P_{\mathcal{Z}_\ell} = Z_\ell Z_\ell^*$
end

3.5 OR Corrections and the Connections with Angles

We recall the definitions and some facts about angles between two vectors, between a vector and a subspace and canonical angles between subspaces. Proofs of the Lemmata 3.5.2 and 3.5.3 and Proposition 3.5.4 can be found in [71] and [22].

Definition 3.5.1. Given two nonzero vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{H}$ the angle $\angle(\boldsymbol{x}, \boldsymbol{y}) \in [0, \pi/2]$ is defined by the relation

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

We define the angle between a nonzero vector $\boldsymbol{x} \in \mathcal{H}$ and a nontrivial subspace $\mathcal{U} \subset \mathcal{H}$ as

$$\angle(\boldsymbol{x}, \mathcal{U}) := \inf_{\boldsymbol{0} \neq \boldsymbol{u} \in \mathcal{U}} \angle(\boldsymbol{x}, \boldsymbol{u}), \quad \text{i.e.,} \quad \cos \angle(\boldsymbol{x}, \mathcal{U}) = \sup_{\boldsymbol{0} \neq \boldsymbol{u} \in \mathcal{U}} \cos \angle(\boldsymbol{x}, \boldsymbol{u}).$$

Given two finite dimensional subspaces $\mathcal{V}$ and $\mathcal{W}$ of $\mathcal{H}$ with $m := \min(\dim \mathcal{V}, \dim \mathcal{W})$, the *canonical* (or *principal*) angles $\{\theta_j\}_{j=1}^m$ between $\mathcal{V}$ and $\mathcal{W}$ are recursively defined by

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

where $\boldsymbol{v} \perp \boldsymbol{v}_1, \dots \boldsymbol{v}_{j-1}$ and $\boldsymbol{w} \perp \boldsymbol{w}_1, \dots \boldsymbol{w}_{j-1}$. *The* angle between the spaces $\mathcal{V}$ and $\mathcal{W}$ is defined as the largest canonical angle

$$\angle(\mathcal{V}, \mathcal{W}) := \theta_m$$

(cf. [30, Section 12.4.3] and [22, p. 259]).

Lemma 3.5.2. *Let $\mathcal{U}$ be a finite dimensional subspace of $\mathcal{H}$ and denote by $P_{\mathcal{U}}$ the orthogonal projector onto $\mathcal{U}$. For each $\boldsymbol{x} \in \mathcal{H}$ there holds*

$$\angle(\boldsymbol{x}, \mathcal{U}) = \angle(\boldsymbol{x}, P_{\mathcal{U}} \boldsymbol{x}) \tag{3.37}$$

and, with $\sin \angle(\boldsymbol{x}, \mathcal{U}) := \sqrt{1 - \cos^2 \angle(\boldsymbol{x}, \mathcal{U})}$,

$$\|P_{\mathcal{U}} \boldsymbol{x}\| = \|\boldsymbol{x}\| \cos \angle(\boldsymbol{x}, \mathcal{U}), \tag{3.38}$$

$$\|(I - P_{\mathcal{U}}) \boldsymbol{x}\| = \|\boldsymbol{x}\| \sin \angle(\boldsymbol{x}, \mathcal{U}). \tag{3.39}$$

Following [22], we define the mth *OR approximation* by the Galerkin condition

$$\boldsymbol{h}_m^{\mathrm{OR}} \in \mathcal{W}_m, \qquad \boldsymbol{r}_m^{\mathrm{OR}} := \boldsymbol{r}_0 - \boldsymbol{h}_m^{\mathrm{OR}} \perp \mathcal{V}_m, \tag{3.40}$$

i.e., we require that the mth approximation error $\boldsymbol{r}_m^{\mathrm{OR}} \in \mathcal{V}_{m+1}$ is orthogonal to the previous error space $\mathcal{V}_m$ (cf. (3.10)).

We summarize existence and uniqueness of such an approximation:

Lemma 3.5.3. *Let $\mathcal{V}, \mathcal{W}$ be subspaces of the Hilbert space $\mathcal{H}$ and $\boldsymbol{r} \in \mathcal{H}$. There exists $\boldsymbol{h} \in \mathcal{W}$ such that $\boldsymbol{r} - \boldsymbol{h} \perp \mathcal{V}$ if and only if $\boldsymbol{r} \in \mathcal{W} + \mathcal{V}^\perp$. Such an $\boldsymbol{h}$ is uniquely determined if and only if $\mathcal{W} \cap \mathcal{V}^\perp = \{\boldsymbol{0}\}$.*

The following statements are equivalent:

(i) *The oblique projection $P_{\mathcal{W}}^{\mathcal{V}}$ exists.*

(ii) $\mathcal{H} = \mathcal{W} \oplus \mathcal{V}^\perp$.

(iii) $\dim \mathcal{V} = \dim \mathcal{W}$ *and* $\angle(\mathcal{V}, \mathcal{W}) < \pi/2$.

(iv) *A unique OR approximation exists and can be written as $\boldsymbol{h}^{\mathrm{OR}} = P_{\mathcal{W}}^{\mathcal{V}} \boldsymbol{r}$.*

The situation, when the OR approximation fails to exist, is referred to as a *Galerkin breakdown* in the literature on Krylov subspace methods.

As mentioned in Section 3.1, the abstract MR and OR approximation problems do not depend on the operator equation (3.1). Therefore we can restate here many results from [22, Sections 2 and 3] regardless of whether A is regular or not. Merely the occurrence of a singular breakdown has to be considered with greater care.

At first we cite some results which are independent of the nested structure of the approximation and residual spaces. Using the notation of Section 3.2 and Lemma 3.5.3 we get

Proposition 3.5.4. *The error of the MR approximation can be expressed as*

$$\|\boldsymbol{r}^{\mathrm{MR}}\| = \|\boldsymbol{r} - \boldsymbol{h}^{\mathrm{MR}}\| = \|(I - \mathcal{W})\boldsymbol{r}\| = \|\boldsymbol{r}\| \sin \angle(\boldsymbol{r}, \mathcal{W}). \tag{3.41}$$

If $\mathcal{H} = \mathcal{W} \oplus \mathcal{V}^\perp$, then there holds

$$P_{\mathcal{W}}^{\mathcal{V}} = (P_{\mathcal{V}} P_{\mathcal{W}})^\dagger, \qquad \|I - P_{\mathcal{W}}^{\mathcal{V}}\| = \frac{1}{\cos \angle(\mathcal{V}, \mathcal{W})} \qquad \textit{and}$$

$$\|\boldsymbol{r}^{\mathrm{OR}}\| = \|\boldsymbol{r} - \boldsymbol{h}^{\mathrm{OR}}\| = \|(I - P_{\mathcal{W}}^{\mathcal{V}})\boldsymbol{r}\| \leq \frac{\|\boldsymbol{r}\|}{\cos \angle(\mathcal{V}, \mathcal{W})}.$$

Applying the notation and results of Section 3.3 (especially the Arnoldi-type decomposition) yields in a representation of the OR correction $\boldsymbol{c}_m^{\mathrm{OR}}$, with $\boldsymbol{h}_m^{\mathrm{OR}} = A\boldsymbol{c}_m^{\mathrm{OR}}$, in the coordinate space:

Proposition 3.5.5. *Let $H_m \in \mathbb{C}^{m \times m}$ denote the square Hessenberg matrix obtained by deleting the last row of $\tilde{H}_m$ as in* (3.14)*. There exists a uniquely determined mth OR approximation if and only if H_m is nonsingular. The correction $\boldsymbol{c}_m^{\mathrm{OR}}$ which corresponds to the approximation $\boldsymbol{h}_m^{\mathrm{OR}} = A\boldsymbol{c}_m^{\mathrm{OR}}$ satisfies*

$$\boldsymbol{c}_m^{\mathrm{OR}} = C_m \boldsymbol{y}_m^{\mathrm{OR}} \qquad \textit{with} \qquad \boldsymbol{y}_m^{\mathrm{OR}} = \beta H_m^{-1} \boldsymbol{u}_1^{(m)}.$$

The following corollary could be also derived noting that $\mathcal{W}_L = \mathcal{V}_L$, i. e., $P_{\mathcal{W}_L}^{\mathcal{V}_L} = P_{\mathcal{W}_L}$, respectively $\mathcal{W}_L = \mathcal{W}_{L-1} \subset \mathcal{V}_L$ (cf. Propositions 3.4.3 and 3.4.4).

Corollary 3.5.6. *If a regular breakdown occurs, the OR approximation and its related correction are uniquely determined and coincide with the corresponding MR quantities. If a singular breakdown occurs the OR approximation does not exists. In any case there holds* $\angle(\mathcal{V}_L, \mathcal{W}_L) = 0$.

The close connection between MR and OR quantities is not restricted to the termination step as we will see in the remainder of this section.

For $m \leq L$ we define $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$ such that $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ is an orthonormal basis of $\mathcal{W}_m$ and $\operatorname{span}\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{m-1}\} = \mathcal{W}_{m-1}$. For notational convenience we set $\boldsymbol{w}_L := 0$ if $\dim \mathcal{W}_L = L - 1$ (i. e., if a singular breakdown occurs). In terms of these vectors, the MR approximation can be expressed as a truncated Fourier expansion

$$\boldsymbol{h}_m^{\mathrm{MR}} = P_{\mathcal{W}_m} \boldsymbol{r}_0 = \sum_{j=1}^{m} (\boldsymbol{r}_0, \boldsymbol{w}_j) \boldsymbol{w}_j.$$

The norm of the approximation error, i. e., the residual $\boldsymbol{r}_m^{\mathrm{MR}} = \boldsymbol{r}_0 - \boldsymbol{h}_m^{\mathrm{MR}}$ is given by

$$\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2 = \|(I - P_{\mathcal{W}_m})\boldsymbol{r}_0\|^2 = \|\boldsymbol{r}_0\|^2 - \sum_{j=1}^{m} |(\boldsymbol{r}_0, \boldsymbol{w}_j)|^2.$$

The mth MR approximation can also be computed by updating the previous approximation:

$$\begin{aligned} \boldsymbol{h}_m^{\mathrm{MR}} &= \sum_{j=1}^{m} (\boldsymbol{r}_0, \boldsymbol{w}_j) \boldsymbol{w}_j \\ &= \boldsymbol{h}_{m-1}^{\mathrm{MR}} + (\boldsymbol{r}_0, \boldsymbol{w}_m) \boldsymbol{w}_m = \boldsymbol{h}_{m-1}^{\mathrm{MR}} + P_{\mathcal{W}_m} \boldsymbol{r}_0 - P_{\mathcal{W}_{m-1}} \boldsymbol{r}_0 \\ &= \boldsymbol{h}_{m-1}^{\mathrm{MR}} + P_{\mathcal{W}_m} (\boldsymbol{r}_0 - \boldsymbol{h}_{m-1}^{\mathrm{MR}}) = \boldsymbol{h}_{m-1}^{\mathrm{MR}} + P_{\mathcal{W}_m} \boldsymbol{r}_{m-1}^{\mathrm{MR}}. \end{aligned}$$

For the residuals it follows that

$$\boldsymbol{r}_m^{\mathrm{MR}} = \boldsymbol{r}_{m-1}^{\mathrm{MR}} - P_{\mathcal{W}_m} \boldsymbol{r}_{m-1}^{\mathrm{MR}} = (I - P_{\mathcal{W}_m}) \boldsymbol{r}_{m-1}^{\mathrm{MR}} \tag{3.42}$$

and, using (3.39),

$$\|\boldsymbol{r}_m^{\mathrm{MR}}\| = \|(I - P_{\mathcal{W}_m}) \boldsymbol{r}_{m-1}^{\mathrm{MR}}\| = \sin \angle(\boldsymbol{r}_{m-1}^{\mathrm{MR}}, \mathcal{W}_m) \|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|. \tag{3.43}$$

The angle which occurs in (3.43) is strongly connected with the QR decomposition (3.22), as the next proposition shows.

Proposition 3.5.7. *Suppose that no singular breakdown occurs in step m ($m \leq L$) and denote by $Q_m = [q_{j,k}^{(m)}]_{j,k=1}^{m+1}$ the uniquely determined unitary matrix in the QR decomposition (3.22) of the Hessenberg matrix $\tilde{H}_m = \left[\eta_{j,k}\right]_{j,k=1}^{m+1,m}$. Then there holds*

$$\sin \angle(\boldsymbol{r}_0, \mathcal{W}_m) = |q_{m+1,1}^{(m)}| \tag{3.44}$$

$$\sin \angle(\boldsymbol{r}_{m-1}^{\mathrm{MR}}, \mathcal{W}_m) = \frac{|q_{m+1,1}^{(m)}|}{|q_{m,1}^{(m-1)}|} = \frac{\eta_{m+1,m}}{\sqrt{|\tau|^2 + \eta_{m+1,m}^2}}, \tag{3.45}$$

where τ is defined by $\begin{bmatrix} \boldsymbol{t} \\ \tau \end{bmatrix} = Q_{m-1}\boldsymbol{h}_m$ and $\boldsymbol{h}_m = [\eta_{j,m}]_{j=1}^{m}$ is the last column of H_m.

Proof. With respect to the orthonormal basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m+1}\}$ of $\mathcal{V}_{m+1}$ the vector $\boldsymbol{r}_0 = \beta\boldsymbol{v}_1$ possesses the coordinates $\beta\boldsymbol{u}_1^{(m+1)}$ and $\mathcal{W}_m = A\mathcal{C}_m$ is represented by $\mathcal{R}(\tilde{H}_m)$. This implies

$$\angle(\boldsymbol{r}_0, \mathcal{W}_m) = \angle(\boldsymbol{v}_1, \mathcal{W}_m) = \angle_2(\boldsymbol{u}_1^{(m+1)}, \mathcal{R}(\tilde{H}_m)),$$

where the last angle is defined with respect to the Euclidean inner product on $\mathbb{C}^{m+1}$, which is indicated by the subscript 2.

Obviously, the columns of $V_{m+1}Q_m^{\mathrm{H}}$ form another orthonormal basis of $\mathcal{V}_{m+1}$ (cf. (3.24)). The coordinate vector of $\boldsymbol{v}_1$ with respect to this basis is $Q_m\boldsymbol{u}_1^{(m+1)}$, the first column of Q_m, and $\mathcal{W}_m$ is represented by all vectors in $\mathbb{C}^{m+1}$ with zero last component, a subspace which we identify with $\mathbb{C}^m$. The orthogonal projection of $\mathbb{C}^{m+1}$ onto $\mathbb{C}^m$ is given by $\begin{bmatrix} I_m & \boldsymbol{0} \\ \boldsymbol{0}^\top & 0 \end{bmatrix}$ and since $Q_m\boldsymbol{u}_1^{(m+1)}$ has unit length we get from (3.39)

$$\begin{aligned} \sin \angle(\boldsymbol{r}_0, \mathcal{W}_m) = \sin \angle_2(Q_m\boldsymbol{u}_1^{(m+1)}, \mathbb{C}^m) = \cdots \\ \cdots = \frac{\left\| \left(I_{m+1} - \begin{bmatrix} I_m & \boldsymbol{0} \\ \boldsymbol{0}^\top & 0 \end{bmatrix} \right) Q_m\boldsymbol{u}_1^{(m+1)} \right\|_2}{\left\| Q_m\boldsymbol{u}_1^{(m+1)} \right\|_2} = |q_{m+1,1}^{(m)}|, \end{aligned}$$

which proves (3.44).

Equations (3.43) and (3.41) can be combined to

$$\sin \angle(\boldsymbol{r}_{m-1}^{\mathrm{MR}}, \mathcal{W}_m) = \frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|}{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|} = \frac{\|(I - P_{\mathcal{W}_m})\boldsymbol{r}_0\|}{\|(I - P_{\mathcal{W}_{m-1}})\boldsymbol{r}_0\|} = \frac{\sin \angle(\boldsymbol{r}_0, \mathcal{W}_m)}{\sin \angle(\boldsymbol{r}_0, \mathcal{W}_{m-1})}, \tag{3.46}$$

where we assume that $\boldsymbol{r}_0 \notin \mathcal{W}_{m-1}$, i.e. no regular breakdown occurs before step m. If a regular breakdown occurs in step $m = L$ then $\angle(\boldsymbol{r}_0, \mathcal{W}_m) = 0$, since $\boldsymbol{r}_0 \in \mathcal{W}_m$ and $\boldsymbol{r}_L^{\mathrm{MR}} = \boldsymbol{0}$, i.e., the quantities in (3.44) and (3.45) become zero.

From (3.18) and (3.20) we conclude $q_{m+1,1}^{(m)} = -s_m e^{i\phi_m} q_{m,1}^{(m-1)}$, where s_m and ϕ_m are as defined in (3.21). Together this yields (3.45). □

Since $\boldsymbol{r}_m^{\mathrm{MR}} \perp P_{\mathcal{W}_m}\boldsymbol{r}_{m-1}^{\mathrm{MR}} \in \mathcal{W}_m$ we can invoke Pythagoras' theorem in (3.42):

$$\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2 = \|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2 - \|P_{\mathcal{W}_m}\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2 = \|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2 - |(\boldsymbol{r}_0, \boldsymbol{w}_m)|^2. \tag{3.47}$$

This shows that the MR approximation is not improved whenever the direction, in which $\mathcal{W}_{m-1}$ is enlarged, is orthogonal to $\boldsymbol{r}_0$, or if $\mathcal{W}_{m-1}$ is not enlarged at all because of a singular breakdown in the mth step (remember the convention $\boldsymbol{w}_m = \boldsymbol{0}$ in this case). Using (3.47) the cosines corresponding to (3.46) can be expressed as

$$c_m = \sqrt{1 - s_m^2} = \sqrt{1 - \frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2}{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|}} = \frac{|(\boldsymbol{w}_m, \boldsymbol{r}_0)|}{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|}, \tag{3.48}$$

which shows, that the MR approximation improves in the mth step if and only if $c_m \neq 0$.

Next, we further investigate the question of when the OR approximation is well-defined, i. e., if $\cos\angle(\mathcal{V}_m, \mathcal{W}_m) \neq 0$ and $\dim \mathcal{W}_m = m$. Noting that the set $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{m-1}, \hat{\boldsymbol{w}}_m\}$ with $\hat{\boldsymbol{w}}_m := \boldsymbol{r}_{m-1}^{\mathrm{MR}}/\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|$ is an orthonormal basis of $\mathcal{V}_m$ we see, that the cosines of the canonical angles between $\mathcal{V}_m$ and $\mathcal{W}_m$ are the singular values of the matrix (cf. [30])[4]

$$\begin{bmatrix} |(\boldsymbol{w}_1, \boldsymbol{w}_1)| & \ldots & |(\boldsymbol{w}_{m-1}, \boldsymbol{w}_1)| & |(\boldsymbol{w}_m, \boldsymbol{w}_1)| \\ \vdots & & \vdots & \vdots \\ |(\boldsymbol{w}_1, \boldsymbol{w}_{m-1})| & \ldots & |(\boldsymbol{w}_{m-1}, \boldsymbol{w}_{m-1})| & |(\boldsymbol{w}_m, \boldsymbol{w}_{m-1})| \\ |(\boldsymbol{w}_1, \hat{\boldsymbol{w}}_m)| & \ldots & |(\boldsymbol{w}_{m-1}, \hat{\boldsymbol{w}}_m)| & |(\boldsymbol{w}_m, \hat{\boldsymbol{w}}_m)| \end{bmatrix} = \begin{bmatrix} I & \boldsymbol{0} \\ \boldsymbol{0}^\top & |(\boldsymbol{w}_m, \hat{\boldsymbol{w}}_m)| \end{bmatrix}.$$

The smallest singular value is

$$|(\boldsymbol{w}_m, \hat{\boldsymbol{w}}_m)| = \frac{|(\boldsymbol{w}_m, \boldsymbol{r}_{m-1}^{\mathrm{MR}})|}{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|} = \frac{|(\boldsymbol{w}_m, \boldsymbol{r}_0 - P_{\mathcal{W}_{m-1}}\boldsymbol{r}_0)|}{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|} = \frac{|(\boldsymbol{w}_m, \boldsymbol{r}_0)|}{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|} = c_m.$$

We summarize this considerations

Proposition 3.5.8. *If no singular breakdown occurs in the mth step, we have*

$$\angle(\mathcal{V}_m, \mathcal{W}_m) = \angle(\boldsymbol{r}_{m-1}^{\mathrm{MR}}, \mathcal{W}_m).$$

In the exceptional case of a singular breakdown the angle between $\mathcal{V}_L$ and $\mathcal{W}_L$ happens to be zero whereas $\boldsymbol{r}_{L-1}^{\mathrm{MR}}$ is orthogonal to $\mathcal{W}_L = \mathcal{W}_{L-1}$. As remarked earlier regarding (3.32) we can arbitrarily define the last Givens rotation G_L in this case. Taking the identity relates G_L to the $\angle(\mathcal{V}_L, \mathcal{W}_L)$ whereas

$$G_L = \begin{bmatrix} I_{L-1} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0}^\top & 0 & 1 \\ \boldsymbol{0}^\top & -1 & 0 \end{bmatrix}$$

[4] Note, that we use a slightly different definition for the canonical angles then in [30], since we take the absolute value of the inner products.

is the Givens rotation belonging to $\angle(\boldsymbol{r}_{L-1}^{\mathrm{MR}}, \mathcal{W}_L)$. But, in view of Corollary 3.5.6 and Proposition 3.4.5, the following consequence of Proposition 3.5.8 holds also true for a singular breakdown in step $m = L$.

Corollary 3.5.9. *The OR approximation of $\boldsymbol{r}_0$ with respect to $\mathcal{V}_m$ and $\mathcal{W}_m$ exists and is uniquely determined if and only if the corresponding MR approximation improves, i. e., if and only if $(\boldsymbol{w}_m, \boldsymbol{r}_0) \neq 0$.*

Remark 3.5.10. We can apply this corollary to the special case that the MR approximation detects a least squares solution of (3.1), which is, in view of Lemma 3.2.6, equivalent to $P_{\mathcal{R}(A)}\boldsymbol{r}_0 \in \mathcal{W}_m$:

If (3.1) is consistent, i. e., $P_{\mathcal{R}(A)}\boldsymbol{r}_0 = \boldsymbol{r}_0$, we conclude from (ii), Proposition 3.4.3, that a regular breakdown occurs in step m.

Suppose now, that $\boldsymbol{r}_0 \notin \mathcal{R}(A)$, i. e., the linear equation (3.1) is inconsistent, and $P_{\mathcal{R}(A)}\boldsymbol{r}_0 \in \mathcal{W}_m$. Then a singular breakdown occurs in step $m \leq L$. Since the MR approximation becomes the global best approximation over $\mathcal{H}$, it can not improve in the subsequent steps $m+1, m+2, \ldots$ and the corresponding OR approximations fail to exist. This holds true, even if we proceed the approximation process after the singular breakdown using the look-ahead strategy described in Remark 3.4.9. Moreover, equation (viii) in Proposition 3.4.4 guarantees $P_{\mathcal{R}(A)}\boldsymbol{r}_0 \in \mathcal{W}_{L-1}$, i. e., there is at least one iteration with stagnation before the singular breakdown.

The question of when an MR method determines a least squares solution can be further investigated. The next proposition generalizes a result which was stated by Brown and Walker in [7] for Krylov spaces, and can be analogously proved for general correction spaces.

Proposition 3.5.11. *A MR method determines a least squares solution without singular breakdown in the mth step if and only if*

$$\dim A^*\mathcal{V}_{m+1} = \dim \mathcal{W}_m = m.$$

Proof. By (ix) of Proposition 3.4.4 we have $\dim \mathcal{W}_m = m$ if and only if no singular breakdown occurs in step m. A MR method determines a least square solution if and only if there exists a correction $\boldsymbol{c} \in \mathcal{C}_m$ such that $\boldsymbol{x}_0 + \boldsymbol{c}$ solves the *normal equation* (see, e. g., [14, Theorem 2.1.2])

$$\boldsymbol{0} = A^*(\boldsymbol{b} - A(\boldsymbol{x}_0 + \boldsymbol{c})) = A^*\boldsymbol{r}_0 - A^*A\boldsymbol{c}.$$

This holds true if and only if $A^*\boldsymbol{r}_0 \in A^*\mathcal{W}_m$ or, equivalently, if and only if $A^*\mathcal{V}_{m+1} = A^*\boldsymbol{r}_0 + A^*\mathcal{W}_m = A^*\mathcal{W}_m$.

Finally, we show that $\dim \mathcal{W}_m = \dim A^*\mathcal{W}_m$: Clearly we have $\dim A^*\mathcal{W}_m \leq \dim \mathcal{W}_m$. Suppose $\dim A^*\mathcal{W}_m < \dim \mathcal{W}_m$, i. e., there exists $\boldsymbol{c} \in \mathcal{C}_m$ such that $\boldsymbol{0} \neq A\boldsymbol{c} \in \mathcal{W}_m$ and $A^*A\boldsymbol{c} = \boldsymbol{0}$. Then $0 = (\boldsymbol{c}, A^*A\boldsymbol{c}) = (A\boldsymbol{c}, A\boldsymbol{c}) = \|A\boldsymbol{c}\|^2$, which is a contradiction. □

We close this section with a theorem which summarizes the connections between MR and OR approximations and solutions.

If $(\boldsymbol{w}_m, \boldsymbol{r}_0) \neq 0$ we can define $\tilde{\boldsymbol{w}}_m = \boldsymbol{r}_{m-1}^{\mathrm{MR}}/(\boldsymbol{w}_m, \boldsymbol{r}_0)$, which implies $(\tilde{\boldsymbol{w}}_m, \boldsymbol{w}_m) = 1$ and $(\tilde{\boldsymbol{w}}_m, \boldsymbol{w}_j) = 0$ $(j = 1, \ldots, m-1)$. Using this definition we can express the oblique OR projection by the expansion

$$P_{\mathcal{W}_m}^{\mathcal{V}_m} = \sum_{j=1}^{m-1} (\cdot, \boldsymbol{w}_j)\boldsymbol{w}_j + (\cdot, \tilde{\boldsymbol{w}}_m)\boldsymbol{w}_m.$$

For the difference of the OR and MR approximations we derive

$$\begin{aligned} \boldsymbol{h}_m^{\mathrm{OR}} - \boldsymbol{h}_m^{\mathrm{MR}} = (P_{\mathcal{W}_m}^{\mathcal{V}_m} - P_{\mathcal{W}_m})\boldsymbol{r}_0 &= \left[\frac{(\boldsymbol{r}_0, \boldsymbol{r}_{m-1}^{\mathrm{MR}})}{(\boldsymbol{w}_m, \boldsymbol{r}_0)} - (\boldsymbol{r}_0, \boldsymbol{w}_m)\right] \boldsymbol{w}_m \\ &= \frac{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2 - |(\boldsymbol{r}_0, \boldsymbol{w}_m)|^2}{(\boldsymbol{w}_m, \boldsymbol{r}_0)} \boldsymbol{w}_m = \frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2}{(\boldsymbol{w}_m, \boldsymbol{r}_0)} \boldsymbol{w}_m, \end{aligned} \tag{3.49}$$

which is the key to prove the following relations:

Theorem 3.5.12. *With the notation of Algorithm* 3.4.11, $s_m = \sin\angle(\boldsymbol{r}_{m-1}^{\mathrm{MR}}, \mathcal{W}_m)$ *and* $c_m = \cos\angle(\boldsymbol{r}_{m-1}^{\mathrm{MR}}, \mathcal{W}_m)$ *the MR approximations satisfy*

$$\|\boldsymbol{r}_m^{\mathrm{MR}}\| = s_m\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\| = s_1 s_2 \cdots s_m \|\boldsymbol{r}_0\|, \tag{3.50}$$

If the mth OR approximation exists, it satisfies

$$\|\boldsymbol{r}_m^{\mathrm{MR}}\| = c_m\|\boldsymbol{r}_m^{\mathrm{OR}}\|, \tag{3.51}$$

$$\|\boldsymbol{r}_m^{\mathrm{OR}}\| = s_1 s_2 \cdots s_m \|\boldsymbol{r}_0\|/c_m, \tag{3.52}$$

$$\boldsymbol{x}_m^{\mathrm{MR}} = s_m^2 \boldsymbol{x}_{m-1}^{\mathrm{MR}} + c_m^2 \boldsymbol{x}_m^{\mathrm{OR}} + \boldsymbol{z}, \qquad \text{with } \boldsymbol{z} \in \mathcal{Z}_\ell \tag{3.53}$$

$$\boldsymbol{h}_m^{\mathrm{MR}} = s_m^2 \boldsymbol{h}_{m-1}^{\mathrm{MR}} + c_m^2 \boldsymbol{h}_m^{\mathrm{OR}} \tag{3.54}$$

$$\boldsymbol{r}_m^{\mathrm{MR}} = s_m^2 \boldsymbol{r}_{m-1}^{\mathrm{MR}} + c_m^2 \boldsymbol{r}_m^{\mathrm{OR}}. \tag{3.55}$$

Moreover, if no breakdown occurs in the mth step, there holds

$$\frac{1}{\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2} \boldsymbol{x}_m^{\mathrm{MR}} = \sum_{j=0}^{m}{}' \frac{1}{\|\boldsymbol{r}_j^{\mathrm{OR}}\|^2} \boldsymbol{x}_j^{\mathrm{OR}} + \boldsymbol{z}, \qquad \text{with } \boldsymbol{z} \in \mathcal{Z}_\ell \tag{3.56}$$

$$\frac{1}{\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2} \boldsymbol{h}_m^{\mathrm{MR}} = \sum_{j=0}^{m}{}' \frac{1}{\|\boldsymbol{r}_j^{\mathrm{OR}}\|^2} \boldsymbol{h}_j^{\mathrm{OR}} \tag{3.57}$$

$$\frac{1}{\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2} \boldsymbol{r}_m^{\mathrm{MR}} = \sum_{j=0}^{m}{}' \frac{1}{\|\boldsymbol{r}_j^{\mathrm{OR}}\|^2} \boldsymbol{r}_j^{\mathrm{OR}}, \tag{3.58}$$

$$\frac{1}{\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2} = \frac{1}{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2} + \frac{1}{\|\boldsymbol{r}_m^{\mathrm{OR}}\|^2} = \sum_{j=0}^{m}{}' \frac{1}{\|\boldsymbol{r}_j^{\mathrm{OR}}\|^2}, \tag{3.59}$$

where the sums $\sum'$ *denote a summation over all* j *for which* $c_j \neq 0$ *(i. e., for which the OR approximation exists).*

Proof. The identity (3.50) follows from an obvious induction applied to (3.46).

From $\boldsymbol{r}_m^{\mathrm{MR}} - \boldsymbol{r}_m^{\mathrm{OR}} = \boldsymbol{h}_m^{\mathrm{OR}} - \boldsymbol{h}_m^{\mathrm{MR}}$ and (3.49) there follows

$$\boldsymbol{r}_m^{\mathrm{MR}} = \boldsymbol{r}_m^{\mathrm{OR}} + \frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2}{(\boldsymbol{w}_m, \boldsymbol{r}_0)} \boldsymbol{w}_m.$$

Since $\boldsymbol{r}_m^{\mathrm{MR}} \perp \boldsymbol{w}_m$, the Pythagorean identity yields

$$\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2 = \left(1 + \frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2}{(\boldsymbol{w}_m, \boldsymbol{r}_0)}\right) \|\boldsymbol{r}_m^{\mathrm{OR}}\|^2 = \frac{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2}{(\boldsymbol{w}_m, \boldsymbol{r}_0)} \|\boldsymbol{r}_m^{\mathrm{OR}}\|^2,$$

where we used (3.47) for the last equality. In view of (3.48), this proves (3.51) and (3.52).

Using $\boldsymbol{h}_m^{\mathrm{MR}} - \boldsymbol{h}_{m-1}^{\mathrm{MR}}$, (3.48), (3.46) and again (3.49) we obtain

$$\begin{aligned}
\boldsymbol{h}_m^{\mathrm{OR}} &= \boldsymbol{h}_m^{\mathrm{MR}} + \frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2}{(\boldsymbol{w}_m, \boldsymbol{r}_0)} \frac{1}{(\boldsymbol{r}_0, \boldsymbol{w}_m)} (\boldsymbol{h}_m^{\mathrm{MR}} - \boldsymbol{h}_{m-1}^{\mathrm{MR}}) \\
&= \boldsymbol{h}_m^{\mathrm{MR}} + \frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2}{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2} \frac{\|\boldsymbol{r}_{m-1}^{\mathrm{MR}}\|^2}{|(\boldsymbol{r}_0, \boldsymbol{w}_m)|^2} (\boldsymbol{h}_m^{\mathrm{MR}} - \boldsymbol{h}_{m-1}^{\mathrm{MR}}) \\
&= \boldsymbol{h}_m^{\mathrm{MR}} + \frac{s_m^2}{c_m^2} (\boldsymbol{h}_m^{\mathrm{MR}} - \boldsymbol{h}_{m-1}^{\mathrm{MR}}),
\end{aligned}$$

which implies by $s_m^2 + c_m^2 = 1$the relationships (3.54) and (3.55).

Since $\boldsymbol{h}_m^{\mathrm{MR}} = \boldsymbol{h}_{m-1}^{\mathrm{MR}}$ if the mth OR approximation does not exist, i. e., if $c_m = 0$, the repeated application of (3.54) leads to

$$\boldsymbol{h}_m^{\mathrm{MR}} = \sum_{\substack{j=0 \\ c_j \neq 0}}^{m} \tau_{m,j}^2 \boldsymbol{h}_j^{\mathrm{OR}}, \tag{3.60}$$

where

$$\tau_{m,0} := \prod_{\substack{\ell=0 \\ c_\ell \neq 0}}^{m} s_\ell \qquad \text{and} \qquad \tau_{m,j} := c_j \prod_{\substack{\ell=j+1 \\ c_\ell \neq 0}}^{m} s_\ell \quad (1 \leq j \leq m \text{ and } c_j \neq 0).$$

Using (3.46) and (3.51) the second term can be rewritten as

$$\tau_{m,j} := \frac{\|\boldsymbol{r}_j^{\mathrm{MR}}\|}{\|\boldsymbol{r}_j^{\mathrm{OR}}\|} \prod_{\substack{\ell=j+1 \\ c_\ell \neq 0}}^{m} \frac{\|\boldsymbol{r}_\ell^{\mathrm{MR}}\|}{\|\boldsymbol{r}_{\ell-1}^{\mathrm{MR}}\|} = \frac{\|\boldsymbol{r}_m^{\mathrm{MR}}\|}{\|\boldsymbol{r}_j^{\mathrm{MR}}\|},$$

which obviously holds true also for $j = 0$. As long as no (regular) breakdown occur, we have $\boldsymbol{r}_m^{\mathrm{MR}} \neq 0$ and can divide (3.60) by $\|\boldsymbol{r}_m^{\mathrm{MR}}\|^2$ to get (3.57). Similarly we can derive (3.58) from (3.55).

The first equality in (3.59) is a reformulation of the Pythagoras identity

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

(cf. (3.46) and (3.51)), and the second follows by induction if we take into account that $\boldsymbol{h}_m^{\mathrm{MR}} = \boldsymbol{h}_{m-1}^{\mathrm{MR}}$ if $c_m = 0$.

Any identity for the approximations $\boldsymbol{h}_m$ implies an analogous result for the corrections $\boldsymbol{c}_m$ or iterates $\boldsymbol{x}_m$: Rewriting the identity as a homogeneous equation in $A|_{\mathcal{C}_m \oplus \mathcal{Z}_\ell}$ (cf. (3.4) and Algorithm 3.4.11) shows that the corrections satisfies the identity up to a difference $\boldsymbol{z} \in \mathcal{Z}_\ell$. Adding $\boldsymbol{x}_0$ supplies the results (3.53) and (3.56). □

4 Krylov Subspace Methods

The most common correction spaces for solving (regular) linear systems are *Krylov spaces* (or *Krylov subspaces*) which are defined by

$$\mathcal{K}_m(A, \boldsymbol{r}_0) := \operatorname{span}\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^{m-1}\boldsymbol{r}_0\}. \tag{4.1}$$

The choice $\mathcal{C}_m = \mathcal{K}_m(A, \boldsymbol{r}_0)$ results in the residual space

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

and the decomposition (3.14) becomes a proper Arnoldi decomposition

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

There are several reasons for the popularity of Krylov spaces. First, they can be simply generated by subsequently multiplying a vector with A in each step. Since the correction space coincides with the preceding residual space, we can use the same basis for both in computations, which reduces the computing and storage requirements. Moreover, if A is regular, a MR or OR method determines the exact solution whenever the Krylov space becomes A-invariant (cf. [60] or [22]). We will see below, that this is no longer true for singular A, that is, there exists combinations of right hand side $\boldsymbol{b}$ and initial guess $\boldsymbol{x}_0$ for which the system $A\boldsymbol{x} = \boldsymbol{b}$ is solvable, i. e., $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0 \in \mathcal{R}(A)$, but no Krylov space contains a correction $\boldsymbol{c}$ with $A\boldsymbol{c} = \boldsymbol{r}_0$.

4.1 The Polynomial Structure of Krylov Spaces

Krylov spaces are closely related to polynomials, as can be seen from the representation

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

where $\mathcal{P}_m$ denotes the space of all complex polynomials of degree at most m. The polynomial $\mathcal{P}_m$ space is isomorphic to the mth Krylov space if and only if there exists no nonzero polynomial $q \in \mathcal{P}_{m-1}$ with $q(A)\boldsymbol{r}_0 = \boldsymbol{0}$. If such a polynomial exists for some m then there also exists a (unique) monic polynomial $m_{\boldsymbol{r}_0,A}$ of minimal degree with $m_{\boldsymbol{r}_0,A}(A)\boldsymbol{r}_0 = \boldsymbol{0}$ which is called the *minimal polynomial of* $\boldsymbol{r}_0$ *with respect to* A. Since (3.12) can be rewritten as

$$A^m\boldsymbol{r}_0 \in \operatorname{span}\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^{m-1}\boldsymbol{r}_0\}, \tag{4.2}$$

the degree of the minimal polynomial equals the breakdown index L (cf. (3.13)), which implies, that L is finite. Conversely, if $L < \infty$, we conclude from (4.2) and the minimality property of L (cf. Propositions 3.4.7 and 3.4.8) that there exists a minimal polynomial $m_{\boldsymbol{r}_0,A}$ of degree L. It is easy to see that this index also equals the smallest integer m such that $\mathcal{K}_m(A,\boldsymbol{r}_0) = \mathcal{K}_{m+1}(A,\boldsymbol{r}_0)$, e. g.

$$\begin{aligned} L &= \min\{m \in \mathbb{N}_0 \,:\, \mathcal{K}_m(A,\boldsymbol{r}_0) = \mathcal{K}_{m+1}(A,\boldsymbol{r}_0)\} \\ &= \min\{\deg q \,:\, q \text{ monic and } q(A)\boldsymbol{r}_0 = \boldsymbol{0}\}. \end{aligned} \tag{4.3}$$

Clearly we have $\mathcal{K}_m(A,\boldsymbol{r}_0) = \mathcal{K}_L(A,\boldsymbol{r}_0)$ for all $m \geq L$ and $\dim \mathcal{K}_m(A,\boldsymbol{r}_0) = m$ if and only if $m \leq L$.

This shows that a singular breakdown cannot occur as long as the sequence of Krylov subspaces does not terminate, i. e., $\mathcal{K}_m(A,\boldsymbol{r}_0) \cap \mathcal{N}(A) = \{\boldsymbol{0}\}$ as long as $m < L$. Thus it is not necessary to employ a look-ahead strategy as described in Remark 3.4.9.

A first consequence of the polynomial connection is a representation of the residuals in terms of special polynomials: Since every vector $\boldsymbol{x}_m \in \boldsymbol{x}_0 + \mathcal{K}_m(A,\boldsymbol{r}_0)$ is of the form $\boldsymbol{x}_m = \boldsymbol{x}_0 + \boldsymbol{c}_m = \boldsymbol{x}_0 + q_{m-1}(A)\boldsymbol{r}_0$ with some $q_{m-1} \in \mathcal{P}_{m-1}$, the corresponding residual $\boldsymbol{r}_m = \boldsymbol{b} - A\boldsymbol{x}_m = \boldsymbol{r}_0 - A\boldsymbol{c}_m$ can be written as

$$\boldsymbol{r}_m = \boldsymbol{r}_0 + Aq_{m-1}(A)\boldsymbol{r}_0 = p_m(A)\boldsymbol{r}_0, \qquad \text{where } p_m(\xi) := 1 - \xi q_{m-1}(\xi) \in \mathcal{P}_m. \tag{4.4}$$

The *residual polynomial* p_m satisfies the normalization condition $p_m(0) = 1$.

Remark 4.1.1. . For nonsingular A characterizations of the residual polynomials which belong to the MR and OR iterates as well as their zeros can be found in the literature (cf. [23] and the references therein). The results about the OR polynomials follow by the usual line of argument, which did not care about if A is singular or not.

The zeros of the MR polynomials can be characterized as the orthogonal section of A^{-1} onto $\mathcal{K}_m(A^{-1}, A^m\boldsymbol{r}_0)$ if A is nonsingular. To obtain analogous results for singular A, we have to substitute the inverse by a suitable generalized inverse.

The operator of choice is the subspace inverse with respect to the mth Krylov subspace, as long as no singular breakdown occur. This limitation is not a critical one since a singular breakdown means, that the MR approximation makes no progress in step $m = L$ (cf. Proposition 3.4.5) and thus the MR residual polynomial remains the same as in the previous step.

Using Propositions 3.2.3 and 3.4.4 we get that $A^{\mathcal{C}_m}A$ is a projection *onto* $\mathcal{C}_m = \mathcal{K}_m(A,\boldsymbol{r}_0)$ if no singular breakdown occur. Noting $A^k\boldsymbol{r}_0 \in \mathcal{R}(A^{\mathcal{C}_m}A)$ for $0 \leq k \leq m-1$ this implies

$$\mathcal{K}_m(A^{\mathcal{C}_m}, A^m\boldsymbol{r}_0) = A\mathcal{K}_m(A,\boldsymbol{r}_0) \quad \text{and} \quad \mathcal{K}_{m+1}(A^{\mathcal{C}_m}, A^m\boldsymbol{r}_0) = \mathcal{K}_{m+1}(A,\boldsymbol{r}_0)$$

As a second tool we need an inverse version of the relation $AV_m = V_{m+1}\tilde{H}_m = \hat{V}_m R_m$ (cf. (3.24)). If no singular breakdown occur in step m, this can be rewritten as

$$A^{C_m}\hat{V}_m = V_m {R_m}^{-1}$$

since R_m is nonsingular (cf. Proposition 3.4.4, (v)) and $A^{\mathcal{C}_m}A$ is a projection onto $\mathcal{C}_m = \mathcal{K}_m(A, \boldsymbol{r}_0) = \operatorname{span}\{V_m\}$. This relations are the essentials to realize that the argumentation and results of [23, Section 3.3] remain valid if we replace the inverse of A by the subspace inverse $A^{\mathcal{C}_m}$.

The disadvantage of $A^{\mathcal{C}_m}$ is its dependence on the mth Krylov subspace (and thus it depends on $\boldsymbol{r}_0$). The dependence on m is not significant since we could just as well choose $A^{\mathcal{C}}$ where $\mathcal{C}$ is the largest Krylov subspace with $\mathcal{K}_m(A, \boldsymbol{r}_0) \cap \mathcal{N}(A) = \{\boldsymbol{0}\}$ (that is $\mathcal{K}_{L-1}(A, \boldsymbol{r}_0)$ or $\mathcal{K}_L(A, \boldsymbol{r}_0)$ depending whether a singular breakdown occurs or not). However, the dependence on $\boldsymbol{r}_0$ persists.

In Remark 4.3.6 we will see how to remedy this deficiency under certain conditions by using a more handy generalized inverse. A careful analysis of our arguments above shows, that as long as no singular breakdown occurs in step m, every generalized inverse A^{X} is suitable, for which $A^{\mathrm{X}}A$ is a projection (i. e., $(A^{\mathrm{X}}A)^2 = A^{\mathrm{X}}A$) and $\mathcal{K}_m(A, \boldsymbol{r}_0) \subseteq \mathcal{R}(A^{\mathrm{X}}A)$.

4.2 Minimal Polynomial and Drazin Inverse

The minimal polynomial of a vector $\boldsymbol{r}$ with respect to A is a generalization of the minimal polynomial m_A of a $n \times n$ matrix A. We recall some known facts to exploit this connection. Let

$$m_A(\xi) = \xi^{k_0} \prod_{i=1}^{\ell} (\xi - \lambda_i)^{k_i}, \tag{4.5}$$

where λ_i $(i = 1, \ldots, \ell)$ are the distinct nonzero eigenvalues of A. The number $d := k_0$ is called the *index of* A. Setting $\lambda_0 = 0$, the numbers k_i are the dimensions of the largest Jordan block associated with the eigenvalue λ_i in the Jordan canonical form of A.

For later use we state three lemmata

Lemma 4.2.1. *The minimal polynomial of a matrix $A \in \mathbb{C}^{n\times n}$ is invariant under similarity transformations.*

Lemma 4.2.2. *If $N \in \mathbb{C}^{n\times n}$ is a nilpotent matrix of index d (i. e., $N^d = O$ and there is no smaller index $m < d$ such that $N^m = O$), then its minimal polynomial is $m_N(\xi) = \xi^d$.*

Lemma 4.2.3. *If $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$ have disjoint spectra, the minimal polynomial of $\begin{bmatrix} A & O \\ O & B \end{bmatrix}$ is the product $m_A(\xi)m_B(\xi)$.*

The *Drazin inverse* A^{D} of a $n \times n$ matrix A can be characterized as the unique solution of the following three equations (cf. Section 2):

$$\begin{aligned} A^{\mathrm{D}}AA^{\mathrm{D}} &= A^{\mathrm{D}}, \\ A^{\mathrm{D}}A &= AA^{\mathrm{D}} \quad \text{and} \\ A^{d+1}A^{\mathrm{D}} &= A^dA^{\mathrm{D}}, \end{aligned} \tag{4.6}$$

where $d = \operatorname{index}(A)$ is the index of A (cf. [14, p. 122]). Given A in the canonical form

$$A = T \begin{bmatrix} C & O \\ O & N \end{bmatrix} T^{-1}, \tag{4.7}$$

where $T \in \mathbb{C}^{n\times n}$ and $C \in \mathbb{C}^{r\times r}$, $r = \operatorname{rank}(A^d)$, are nonsingular and $N \in \mathbb{C}^{(n-r)\times(n-r)}$ is nilpotent of index d, the Drazin inverse of A can be written as

$$A^{\mathrm{D}} = T \begin{bmatrix} C^{-1} & O \\ O & O \end{bmatrix} T^{-1} \tag{4.8}$$

(cf. [14, Theorem 7.2.1.]). This implies $A^{\mathrm{D}} = O$ if and only if A is nilpotent.

The Drazin inverse of A can represented as a polynomial in A. More precisely, there holds

Proposition 4.2.4. *Let $A \in \mathbb{C}^{n\times n}$ with minimal polynomial as in (4.5). Let p denote the (uniquely determined) Hermite interpolation polynomial of degree $M = d+\sum_{i=1}^{\ell} k_i - 1 = \deg(m_A) - 1$ with*

$$\begin{aligned} p(0) = p'(0) = \cdots = p^{(d-1)}(0) = 0 \\ p^{(j)}(\lambda_i) = \frac{(-1)^j j!}{\lambda_i^{j+1}} \quad (j = 0, 1, \ldots, k_i - 1 \text{ and } i = 1, 2, \ldots, \ell). \end{aligned} \tag{4.9}$$

Then p is the unique polynomial of smallest degree, which satisfies

$$p(A) = A^{\mathrm{D}}. \tag{4.10}$$

Proof. The transformation of A to Jordan canonical form is given by

$$A = T \begin{bmatrix} J & O \\ O & N \end{bmatrix} T^{-1} \tag{4.11}$$

where the diagonal blocks J and N are block diagonal, namely $J = \operatorname{diag}(J_1, \ldots, J_h)$ and $N = \operatorname{diag}(N_1, \ldots, N_g)$. Each J_k is a Jordan block associated with a non-zero eigenvalue λ_i. The matrices N_j are the Jordan blocks associated with the zero eigenvalue which implies $N_j \in \mathbb{C}^{m\times m}$ with $m \leq d = k_0$. Moreover there is at least one $d \times d$ Jordan block in N. Thus, N is nilpotent of index d, which shows, that the Jordan canonical form is a special form of (4.7). Lemma 4.2.2 implies that $m_N(\xi) = \xi^d$ is the minimal polynomial of N.

Let p be an arbitrary polynomial. Then there holds $p(A) = T \begin{bmatrix} p(J) & 0 \\ 0 & p(N) \end{bmatrix} T^{-1}$, and, observing (4.8) and the identities

$$\begin{aligned} J^{-1} &= \operatorname{diag}\left({J_1}^{-1}, \ldots, {J_h}^{-1}\right), \\ p(J) &= \operatorname{diag}\left(p(J_1), \ldots, p(J_h)\right), \end{aligned}$$

we conclude, that equation (4.10) is equivalent to $p(J_k) = {J_k}^{-1}$ for $k = 1, \ldots, h$ and $p(N) = O$. Therefore p is divided by ξ^d, which can be rewritten as $p(0) = p'(0) = \cdots = p^{(d-1)} = 0$. Given

$$J_k = \begin{bmatrix} \lambda_i & 1 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & & 0 \\ \vdots & & \ddots & \ddots & \\ 0 & 0 & & \lambda_i & 1 \\ 0 & 0 & \ldots & 0 & \lambda_i \end{bmatrix} \in \mathbb{C}^{m \times m}$$

an easy computation shows that

$$p(J_k) = \begin{bmatrix} p(\lambda_i) & \frac{p'(\lambda_i)}{1!} & \frac{p''(\lambda_i)}{2!} & \ldots & \frac{p^{(m-1)}(\lambda_i)}{(m-1)!} \\ 0 & p(\lambda_i) & \frac{p'(\lambda_i)}{1!} & \ldots & \frac{p^{(m-1)}(\lambda_i)}{(m-1)!} \\ \vdots & & \ddots & & \vdots \\ 0 & & & p(\lambda_i) & \frac{p'(\lambda_i)}{1!} \\ 0 & 0 & \ldots & 0 & p(\lambda_i) \end{bmatrix} \quad \text{and}$$

$$J_k{}^{-1} = \begin{bmatrix} \frac{1}{\lambda_i} & \frac{-1}{\lambda_i^2} & \frac{1}{\lambda_i^3} & \ldots & \frac{(-1)^{m-1}}{\lambda_i^m} \\ 0 & \frac{1}{\lambda_i} & \frac{-1}{\lambda_i^2} & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & \ddots & \frac{-1}{\lambda_i^2} \\ 0 & 0 & \ldots & 0 & \frac{1}{\lambda_i} \end{bmatrix}.$$

Since the Jordan blocks corresponding to λ_i have maximum dimension k_i, and there is a block with exactly this dimension, we obtain the k_i equations $p^{(j)}(\lambda_i) = \frac{(-1)^j j!}{\lambda_i^{j+1}}$ $(j = 0, 1, \ldots, k_i - 1)$ for every eigenvalue λ_i.

In summary, given the matrix A with the above Jordan decomposition, (4.10) is equivalent to (4.9) for every polynomial p. Thus, the uniqueness of the Hermite interpolation polynomial shows our assertion. □

Note that the degree of p can be smaller than $M = \deg(m_A) - 1$.

Example 4.2.5. Let $A = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ with $m_A(\xi) = \xi^2$. Then $A^{\mathrm{D}} = O$ and the polynomial $p(\xi) = 0$ of degree 0 satisfies (4.10).

But this example illustrates – in essence – the only case where p has a degree smaller than $M = \deg(m_A) - 1$, a fact which is less often stated in the literature.

Proposition 4.2.6. *Let $A \in \mathbb{C}^{n \times n}$, m_A its minimal polynomial and $d = \mathrm{index}(A)$. If A is not nilpotent the degree of a polynomial p with $p(A) = A^{\mathrm{D}}$ can not be smaller than $M = \deg(m_A) - 1$.*

Proof. If $d = 0$, the matrix A is regular. Suppose $p(A) = A^{\mathrm{D}} = A^{-1}$ for some polynomial p. By multiplying with A and rearranging this equation we see that $s(\xi) := 1 - \xi p(\xi)$ is divided by m_A, i. e., $\deg p \geq \deg(m_A) - 1$

Now, let $d > 0$. The coefficients of the Hermite interpolation polynomial which satisfies (4.10) can be expressed by determinants.[1] Consider

$$\begin{vmatrix} p(\xi) & 1 & \xi & \xi^2 & \dots & \xi^{d-1} & \xi^d & \dots & \xi^M \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & & & & & \\ 0 & 0 & 0 & 2 & & & & & \\ \vdots & \vdots & & & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & & (d-1)! & 0 & \dots & 0 \\ \hline p^{(j)}(\lambda_i) & & & & & & \frac{d!}{(d-j)!}\lambda_i^{d-j} & & \frac{M!}{(M-j)!}\lambda_1^{M-j} \end{vmatrix}.$$

Each of the lower $M + 1$ rows corresponds to one of the interpolation conditions in (4.9). The first column consists of the right hand sides of the interpolation conditions and the subsequent columns contain the values/derivatives of the polynomials in the first row evaluated at λ_i. It is easy to see that the Hermite interpolation polynomial can be computed by expanding this determinant along the first row and isolating $p(\xi)$. The uniqueness of the Hermite interpolation guarantees that the lower right sub-determinant (which is the coefficient of $p(\xi)$ in the expansion) is nonzero, i. e., its columns are linearly independent. The coefficient of ξ^M is zero if and only if

$$\begin{vmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & & & & & \\ 0 & 0 & 0 & 2 & & & & & \\ \vdots & \vdots & & & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & & (d-1)! & 0 & \dots & 0 \\ \frac{1}{\lambda_1} & 1 & \lambda_1 & \lambda_1^2 & \dots & & & & \lambda_1^{M-1} \\ \frac{-1}{\lambda_1^2} & 0 & 1 & 2\lambda_1 & \dots & & & & (M-1)\lambda_1^{M-2} \\ \vdots & & & & & & & & \vdots \\ (-1)^{(k_1-1)}\frac{(k_1-1)!}{\lambda_1^{k_1}} & 0 & 0 & 0 & & & & & \frac{(M-1)!}{(M-k_1)!}\lambda_1^{M-k_1} \\ \frac{1}{\lambda_2} & 1 & \lambda_2 & \lambda_2^2 & \dots & & & & \lambda_2^{M-1} \\ \hline \frac{(-1)^j j!}{\lambda_i^{j+1}} & & & & & & \frac{d!}{(d-j)!}\lambda_i^{d-j} & & \frac{(M-1)!}{(M-1-j)!}\lambda_i^{M-1-j} \\ \hline (-1)^{(k_\ell-1)}\frac{(k_\ell-1)!}{\lambda_\ell^{k_\ell}} & 0 & 0 & & & & & & \frac{(M-1)!}{(M-k_\ell)!}\lambda_\ell^{M-k_\ell} \end{vmatrix} = 0.$$

[1] This idea was found in a Usenet article by David Petry in `sci.math.research`, Message-ID: `41qo7a$h3l@ixnews3.ix.netcom.com`.

Since we only ask if the determinant is zero or not, we can scale each row or column by a suitable nonzero factor. Thus, in the first $d-1$ rows the only nonzero entry can be scaled to 1 and multiplying the rows, which start with $\frac{(-1)^j j!}{\lambda_i^{j+1}}$ by $\frac{\lambda_i^{j+1}}{j!} \neq 0$ leads to

$$\begin{vmatrix}
0 & 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\
0 & 0 & 1 & 0 & & & & & \\
0 & 0 & 0 & 1 & & & & & \\
\vdots & \vdots & & & \ddots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \dots & & 1 & 0 & \dots & 0 \\
1 & \lambda_1 & \lambda_1^2 & \lambda_1^3 & \dots & & & & \lambda_1^M \\
-1 & 0 & \lambda_1^2 & 2\lambda_1^3 & \dots & & & & (M-1)\lambda_1^M \\
\vdots & & & & & & & & \vdots \\
(-1)^{(k_1-1)} & 0 & 0 & 0 & & & & & \binom{M-1}{k_1-1}\lambda_1^M \\
1 & \lambda_2 & \lambda_2^2 & \lambda_2^3 & \dots & & & & \lambda_2^M \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
(-1)^j & & & & & & \binom{d}{j}\lambda_i^{d+1} & & \binom{M-1}{j}\lambda_i^M \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
(-1)^{(k_\ell-1)} & 0 & 0 & & & & & & \binom{M-1}{k_\ell-1}\lambda_\ell^M
\end{vmatrix}.$$

If the eigenvalues λ_i are sorted according to the size of the largest Jordan block, i. e., such that $k_i \leq k_j$ for $i > j$, we can rearrange the rows as follows

$$\begin{vmatrix}
0 & 1 & & 0 & 0 & \dots & 0 \\
\vdots & & \ddots & & \vdots & & \vdots \\
0 & 0 & & 1 & 0 & \dots & 0 \\
1 & \lambda_1 & \lambda_1^2 & \lambda_1^3 & \dots & \dots & \lambda_1^M \\
\vdots & & & & & & \vdots \\
1 & \lambda_\ell & \lambda_\ell^2 & \lambda_\ell^3 & \dots & \dots & \lambda_\ell^M \\
-1 & 0 & \lambda_1^2 & 2\lambda_1^3 & \dots & \dots & (M-1)\lambda_1^M \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
(-1)^j & 0 & 0 & & \lambda_1^{j+1} & \binom{d}{j}\lambda_1^{d+1} & \binom{M-1}{j}\lambda_1^M \\
\vdots & & & & & & \vdots \\
(-1)^j & 0 & 0 & & \lambda_{\ell_j}^{j+1} & \binom{d}{j}\lambda_{\ell_j}^{d+1} & \binom{M-1}{j}\lambda_{\ell_j}^M \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
(-1)^{(k_1-1)} & 0 & 0 & & & & \binom{M-1}{k_1-1}\lambda_1^M \\
\vdots & & & & & & \vdots \\
(-1)^{(k_1-1)} & 0 & 0 & & & & \binom{M-1}{k_1-1}\lambda_{\ell_{k_1-1}}^M
\end{vmatrix},$$

where we denote by $\ell_j = \max\{i \ : \ 1 \le i \le \ell,\ k_i > j$. Each j, $0 \le j < k_1$, corresponds to a row block in this determinant. Obviously, the rows in different blocks are linearly independent. On the other hand, the rows in each block are linearly independent, since they are – up to suitable column scaling and additional zero columns – rows of Vandermonde determinants. Thus, the determinant cannot be zero unless it reduces to the upper left $d \times d$ block, which has an all zero last row. □

A third equivalent definition of the Drazin inverse can be given in functional terms (cf. [14, Section 7.2]): The index of A can be characterized as the smallest integer for which $\mathcal{R}(A^d) = \mathcal{R}(A^{d+1})$ and $\mathcal{N}(A^d) = \mathcal{N}(A^{d+1})$. Moreover there holds

$$\mathcal{H} = \mathcal{R}(A^d) \oplus \mathcal{N}(A^d) \tag{4.12}$$

and the restriction $A|_{\mathcal{R}(A^d)}$ of A to the range of A^d is an invertible linear operator on $\mathcal{R}(A^d)$. Given a vector $\boldsymbol{x} \in \mathcal{H}$, with the unique decomposition $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$, $\boldsymbol{y} \in \mathcal{R}(A^d)$, $\boldsymbol{z} \in \mathcal{N}(A^d)$, the Drazin inverse is defined by $A^{\mathrm{D}}\boldsymbol{x} = (A|_{\mathcal{R}(A^d)})^{-1}\boldsymbol{y}$.

Remark 4.2.7. With regard to the canonical forms (4.7) and (4.8) we note that the matrix C represents the operator A restricted to $\mathcal{R}(A^d)$ with respect to the basis T (recall that the columns of T forms a basis of $\mathcal{H} = \mathbb{C}^n$ if T is nonsingular). Similarly, N describes the restriction of A to $\mathcal{N}(A^d)$ in terms of this basis. We retain the distinction between $\mathcal{H}$ and the coordinate vectors and matrices even if $\mathcal{H} = \mathbb{C}^n$ is assumed. Thus we write, e. g., A^* for the adjoint of an operator A while we use C^{H} to denote the conjugate transpose of the coordinate matrix C.

For the reader's convenience we state an important property of the Drazin inverse, which can be found e. g. in [14].

Lemma 4.2.8. *Let d denote the index of A. The (uniquely determined) projection onto $\mathcal{R}(A^d)$ along $\mathcal{N}(A^d)$ is given by $P_{\mathcal{R}(A^d),\mathcal{N}(A^d)} = A^{\mathrm{D}}A$ and there holds*

$$\begin{aligned} &\mathcal{R}(A^{\mathrm{D}}A) = \mathcal{N}(I - A^{\mathrm{D}}A) = \mathcal{R}(A^{\mathrm{D}}) = \mathcal{R}(A^d) \qquad \textit{and} \\ &\mathcal{N}(A^{\mathrm{D}}A) = \mathcal{R}(I - A^{\mathrm{D}}A) = \mathcal{N}(A^{\mathrm{D}}) = \mathcal{N}(A^d). \end{aligned}$$

For $j \le d$, an analogous decomposition to (4.12) can be given for $\mathcal{R}(A^{d-j})$.

Lemma 4.2.9. *Denote by d the index of A and let $j \in \mathbb{N}_0$, $0 \le j \le d$. Then there holds*

$$\mathcal{R}(A^{d-j}) = \mathcal{R}(A^d) \oplus \mathcal{N}(A^j).$$

Proof. Let $\boldsymbol{v} \in \mathcal{R}(A^{d-j})$. This means there exists a vector $\boldsymbol{w} \in \mathcal{H}$ with $\boldsymbol{v} = A^{d-j}\boldsymbol{w}$. Using (4.12) this vector has the unique decomposition $\boldsymbol{w} = \boldsymbol{r} + \boldsymbol{s}$ with $\boldsymbol{r} \in \mathcal{R}(A^d)$ and $\boldsymbol{s} \in \mathcal{N}(A^d)$. Applied to $\boldsymbol{v}$ this gives

$$\boldsymbol{v} = A^{d-j}\boldsymbol{w} = A^{d-j}\boldsymbol{r} + A^{d-j}\boldsymbol{s} =: \boldsymbol{y} + \boldsymbol{z},$$

where $\boldsymbol{y} = A^{d-j}\boldsymbol{r} \in \mathcal{R}(A^d)$ and $\boldsymbol{z} = A^{d-j}\boldsymbol{s} \in \mathcal{N}(A^d)$ are the unique components of $\boldsymbol{v}$. One computes easily

$$A^j\boldsymbol{z} = A^j A^{d-j}\boldsymbol{s} = A^d\boldsymbol{s} = \boldsymbol{0},$$

which shows that $\boldsymbol{v} \in \mathcal{R}(A^d) \oplus \mathcal{N}(A^j)$, i. e., $\mathcal{R}(A^{d-j}) \subseteq \mathcal{R}(A^d) \oplus \mathcal{N}(A^j)$.

Let now $\boldsymbol{v} = \boldsymbol{y}+\boldsymbol{z}$ with $\boldsymbol{y} \in \mathcal{R}(A^d)$ and $\boldsymbol{z} \in \mathcal{N}(A^j) \subseteq \mathcal{N}(A^d)$. Now, choose $\boldsymbol{r} \in \mathcal{R}(A^d)$ and $\boldsymbol{s} \in \mathcal{N}(A^d)$ such that $\boldsymbol{y} = A^{d-j}\boldsymbol{r}$ and $\boldsymbol{z} = A^{d-j}\boldsymbol{s}$. Then $\boldsymbol{v} = \boldsymbol{y}+\boldsymbol{z} = A^{d-j}(\boldsymbol{r}+\boldsymbol{s}) \in \mathcal{R}(A^{d-j})$, which completes the proof. □

A well-known property of the minimal polynomial of a matrix A is that m_A divides every nonzero polynomial p with $p(A) = O$. The following result is a generalization of this fact to the minimal polynomial of a vector $\boldsymbol{r}$ with respect to A.

Proposition 4.2.10. *The minimality polynomial $m_{\boldsymbol{r},A}$ divides every nonzero polynomial p which satisfies $p(A)\boldsymbol{r} = \boldsymbol{0}$.*

Proof. The proof is an application of Euclid's algorithm. Clearly we have $\deg p \geq \deg m_{\boldsymbol{r},A}$. We divide p by $m_{\boldsymbol{r},A}$ to obtain

$$p(\xi) = s(\xi)m_{\boldsymbol{r},A}(\xi) + r(\xi),$$

with $\deg r < \deg m_{\boldsymbol{r},A}$. Since $r(A)\boldsymbol{r} = p(A)\boldsymbol{r} - s(A)m_{\boldsymbol{r},A}(A)\boldsymbol{r} = \boldsymbol{0}$, there is a contradiction to the minimal property of $m_{\boldsymbol{r},A}$ unless r is identically zero. □

In particular, this shows that the minimal polynomial $m_{\boldsymbol{r},A}$ of any vector $\boldsymbol{r}$ divides the minimal polynomial of A. Using (4.5) this leads to the representation

$$m_{\boldsymbol{r},A}(\xi) = \xi^q \prod_{i=1}^{k} (\xi - \lambda_i)^{n_i}, \tag{4.13}$$

where $\lambda_i \neq 0$, $i = 1, \ldots, k$, are nonzero eigenvalues of A. The quantity $q = \operatorname{index}(\boldsymbol{r}, A)$ is called the *index of $\boldsymbol{r}$ with respect to A*. It can be characterized independently of the above representation.

Proposition 4.2.11. *If q denotes the index of $\boldsymbol{r}$ with respect to $A \in \mathbb{C}^{n\times n}$ then there holds*

$$q = \min\{j \in \mathbb{N}_0 \ : \ 0 \leq j \leq d \text{ and } \boldsymbol{r} \in \mathcal{R}(A^{d-j})\}, \tag{4.14}$$

where $d = \operatorname{index}(A)$. Moreover there holds

$$\boldsymbol{r} \in \mathcal{R}(A^{d-j}) = \mathcal{R}(A^d) \oplus \mathcal{N}(A^j) \Longleftrightarrow q \leq j \leq d. \tag{4.15}$$

Proof. Let A be given in the form (4.7). Then the decomposition (4.12) of $\boldsymbol{r}$ is given by

$$\boldsymbol{r} = T \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{0} \end{bmatrix} + T \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{z} \end{bmatrix}.$$

Denote the minimal polynomial of $\boldsymbol{r}$ with respect to A by $m_{\boldsymbol{r},A} = \xi^q c(\xi)$ with $c(0) \neq 0$. Then there holds

$$\begin{aligned}\boldsymbol{0} = m_{\boldsymbol{r},A}(A)\boldsymbol{r} &= T \begin{bmatrix} m_{\boldsymbol{r},A}(C) & O \\ O & m_{\boldsymbol{r},A}(N) \end{bmatrix} \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{0} \end{bmatrix} + T \begin{bmatrix} m_{\boldsymbol{r},A}(C) & O \\ O & m_{\boldsymbol{r},A}(N) \end{bmatrix} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{z} \end{bmatrix} \\ &= T \begin{bmatrix} m_{\boldsymbol{r},A}(C)\boldsymbol{y} \\ \boldsymbol{0} \end{bmatrix} + T \begin{bmatrix} \boldsymbol{0} \\ m_{\boldsymbol{r},A}(N)\boldsymbol{z} \end{bmatrix},\end{aligned}$$

i. e., $m_{\boldsymbol{r},A}(C)\boldsymbol{y} = \boldsymbol{0}$ and $m_{\boldsymbol{r},A}(N)\boldsymbol{z} = \boldsymbol{0}$. Hence the minimal polynomials $m_{\boldsymbol{y},C}$ and $m_{\boldsymbol{z},N}$ divide $m_{\boldsymbol{r},A}$. Since N is nilpotent, $m_{\boldsymbol{z},N}(\xi)$ has to be of the form ξ^j (cf. Lemma 4.2.2). The fact, that $\xi^q c(\xi)$ is a multiple of $m_{\boldsymbol{z},N}(\xi) = \xi^j$ implies $j \leq q$. Now the assumption $j < q$ contradicts the minimality property of $m_{\boldsymbol{r},A}$ – consider $p(\xi) := \xi^j c(\xi)$. Hence $m_{\boldsymbol{z},N}(\xi) = \xi^q$, which implies $N^q\boldsymbol{z} = \boldsymbol{0}$ and, using (4.7),

$$A^q T \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{z} \end{bmatrix} = T \begin{bmatrix} \boldsymbol{0} \\ N^q\boldsymbol{z} \end{bmatrix} = \boldsymbol{0}, \quad \text{i. e.,} \quad T \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{z} \end{bmatrix} \in \mathcal{N}(A^q).$$

This proves $\boldsymbol{r} \in \mathcal{R}(A^d) \oplus \mathcal{N}(A^q) = \mathcal{R}(A^{d-q})$, where the last equation follows from Lemma 4.2.9. Moreover, since ξ^q is the minimal polynomial of $\boldsymbol{z}$ with respect to N, there can not exist a smaller index $j < q$, such that $T \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{z} \end{bmatrix} \in \mathcal{N}(A^j)$.

All assertions now follow from the obvious equivalence

$$\mathcal{R}(A^{d-q}) \subseteq \mathcal{R}(A^{d-j}) \Longleftrightarrow q \leq j. \qquad \square$$

Corollary 4.2.12. *If q denotes the index of $\boldsymbol{r}$ with respect to A and $d = \operatorname{index}(A)$, then*

$$\mathcal{K}_m(A, \boldsymbol{r}) \subseteq \mathcal{R}(A^{d-q}) \quad \textit{for all } m.$$

Remark 4.2.13. If $\mathcal{H}$ is not finite dimensional, we can define the Drazin inverse for bounded linear operators which possess a decomposition like (4.12). We call the sequence of subspaces

$$\mathcal{H} = \mathcal{R}(A^0) \supseteq \mathcal{R}(A) \supseteq \mathcal{R}(A^2) \supseteq \cdots \supseteq \mathcal{R}(A^k) \supseteq \cdots$$

the *range chain* of A. If there exists an index $k < \infty$ for which equality holds, i. e., $\mathcal{R}(A^k) = \mathcal{R}(A^{k+1})$, we have $\mathcal{R}(A^m) = \mathcal{R}(A^k)$ for all $m \leq k$. If such an index exists we define

$$a = \min\{n \in \mathbb{N}_0 \,:\, \mathcal{R}(A^n) = \mathcal{R}(A^{n+1})\}$$

and set $a = \infty$ otherwise.

Similarly we call

$$\{\boldsymbol{0}\} = \mathcal{N}(A^0) \subseteq \mathcal{N}(A) \subseteq \mathcal{N}(A^2) \subseteq \cdots \subseteq \mathcal{N}(A^m) \subseteq \cdots$$

the *nullspace chain* of A and define as its length the smallest integer n with $\mathcal{N}(A^n) = \mathcal{N}(A^{n+1})$.

If both chain lengths, a and n are finite, they must coincide (cf. [38, Satz 72.3]) and we define the *index of* A by this integer, $d := a = n$. The existence of a finite index d is equivalent to the existence of the direct decomposition $\mathcal{H} = \mathcal{R}(A^d) \oplus \mathcal{N}(A^d)$ (cf. [38, Satz 72.4]) and the restriction $A|_{\mathcal{R}(A^d)}$ of A to $\mathcal{R}(A^d)$ is invertible.

The Drazin inverse now may be defined in functional terms as $A^{\mathrm{D}}\boldsymbol{v} = {A|_{\mathcal{R}(A^d)}}^{-1}\boldsymbol{y}$, where $\boldsymbol{v}$ has the unique decomposition $\boldsymbol{v} = \boldsymbol{y} + \boldsymbol{z}$ with $\boldsymbol{y} \in \mathcal{R}(A^d)$ and $\boldsymbol{z} \in \mathcal{N}(A^d)$. Equivalently, we can define A^{D} as the unique solution of (4.6). The index q of a vector $\boldsymbol{r}$ with respect to A can be defined by (4.14).

4.3 Termination of Krylov Methods

Recall that the degree of $m_{\boldsymbol{r},A}$ coincides with the termination index of the sequence of Krylov spaces. In terms of the representation (4.13) this means

$$L = q + \sum_{i=1}^{k} n_i.$$

Another characterization of L is, that it is the index of the first Krylov space which contains $A^{\mathrm{D}}\boldsymbol{r}$.

Proposition 4.3.1. *Suppose that* $\boldsymbol{r} \notin \mathcal{N}(A^d)$, $d = \mathrm{index}(A)$. *Then there holds*

$$L = \min\{m \,:\, A^{\mathrm{D}}\boldsymbol{r} \in \mathcal{K}_m(A, \boldsymbol{r})\}. \tag{4.16}$$

Proof. Denote by L the degree of the minimal polynomial of $\boldsymbol{r}$ with respect to A (or equivalently let L as defined in (3.13)). We first show that the Krylov space $\mathcal{K}_L(A, \boldsymbol{r})$ always contains $A^{\mathrm{D}}\boldsymbol{r}$. Since $\mathcal{K}_L(A, \boldsymbol{r}) = \{p(A)\boldsymbol{r} \,:\, p \in \mathcal{P}_{L-1}\}$, this is equivalent to the existence of a polynomial of degree at most $L-1$ which satisfies $A^{\mathrm{D}}\boldsymbol{r} = p(A)\boldsymbol{r}$. This polynomial is given by the Hermite interpolation polynomial with

$$p(0) = p'(0) = \cdots = p^{(q-1)} = 0$$
$$p^{(j)}(\lambda_i) = \frac{(-1)^j j!}{\lambda_i^{j+1}} \quad (j = 0, 1, \ldots, n_i - 1 \text{ and } i = 1, 2, \ldots, \ell),$$

which can be seen using the Jordan canonical form of A as follows: There exists a permutation P for the similarity transformation T in (4.11) such that A and $\boldsymbol{r}$ can be written as

$$A = TP \begin{bmatrix} \hat{J} & O & R_J \\ O & \hat{N} & R_N \\ O & O & \hat{R} \end{bmatrix} P^* T^{-1} \qquad \text{and} \qquad \boldsymbol{r} = TP \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{z} \\ \boldsymbol{0} \end{bmatrix},$$

where the Jordan blocks in the block diagonal matrix $\hat{J}$ corresponding to an eigenvalue λ_i have maximal dimension n_i and $\hat{N}$ is nilpotent of index q. Using that we can proceed as in the proof of Proposition 4.2.4.

By the same argument as in Proposition 4.2.6 we conclude that $\deg p = L - 1$ if $L > q$. □

Remark 4.3.2. Consider a Krylov subspace method for solving (3.1) with initial residual $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$. If $\boldsymbol{r}_0 \in \mathcal{N}(A^d) = \mathcal{N}(A^{\mathrm{D}})$ we have $A^{\mathrm{D}}\boldsymbol{r}_0 = \boldsymbol{0}$ and $\mathcal{K}_m(A, \boldsymbol{r}_0) \subseteq \mathcal{N}(A^d)$ for all m.

The index of $\boldsymbol{r}_0$ with respect to A is the smallest integer q for which $\boldsymbol{r}_0 \in \mathcal{N}(A^q)$ (cf. (4.15)). If $\boldsymbol{r}_0 \neq \boldsymbol{0}$ (and thus $d \neq 0$) we have $q > 0$. Since $\mathcal{R}(A^d) \cap \mathcal{N}(A^d) = \{\boldsymbol{0}\}$ there holds $m_{\boldsymbol{r}_0,A}(\xi) = \xi^q$ and the sequence of Krylov spaces terminates with $L = q$. Using (4.4) we get $\boldsymbol{r}_m = p_m(A)\boldsymbol{r}_0$ with $p_m(\xi) := 1 - \xi q_{m-1}(\xi) \in \mathcal{P}_m$ and $q_{m-1}(A)\boldsymbol{r}_0 = \boldsymbol{c}_m$ for each correction $\boldsymbol{c}_m \in \mathcal{K}_m(A, \boldsymbol{r}_0)$. Therefore the residual is nonzero if $m < L$ and $\boldsymbol{r}_L = \boldsymbol{r}_q = \boldsymbol{0}$. Moreover, the breakdown in step $L = q$ is singular since $A^q\boldsymbol{r}_0 = \boldsymbol{0}$ and $\mathcal{W}_L = A\mathcal{K}_L(A, \boldsymbol{r}_0) = \operatorname{span}\{A\boldsymbol{r}_0, \ldots, A^{q-1}\boldsymbol{r}_0, A^q\boldsymbol{r}_0\} = \operatorname{span}\{A\boldsymbol{r}_0, \ldots, A^{q-1}\boldsymbol{r}_0\} = A\mathcal{K}_{L-1}(A, \boldsymbol{r}_0) = \mathcal{W}_{L-1}$ (cf. (viii) in Proposition 3.4.4).

The inclusion of $A^{\mathrm{D}}\boldsymbol{r}$ in the largest Krylov space $\mathcal{K}_L(A, \boldsymbol{r})$ and the A-invariance of this space yields that it is invariant under A^{D} too (and it is not difficult to see, that L is the smallest index for which this invariance holds, provided that $\boldsymbol{r} \notin \mathcal{N}(A^d)$). This leads to a decomposition of $\mathcal{K}_L(A, \boldsymbol{r})$ into the components on range and nullspace of A^d (cf. Lemmata 4.2.8 and 4.2.9 and Corollary 4.2.12).

Corollary 4.3.3. *Denote by L the termination index of the sequence of Krylov subspaces of $\boldsymbol{r}$ with respect to A and by q the index of $\boldsymbol{r}$ with respect to A. Let $\boldsymbol{r} = \boldsymbol{s} + \boldsymbol{z}$ with $\boldsymbol{s} \in \mathcal{R}(A^d)$, i. e., $\boldsymbol{s} = P_{\mathcal{R}(A^d),\mathcal{N}(A^d)}\boldsymbol{r}$, and $\boldsymbol{z} = P_{\mathcal{N}(A^d),\mathcal{R}(A^d)}\boldsymbol{r} \in \mathcal{N}(A^d)$. Then here holds*

$$\mathcal{K}_L(A, \boldsymbol{s}) = A^{\mathrm{D}}A\mathcal{K}_L(A, \boldsymbol{r}) = \mathcal{R}(A^d) \cap \mathcal{K}_L(A, \boldsymbol{r}) \subseteq \mathcal{K}_L(A, \boldsymbol{r}), \tag{4.17}$$

$$\mathcal{K}_L(A, \boldsymbol{z}) = \mathcal{N}(A^d) \cap \mathcal{K}_L(A, \boldsymbol{r}) = \mathcal{N}(A^q) \cap \mathcal{K}_L(A, \boldsymbol{r}) \subseteq \mathcal{K}_L(A, \boldsymbol{r}) \quad \textit{and}$$

$$\mathcal{K}_L(A, \boldsymbol{s}) \oplus \mathcal{K}_L(A, \boldsymbol{z}) = \mathcal{K}_L(A, \boldsymbol{r}). \tag{4.18}$$

In the rightmost relation of (4.17) *equality holds if and only if $q = 0$, which is equivalent to $\boldsymbol{z} = 0$ or $\mathcal{N}(A^d) \cap \mathcal{K}_L(A, \boldsymbol{r}) = \{\boldsymbol{0}\}$.*

The only relation in (4.17), which is *not* true in general (for $m < L$ and $q > 0$) is $\mathcal{K}_m(A, P_{\mathcal{R}(A^d),\mathcal{N}(A^d)}\boldsymbol{r}) \subseteq \mathcal{K}_m(A, \boldsymbol{r})$. The decomposition (4.18) corresponds to that in Lemma 4.2.9 restricted to $\mathcal{K}_L(A, \boldsymbol{r})$.

As already noted, the largest Krylov space is invariant under the Drazin inverse, more precisely, in view of Corollary 4.3.3, we have $A^{\mathrm{D}}\mathcal{K}_L(A, \boldsymbol{r}) = \mathcal{K}_L(A, \boldsymbol{s}) = A^{\mathrm{D}}A\mathcal{K}_L(A, \boldsymbol{r}) = \mathcal{R}(A^d) \cap \mathcal{K}_L(A, \boldsymbol{r})$. In essence, we are able to express quantities related with A^{D} using the Drazin inverse of H_L, the orthogonal section of A onto the invariant Krylov subspace.

Proposition 4.3.4. *Suppose, for $\boldsymbol{r} \in \mathcal{H}$ with $q = \operatorname{index}(\boldsymbol{r}, A)$, that we have established the Arnoldi decomposition $AV_L = V_L H_L$ with $\operatorname{span}\{V_L\} = \mathcal{K}_L(A, \boldsymbol{r})$. Then the index of H_L equals q and for all $\boldsymbol{v} \in \mathcal{K}_L(A, \boldsymbol{r})$ with $\boldsymbol{v} = V_L \boldsymbol{y}$, $\boldsymbol{y} \in \mathbb{C}^L$, there holds*

$$\begin{aligned} P_{\mathcal{R}(A^d),\mathcal{N}(A^d)}\boldsymbol{v} &= V_L H_L^{\mathrm{D}} H_L \boldsymbol{y} && \in \mathcal{R}(A^d) \cap \mathcal{K}_L(A, \boldsymbol{r}), \\ P_{\mathcal{N}(A^d),\mathcal{R}(A^d)}\boldsymbol{v} &= V_L (I - H_L^{\mathrm{D}} H_L)\boldsymbol{y} && \in \mathcal{N}(A^q) \cap \mathcal{K}_L(A, \boldsymbol{r}) \qquad \textit{and} \\ A^{\mathrm{D}}\boldsymbol{v} &= V_L H_L^{\mathrm{D}} \boldsymbol{y}. \end{aligned}$$

Using the QR decomposition (3.32) of H_L we can compute a basis of the nullspace of H_L^q (and at the same time a basis of $\mathcal{N}(A^q) \cap \mathcal{K}_L(A, \boldsymbol{r})$). Assume $q > 0$, i. e., H_L is singular. Recall from (3.31) and (3.36), that the nullspace of H_L is spanned by

$$\boldsymbol{g}^{(1)} := \tilde{\boldsymbol{g}} = \begin{bmatrix} \boldsymbol{g} \\ -1 \end{bmatrix} = \begin{bmatrix} \tilde{H}_{L-1}^{\dagger} \boldsymbol{h}_L \\ -1 \end{bmatrix} = \begin{bmatrix} R_{L-1}^{-1} \boldsymbol{t} \\ -1 \end{bmatrix}.$$

Thus, a vector $\boldsymbol{g}^{(2)}$ for which $H_L^2 \boldsymbol{g}^{(2)} = \boldsymbol{0}$ but $H_L \boldsymbol{g}^{(2)} \neq \boldsymbol{0}$ must satisfy

$$H_L \boldsymbol{g}^{(2)} = \zeta \boldsymbol{g}^{(1)} \tag{4.19}$$

for some $\zeta \neq 0$. Such a vector exists if and only if $\operatorname{index}(H_L) > 1$, i. e., if and only if $q \geq 2$. Multiplying (4.19) by Q_{L-1} and using (3.32) we get

$$\begin{bmatrix} R_{L-1} & \boldsymbol{t} \\ \boldsymbol{0}^{\top} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{s} \\ \sigma \end{bmatrix} = \zeta \begin{bmatrix} \boldsymbol{f} \\ \varphi \end{bmatrix}, \quad \text{where} \quad Q_{L-1}\tilde{\boldsymbol{g}} = \begin{bmatrix} \boldsymbol{f} \\ \varphi \end{bmatrix} \quad \text{and} \quad \boldsymbol{y} = \begin{bmatrix} \boldsymbol{s} \\ \sigma \end{bmatrix}.$$

Thus, (4.19) is solvable if and only if φ, which is the last component of $Q_{L-1}\tilde{\boldsymbol{g}}$, happens to be zero. If we choose the normalization $\zeta = 1$ and $\sigma = -1$ we result in the solution

$$\boldsymbol{g}^{(2)} = \begin{bmatrix} \boldsymbol{s} \\ \sigma \end{bmatrix} = \begin{bmatrix} R_{L-1}^{-1}\boldsymbol{f} + \boldsymbol{g} \\ -1 \end{bmatrix} = \begin{bmatrix} R_{L-1}^{-1}\boldsymbol{f} \\ 0 \end{bmatrix} + \boldsymbol{g}^{(1)}$$

and $\operatorname{span}\{\boldsymbol{g}^{(1)}, \boldsymbol{g}^{(2)}\} = \mathcal{N}(H_L^2)$. An obvious induction shows, that we can compute a basis $\{\boldsymbol{g}^{(1)}, \ldots, \boldsymbol{g}^{(q)}\}$ of $\mathcal{N}(H_L^q)$ and the basis $\{V_L\boldsymbol{g}^{(1)}, \ldots, V_L\boldsymbol{g}^{(q)}\}$ of $\mathcal{N}(A^q) \cap \mathcal{K}_L(A, \boldsymbol{r}) = \mathcal{N}(A^d) \cap \mathcal{K}_L(A, \boldsymbol{r})$ in this way.

The occurrence of a regular or singular breakdown only depends on the index q of the initial residual with respect to A:

Theorem 4.3.5. *Let A, $\boldsymbol{b}$ and $\boldsymbol{x}_0$ be given and apply an MR or OR method with $\mathcal{C}_m = \mathcal{K}_m(A, \boldsymbol{r}_0)$, $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$, to compute approximate solutions $\boldsymbol{x}_m^{\mathrm{MR}}$ or $\boldsymbol{x}_m^{\mathrm{OR}}$ of (3.1). Then there holds:*

A regular breakdown occurs if and only if $q = \operatorname{index}(\boldsymbol{r}_0, A) = 0$, i. e., if $\boldsymbol{r}_0 \in \mathcal{R}(A^d)$. In this case, the corresponding MR and OR iterates are given by

$$\boldsymbol{x}_L^{\mathrm{MR}} = \boldsymbol{x}_L^{\mathrm{OR}} = A^{\mathrm{D}}\boldsymbol{b} + (I - A^{\mathrm{D}}A)\boldsymbol{x}_0 = \boldsymbol{x}_0 + A^{\mathrm{D}}\boldsymbol{r}_0.$$

Proof. A regular breakdown occurs if and only if $\boldsymbol{r}_0 \in \mathcal{W}_m = A\mathcal{K}_m(A, \boldsymbol{r}_0)$ (cf. (ii) in Proposition 3.4.3). Then there exists a polynomial p of degree $m-1$ such that $\boldsymbol{r}_0 - Ap(A)\boldsymbol{r}_0 = \boldsymbol{0}$. Thus, the polynomial $r(\xi) := 1 - \xi p(\xi)$ is divided by the minimal polynomial of $\boldsymbol{r}_0$ with respect to A (cf. Lemma 4.2.10), which shows that the latter has a nonzero constant term, i. e., $q = 0$.

Suppose now, that $q = 0$ and denote the minimal polynomial by

$$m_{\boldsymbol{r}_0,A}(\xi) = \xi^L + \alpha_{L-1}\xi^{L-1} + \cdots + \alpha_1\xi + \alpha_0,$$

where $\alpha_0 \neq 0$. Consider

$$p(\xi) := \frac{1}{\alpha_0}\xi^{L-1} + \frac{\alpha_{L-2}}{\alpha_0}\xi^{L-1} + \cdots + \frac{\alpha_2}{\alpha_0}\xi + \frac{\alpha_1}{\alpha_0}.$$

Obviously, $\boldsymbol{r}_0 - Ap(A)\boldsymbol{r}_0 = \boldsymbol{0}$ and, since $p(A)\boldsymbol{r}_0 \in \mathcal{K}_L(A, \boldsymbol{r}_0)$, we conclude that $\boldsymbol{r}_0 \in A\mathcal{K}_L(A, \boldsymbol{r}_0)$. From the properties of the minimal polynomial it follows also that there exists no smaller index $m < L$ with $\boldsymbol{r}_0 \in A\mathcal{K}_m(A, \boldsymbol{r}_0) = \mathcal{W}_m$.

Since $A^{\mathrm{D}}\boldsymbol{r}_0 \in \mathcal{K}_L(A, \boldsymbol{r}_0)$ a suitable correction is $\boldsymbol{c}_L = A^{\mathrm{D}}\boldsymbol{r}_0$. We compute

$$\boldsymbol{r}_L = \boldsymbol{r}_0 - A\boldsymbol{c}_L = \boldsymbol{r}_0 - AA^{\mathrm{D}}\boldsymbol{r}_0 = (I - AA^{\mathrm{D}})\boldsymbol{r}_0 = \boldsymbol{0}, \tag{4.20}$$

where we made use of the fact, that $\boldsymbol{r}_0 \in \mathcal{R}(A^d)$ (cf. Lemma 4.2.11). This shows that $\boldsymbol{c}_L$ is the MR and OR correction which is uniquely determined (cf. Proposition 3.4.2 and Corollary 3.5.6). □

Note that we have only made use of the fact that the largest Krylov space contains $A^{\mathrm{D}}\boldsymbol{r}_0$. The minimality property (4.16) is not needed in general and in the *regular case* $q = 0$ this property is a simple consequence of (4.20) and Proposition 3.4.7 (see also Lemma 3.1.1).

Theorem 4.3.5 is a reformulation of a result by Ipsen and Meyer [44]: A Krylov space contains a solution of (3.1) if and only if the right hand side $\boldsymbol{b}$ belongs to the range of A^d and the initial guess $\boldsymbol{x}_0$ is chosen such that $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0 \in \mathcal{R}(A^d)$. For $\boldsymbol{b} \neq \boldsymbol{0}$ this can always be achieved by $\boldsymbol{x}_0 = \boldsymbol{0}$.

In other words: If the linear equation (3.1) is consistent but $\boldsymbol{b} \notin \mathcal{R}(A^d)$, then *no* Krylov space can contain a solution, unless $\boldsymbol{x}_0$ is chosen such that $A\boldsymbol{x}_0$ has the same components as $\boldsymbol{b}$ in the direction of $\mathcal{N}(A^d)$, i. e., $P_{\mathcal{N}(A^d),\mathcal{R}(A^d)}\boldsymbol{b} = P_{\mathcal{N}(A^d),\mathcal{R}(A^d)}A\boldsymbol{x}_0$. But in general, computing of the projection $P_{\mathcal{N}(A^d),\mathcal{R}(A^d)} = (I - AA^{\mathrm{D}})$ is as difficult as the original problem of solving a linear equation.

Remark 4.3.6. If $\boldsymbol{r}_0 \in \mathcal{R}(A^d)$ the Drazin inverse is an equation solving inverse on $\mathcal{K}_m(A, \boldsymbol{r}_0) \subseteq \mathcal{R}(A^d)$ and no Krylov subspace contains components in the nullspace of A, i. e., $\mathcal{N}(A) \cap \mathcal{K}_m(A, \boldsymbol{r}_0) = \{\boldsymbol{0}\}$ for all $m \leq L$. Thus it is not surprising that all known results about Krylov subspace methods for a nonsingular operator (cf. [22] and [23]) remain valid for a singular operator A and an initial residual $\boldsymbol{r}_0$ such that $q = \mathrm{index}(\boldsymbol{r}_0, A) = 0$, if we replace the proper inverse by the Drazin inverse. In particular, the MR residual polynomials can be characterized in terms of Drazin inverse A^{D} (instead of the subspace inverse) using the same line of argument as in Remark 4.1.1.

4.4 Least Squares Solutions in Krylov Spaces

As we have seen so far, the Drazin inverse is closely connected with Krylov subspace methods. In practice a least squares solution is often desired. Therefore we want to investigate now under which conditions a Krylov subspace method leads to such a solution. We have already described this situation in Proposition 3.5.11 and Remark 3.5.10 for general correction spaces: If (3.1) is consistent, a (least squares) solution is generated by a Krylov method if and only if $\operatorname{index}(\boldsymbol{r}_0, A) = 0$ (see also Proposition 3.4.2, Theorem 4.3.5 and Corollary 3.5.6). If (3.1) is inconsistent and the mth Krylov space contains a least squares solution, the MR method makes no progress in the subsequent steps up to the singular breakdown in step $L > m$. An extremal example for this behavior is provided by the backward shift operator with the first unit vector as right hand side and the all zero vector as initial guess (cf. [7, Example 1.2]).

A natural question is, under which conditions the possible Krylov correction $A^{\mathrm{D}}\boldsymbol{r}_0$ in the Lth step represents a least squares solution and thus will be generated by an MR method. We will answer this later in this section.

In [7] it is shown that GMRES yields in the least squares solution of (3.1) for all right hand sides $\boldsymbol{b}$ and all initial vectors $\boldsymbol{x}$ if and only if $\mathcal{N}(A) = \mathcal{N}(A^*)$. We cite from Campbell and Meyer several characterizations of this situation (cf. [14, Definition 4.3.1 and Theorems 4.3.1 and 7.3.4]).

Theorem 4.4.1. *Suppose $A \in \mathbb{C}^{n\times n}$. Then the following statements are equivalent:*

(i) $A^\dagger = A^{\mathrm{D}}$

(ii) $A^\dagger A = AA^\dagger$

(iii) $\mathcal{R}(A) = \mathcal{R}(A^*)$

(iv) $\mathcal{R}(A) \perp \mathcal{N}(A)$ *(and $\mathbb{C}^n = \mathcal{R}(A) \oplus \mathcal{N}(A)$)*

(v) *There exists an unitary matrix U and an nonsingular matrix $C \in \mathbb{C}^{r\times r}$, where r is the rank of A, such that*

$$A = U \begin{bmatrix} C & O \\ O & O \end{bmatrix} U^*$$

Remark 4.4.2. A matrix with the property (iii) is called *range hermitian*. It is easy to see, that this is equivalent to the condition $\mathcal{N}(A) = \mathcal{N}(A^*)$ of Brown and Walker. Note also, that (iv) can be rewritten as

$$\mathcal{R}(A)^\perp = \mathcal{N}(A),$$

which is used by Hayami in [34].

With regard to condition (iv) we note that $\mathcal{R}(A) \perp \mathcal{N}(A)$ alone implies that $\operatorname{index}(A) \leq 1$. This may be seen directly or one observes that the orthogonality implies $\mathcal{R}(A) \cap \mathcal{N}(A) = \{\boldsymbol{0}\}$ and uses Proposition 4.2.11 below. The direct decomposition is then provided by (4.12).

From the condition $\operatorname{index}(A) \leq 1$ we further conclude that the Drazin and, by (i), also the Moore-Penrose inverse coincide with the group inverse (cf. Section 2), i. e., $A^\dagger = A^{\mathrm{D}} = A^\#$.

Now, if (3.1) is consistent it follows immediately that $\boldsymbol{r}_0 \in \mathcal{R}(A)$ for every choice of $\boldsymbol{x}_0$ and this means, by (4.15), that the index q of $\boldsymbol{r}_0$ with respect to A is zero. Applying Theorem 4.3.5 shows that any Krylov method breaks down regularly with the solution

$$\boldsymbol{x}_L^{\mathrm{MR}} = \boldsymbol{x}_L^{\mathrm{OR}} = \boldsymbol{x}_0 + A^\# \boldsymbol{r}_0 = A^\# \boldsymbol{b} + (I - A^\# A)\boldsymbol{x}_0 = A^\dagger \boldsymbol{b} + P_{\mathcal{N}(A)} \boldsymbol{x}_0. \tag{4.21}$$

If, on the other hand, the equation is inconsistent, we have $q = 1$ for every initial guess $\boldsymbol{x}_0$ and an MR Krylov methods breaks down singularly in step L. Since $A^\dagger \boldsymbol{r}_0 = A^{\mathrm{D}} \boldsymbol{r}_0 \in \mathcal{K}_L(A, \boldsymbol{r}_0) = \mathcal{C}_L$ we have

$$A A^\dagger \boldsymbol{r}_0 = P_{\mathcal{R}(A)} \boldsymbol{r}_0 \in A\mathcal{K}_L(A, \boldsymbol{r}_0) = \mathcal{W}_L.$$

Moreover, as stated in Remark 3.5.10, there exists at least one index $m < L$ for which $P_{\mathcal{R}(A)} \boldsymbol{r}_0 \in \mathcal{W}_m = A\mathcal{K}_m(A, \boldsymbol{r}_0)$ and thus an MR subspace correction method yields a least squares solution. By Lemma 3.2.6 the correction is of the form $\boldsymbol{c}_m^{\mathrm{MR}} = A^\dagger \boldsymbol{r}_0 + \tilde{\boldsymbol{z}}$ with $\tilde{\boldsymbol{z}} \in \mathcal{N}(A)$ and the corresponding least squares solution is

$$\boldsymbol{x}_m^{\mathrm{MR}} = A^\dagger \boldsymbol{b} + \boldsymbol{z} \qquad \text{with } \boldsymbol{z} \in \mathcal{N}(A).$$

In the terminating step L we can freely choose the nullspace component $\boldsymbol{z}$ since $\mathcal{N}(A) \cap \mathcal{K}_L(A, \boldsymbol{r}_0) \neq \{\boldsymbol{0}\}$. Using (3.36) we may compute the subspace inverse solution. Setting $\boldsymbol{c}_L^{\mathrm{MR}} = A^\# \boldsymbol{r}_0 = A^\dagger \boldsymbol{r}_0$ leads to

$$\boldsymbol{x}_L^{\mathrm{MR}} = A^\# \boldsymbol{b} + (I - A^\# A)\boldsymbol{x}_0 = A^\dagger \boldsymbol{b} + P_{\mathcal{N}(A)} \boldsymbol{x}_0,$$

which coincides with the MR solution for a consistent problem. For $\boldsymbol{x}_0 \in \mathcal{R}(A) \perp \mathcal{N}(A)$ we get the pseudoinverse solution $A^\dagger \boldsymbol{b}$ in both cases.

Note that our considerations prove one direction in the result of Brown and Walker: If A is range hermitian, an MR Krylov method yields a (least squares) solution for arbitrary vectors $\boldsymbol{b}$ and $\boldsymbol{x}_0$. The other implication is provided by a counterexample.

Example 4.4.3. Consider the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 3}.$$

Note first that $A^2 = A$, in particular $\operatorname{index}(A) = 1$, and $A = A^{\mathrm{D}} = A^{\#}$. Moreover,

$$A^{\dagger} = \begin{bmatrix} 2/3 & 1/3 & 1/3 \\ 0 & 0 & 0 \\ 1/3 & -1/3 & 2/3 \end{bmatrix}$$

and $\mathcal{N}(A) = \operatorname{span}\left\{\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^{\top}\right\}$. Obviously, the right hand side $\boldsymbol{b} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^{\top}$ does not lie in the range of A. We apply a Krylov method with the zero vector as initial guess. In the second step the Arnoldi process breaks down singularly and $\boldsymbol{c}_2^{\mathrm{MR}} = \boldsymbol{c}_1^{\mathrm{MR}} + \boldsymbol{z}$ with $\boldsymbol{z} \in \mathcal{K}_2(A, \boldsymbol{b}) \cap \mathcal{N}(A) = \mathcal{N}(A)$. The first MR correction is $\boldsymbol{c}_1^{\mathrm{MR}} = \alpha \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^{\top}$ where α minimizes

$$\left\| \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - A(\alpha \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}) \right\| = \left\| \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \alpha \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\|,$$

i. e., $\alpha = 1$ and the residual norm is 1. This is not a least squares solution since $A^{\dagger}\boldsymbol{b} = A^{\dagger}\boldsymbol{r}_0 = \begin{bmatrix} 4/3 & 0 & 2/3 \end{bmatrix}^{\top}$ has residual norm $1/\sqrt{3}$. For the second MR approximation, we can minimize the norm of

$$\|\boldsymbol{x}_2^{\mathrm{MR}}\| = \|\boldsymbol{c}_2^{\mathrm{MR}}\| = \|\boldsymbol{x}_1^{\mathrm{MR}} - \gamma \boldsymbol{z}\| \left\| \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \gamma \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\|,$$

which is done by $\gamma = 1$ and results in $\boldsymbol{c}_2^{\mathrm{MR}} = A^{\mathcal{K}_2(A,\boldsymbol{b})}\boldsymbol{b} = A^{\mathrm{D}}\boldsymbol{b}$, i. e., the subspace inverse solution coincide with the Drazin inverse solution.

But also this need not to be true in general as $\tilde{\boldsymbol{b}} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^{\top} \notin \mathcal{R}(A)$ shows. Again we get a singular breakdown after two steps. Now $\boldsymbol{c}_1^{\mathrm{MR}} = \begin{bmatrix} 1/2 & 0 & 0 \end{bmatrix}^{\top}$ and this is also the solution with minimal norm after the second step, but $A^{\mathrm{D}}\tilde{\boldsymbol{b}} = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^{\top}$ and $A^{\dagger}\tilde{\boldsymbol{b}} = \begin{bmatrix} 2/3 & 0 & 1/3 \end{bmatrix}^{\top}$.

This example shows that, if A is not range hermitian, an (MR) Krylov method may converge neither to a least squares solution nor to the Drazin inverse solution – even if the index of the operator is one.

Though, the case $\operatorname{index}(A) = 1$ deserves closer attention. Of course, the index is zero if and only if A is nonsingular. Thus, stating $\operatorname{index}(A) = 1$ is equivalent to demand $\operatorname{index}(A) \leq 1$ for singular A.

Proposition 4.4.4. *The following statements are equivalent:*

(i) $\operatorname{index}(A) \leq 1$,

(ii) $\mathcal{R}(A) \oplus \mathcal{N}(A) = \mathcal{H}$,

(iii) $\mathcal{R}(A) \cap \mathcal{N}(A) = \{\boldsymbol{0}\}$,

(iv) $A^{\#}$ *exists.*

Proof. For the equivalence of (i) and (iv) we refer to [14, Definition 7.2.4 and Theorem 7.2.5].

If A is nonsingular, the assertion is trivial since $\mathcal{N}(A) = \{\boldsymbol{0}\}$, $\mathcal{R}(A) = \mathcal{H}$ and $\text{index}(A) = 0$. Therefore we assume A is singular or, equivalently, $\text{index}(A) \geq 1$.

The implications (i)⇒(ii)⇒(iii) are direct consequences of (4.12).

Now, suppose (iii) and denote by d the index of A. From Lemma 4.2.9 we obtain

$$\mathcal{R}(A) = \mathcal{R}(A^d) \oplus \mathcal{N}(A^{d-1}).$$

The intersections of $\mathcal{N}(A)$ with each component of this direct decomposition are assumed to be the trivial space containing only $\boldsymbol{0}$. For $\mathcal{R}(A^d)$ this is evidently true (regardless of $d = 1$ or greater). Thus, the only restriction for d arise from $\mathcal{N}(A^{d-1}) \cap \mathcal{N}(A) = \{\boldsymbol{0}\}$. If $d = 1$, even this is evident. On the other hand, if $d > 1$ there holds $\mathcal{N}(A) \subseteq \mathcal{N}(A^{d-1})$ and we have a contradiction. □

If A is not range hermitian but $\text{index}(A) = 1$, a similar result as in (4.21) is implied along the same line of argument if we replace $P_{\mathcal{N}(A)}$ by the oblique projection $P_{\mathcal{N}(A),\mathcal{R}(A)}$.

Proposition 4.4.5. *If* $\text{index}(A) = 1$ *and* $\boldsymbol{b} \in \mathcal{R}(A)$, *then no singular breakdown occurs and there holds*

$$\boldsymbol{x}_L^{\text{MR}} = \boldsymbol{x}_L^{\text{OR}} = \boldsymbol{x}_0 + A^{\#}\boldsymbol{r}_0 = A^{\#}\boldsymbol{b} + (I - A^{\#}A)\boldsymbol{x}_0 = A^{\#}\boldsymbol{b} + P_{\mathcal{N}(A),\mathcal{R}(A)}\boldsymbol{x}_0,$$

which is a solution of (3.1), *i. e.,* $A\boldsymbol{x}_L^{\text{MR}} = A\boldsymbol{x}_L^{\text{OR}} = \boldsymbol{b}$.

This was first observed by Freund and Hochbruck in [29, Corollary 3] in the context of the QMR method. Similar results, though without stating the representation of the solution in terms of $A^{\#}$, are provided by Brown and Walker for GMRES (cf. [7, Theorem 2.6]) and by Hayami for the conjugate residual (CR) method (cf. [34, Theorem 3.21]) and GCR (cf. [35]).

Assume, a Krylov subspace method yields a solution of (3.1) for *any* initial guess $\boldsymbol{x}_0$, then for some m there holds $\boldsymbol{r}_0 \in \mathcal{W}_m = A\mathcal{K}_m(A, \boldsymbol{r}_0)$ and the method breaks down regularly (cf. Proposition 3.4.3). By Theorem 4.3.5 this is equivalent to $\boldsymbol{r}_0 \in \mathcal{R}(A^d)$, where $d = \text{index}(A)$ and $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$. This must hold true for arbitrary $\boldsymbol{x}_0$. By choosing $\boldsymbol{x}_0 = \boldsymbol{0}$ we conclude $\boldsymbol{b} \in \mathcal{R}(A^d)$. Thus, for any $\boldsymbol{x}_0$ we have $A\boldsymbol{x}_0 \in \mathcal{R}(A^d)$.

We decompose $\boldsymbol{x}_0 = \boldsymbol{s} + \boldsymbol{t}$ subject to $\boldsymbol{s} \in \mathcal{R}(A^d)$ and $\boldsymbol{t} \in \mathcal{N}(A^d)$. Then $A\boldsymbol{x}_0 = A\boldsymbol{s} + A\boldsymbol{t} = A\boldsymbol{s}$, i. e., $A\boldsymbol{t} = 0$. This applies for any $\boldsymbol{t} \in \mathcal{N}(A^d)$, which is only possible if $d = 1$ or $d = 0$.

Thus we have proved: If an (MR or OR) Krylov method yields a solution of (3.1) for any initial guess $\boldsymbol{x}_0$, then $\text{index}(A) \leq 1$ and $\boldsymbol{b} \in \mathcal{R}(A)$. Vice versa, if $\text{index}(A) \leq 1$ and $\boldsymbol{b} \in \mathcal{R}(A)$, we already know from Proposition 4.4.5 that the Krylov method determines a solution regardless of $\boldsymbol{x}_0$. Let us summarize:

Theorem 4.4.6. *An MR or OR Krylov method yields a solution of* (3.1) *for any initial guess* $\boldsymbol{x}_0$, *if and only if* $\text{index}(A) \leq 1$ *and* $\boldsymbol{b} \in \mathcal{R}(A)$.

Note the difference between our result and that in [7, Theorem 2.4]: We ask for proper solutions, which requires consistency, while Brown and Walker seek for a least squares solution.

Theorem 4.4.1 characterizes the matrices for which $A^\dagger = A^\mathrm{D}$, in other words, it answers the question when there holds

$$A^\dagger \boldsymbol{r} = A^\mathrm{D} \boldsymbol{r} \tag{4.22}$$

for any vector $\boldsymbol{r}$. We now look for a characterization of the vectors for which, given the matrix A, the relation (4.22) is satisfied. We first assume $\boldsymbol{r} \in \mathcal{R}(A)$, i. e., $\boldsymbol{r} = A\boldsymbol{y}$ for some $\boldsymbol{y} \in \mathbb{C}^n$. Using (4.22) and $AA^\dagger A = A$ we obtain

$$\boldsymbol{r} = A\boldsymbol{y} = AA^\dagger A\boldsymbol{y} = AA^\dagger \boldsymbol{r} = AA^\mathrm{D}\boldsymbol{r},$$

which shows that $\boldsymbol{r} \in \mathcal{R}(A^d)$, where $d = \mathrm{index}(A)$. Similarly, for $\boldsymbol{r} \in \mathcal{R}(A)^\perp = \mathcal{N}(A^*) = \mathcal{N}(A^\dagger)$ we see from

$$(I - AA^\mathrm{D})\boldsymbol{r} = \boldsymbol{r} - AA^\mathrm{D}\boldsymbol{r} = \boldsymbol{r} - AA^\dagger \boldsymbol{r} = \boldsymbol{r}$$

that $\boldsymbol{r} \in \mathcal{N}(A^d)$.

Proposition 4.4.7. *Suppose $\boldsymbol{r}$ satisfies (4.22) and $d = \mathrm{index}(A)$. Then there holds*

$$P_{\mathcal{R}(A)}\boldsymbol{r} = P_{\mathcal{R}(A^d),\mathcal{N}(A^d)}\boldsymbol{r} \qquad \textit{(or equivalently } P_{\mathcal{N}(A^*)}\boldsymbol{r} = P_{\mathcal{N}(A^d),\mathcal{R}(A^d)}\boldsymbol{r}\textit{)} \tag{4.23}$$

Proof. As seen above we have $\boldsymbol{s} := P_{\mathcal{R}(A)}\boldsymbol{r} \in \mathcal{R}(A^d)$ and $\boldsymbol{t} := P_{\mathcal{N}(A^*)}\boldsymbol{r} \in \mathcal{N}(A^d)$. Thus we have established a decomposition $\boldsymbol{r} = \boldsymbol{s} + \boldsymbol{t}$ into the components on $\mathcal{R}(A^d)$ and $\mathcal{N}(A^d)$. Since $\mathcal{R}(A^d) \oplus \mathcal{N}(A^d) = \mathbb{C}^n$ (cf. (4.12)) this decomposition is unique and the components are given by the oblique projection $P_{\mathcal{R}(A^d),\mathcal{N}(A^d)} = AA^\mathrm{D}$ of $\boldsymbol{r}$ onto $\mathcal{R}(A^d)$ along $\mathcal{N}(A^d)$ and its complementary projection $I - AA^\mathrm{D} = P_{\mathcal{N}(A^d),\mathcal{R}(A^d)}$ (cf. Lemma 4.2.8). □

Proposition 4.4.7 implies, that (4.22) is sufficient for $A^\mathrm{D}\boldsymbol{r}_0$ to be a least squares solution of the residual equation, i. e., for $\boldsymbol{x}_0 + A^\mathrm{D}\boldsymbol{r}_0$ to be a least square solution of (3.1).

Proposition 4.4.8. *The correction $\boldsymbol{c}_L^\mathrm{MR} = A^\mathrm{D}\boldsymbol{r}$ is a least squares solution of $A\boldsymbol{c} = \boldsymbol{r}$ and is thus generated by a MR Krylov method in the termination step L if and only if*

$$AA^\mathrm{D}\boldsymbol{r} = AA^\dagger \boldsymbol{r} \qquad \textit{(or equivalently } \left(A^\mathrm{D} - A^\dagger\right)\boldsymbol{r} \in \mathcal{N}(A)\textit{)} \tag{4.24}$$

Proof. Note that (4.24) is merely another formulation for (4.23). A vector $\boldsymbol{c}^\mathrm{MR}$ is a least squares solution of $A\boldsymbol{c} = \boldsymbol{r}$ if and only if $\|P_{\mathcal{R}(A)}\boldsymbol{r} - A\boldsymbol{c}^\mathrm{MR}\| = 0$ or, equivalently, if and only if $P_{\mathcal{R}(A)}\boldsymbol{r} = AA^\dagger \boldsymbol{r} = A\boldsymbol{c}^\mathrm{MR}$. Using the last equation, we get: If $\boldsymbol{c}_L^\mathrm{MR} = A^\mathrm{D}\boldsymbol{r}$ is a least squares solution, then there holds (4.24) (and (4.23)).

If, on the other hand, $AA^\mathrm{D}\boldsymbol{r} = AA^\dagger \boldsymbol{r}$ holds true, we choose $\boldsymbol{c}^\mathrm{MR} = A^\mathrm{D}\boldsymbol{r}$, which is possible in the termination step L. A simple calculation shows $A\boldsymbol{c}^\mathrm{MR} = AA^\mathrm{D}\boldsymbol{r} = AA^\dagger \boldsymbol{r} = P_{\mathcal{R}(A)}\boldsymbol{r}$, which implies that $\boldsymbol{c}^\mathrm{MR}$ is a least squares solution. □

Another way to generalize the results of Brown and Walker in [7] is to investigate under which conditions a Krylov method constructs a least squares solution up to some projection. The close connection of Krylov spaces and the Drazin inverse suggests the projection $P_{\mathcal{R}(A^d),\mathcal{N}(A^d)} = A^{\mathrm{D}}A$.

Proposition 4.4.9. *Let $\mathcal{H}$ be finite dimensional. If $\mathcal{R}(A^d) \perp \mathcal{N}(A^d)$, there holds*

$$A^{\mathrm{D}} = A^{\dagger}AA^{\mathrm{D}} = A^{\mathrm{D}}AA^{\dagger}. \tag{4.25}$$

Remark 4.4.10. Note, that the identities in (4.25) can be rewritten in several ways using $P_{\mathcal{R}(A^d),\mathcal{N}(A^d)} = A^{\mathrm{D}}A = AA^{\mathrm{D}}$, $P_{\mathcal{R}(A)} = AA^{\dagger}$ and $P_{\mathcal{R}(A^*)} = A^{\dagger}A$. In particular, (4.25) is equivalent to $A^{\mathrm{D}} = P_{\mathcal{R}(A^d),\mathcal{N}(A^d)}A^{\dagger} = A^{\dagger}P_{\mathcal{R}(A^d),\mathcal{N}(A^d)}$, which means that the Drazin and the Moore-Penrose inverse coincide up to the projection $P_{\mathcal{R}(A^d),\mathcal{N}(A^d)}$, which can be seen as a generalization of (i) in Theorem 4.4.1.

Note further that the orthogonality assumption implies that $P_{\mathcal{R}(A^d),\mathcal{N}(A^d)} = A^{\mathrm{D}}A$ is the orthogonal projection $P_{\mathcal{R}(A^d)}$.

Proof. Let U_C be an orthonormal basis of $\mathcal{R}(A^d) \subseteq \mathcal{R}(A)$. We can extend U_C to an orthonormal basis $W = \begin{bmatrix} U_C & U_W \end{bmatrix}$ of $\mathcal{R}(A)$. Since $\mathcal{H} = \mathcal{R}(A) \oplus \mathcal{N}(A^*)$ and $\mathcal{R}(A) \perp \mathcal{N}(A^*)$, we can further extend this to an orthonormal basis $\begin{bmatrix} U_C & U_W & Y \end{bmatrix}$ of $\mathcal{H}$, where Y is a basis of $\mathcal{N}(A^*)$. We set $U_Y := \begin{bmatrix} U_W & Y \end{bmatrix}$. Obviously, $\mathcal{H} = \operatorname{span}\{U_C\} \oplus \operatorname{span}\{U_Y\}$ is an orthogonal decomposition of $\mathcal{H}$ and, since U_C is a basis of $\mathcal{R}(A^d)$, its orthogonal complement $\mathcal{N}(A^d)$ is spanned by the basis U_Y.

Now, denote by Z an orthonormal basis of $\mathcal{N}(A) \subseteq \mathcal{N}(A^d)$ and extend it to an orthonormal basis $U_Z = \begin{bmatrix} U_V & Z \end{bmatrix}$ of $\mathcal{N}(A^d)$. Since $\mathcal{R}(A^d) \perp \mathcal{N}(A^d)$ we may join U_C with U_Z to obtain another orthonormal basis $\begin{bmatrix} U_C & U_Z \end{bmatrix} = \begin{bmatrix} U_C & U_V & Z \end{bmatrix}$ of $\mathcal{H}$. We set $V := \begin{bmatrix} U_C & U_V \end{bmatrix}$. Obviously, $\mathcal{H} = \operatorname{span}\{V\} \oplus \operatorname{span}\{Z\}$ is an orthogonal decomposition of $\mathcal{H}$. But $\operatorname{span}\{Z\}$ is the nullspace of A and thus the orthogonal complement has to coincide with $\mathcal{R}(A^*)$. In other words, V is an orthonormal basis of $\mathcal{R}(A^*)$.

The remainder of the proof will be done by observing, how the operators act on these orthonormal bases. Note first, that

$$\begin{aligned} A\begin{bmatrix} U_C & U_Y \end{bmatrix} &= \begin{bmatrix} U_C & U_Y \end{bmatrix}\begin{bmatrix} C & O \\ O & N_Y \end{bmatrix} \\ A\begin{bmatrix} U_C & U_Z \end{bmatrix} &= \begin{bmatrix} U_C & U_Z \end{bmatrix}\begin{bmatrix} C & O \\ O & N_Z \end{bmatrix} \end{aligned} \tag{4.26}$$

and

$$\begin{aligned} A^{\mathrm{D}}\begin{bmatrix} U_C & U_W & Y \end{bmatrix} &= \begin{bmatrix} U_C & U_W & Y \end{bmatrix}\begin{bmatrix} C^{-1} & O & O \\ O & O & O \\ O & O & O \end{bmatrix} = \begin{bmatrix} U_C C^{-1} & O & O \end{bmatrix} \\ A^{\mathrm{D}}\begin{bmatrix} U_C & U_V & Z \end{bmatrix} &= \begin{bmatrix} U_C & U_V & Z \end{bmatrix}\begin{bmatrix} C^{-1} & O & O \\ O & O & O \\ O & O & O \end{bmatrix} = \begin{bmatrix} U_C C^{-1} & O & O \end{bmatrix} \end{aligned}$$

Since $AA^\dagger$ is the orthogonal projection onto $\mathcal{R}(A)$ it follows that

$$AA^\dagger \begin{bmatrix} W & Y \end{bmatrix} = \begin{bmatrix} W & O \end{bmatrix} = \begin{bmatrix} U_C & U_W & Y \end{bmatrix} \begin{bmatrix} I & O & O \\ O & I & O \\ O & O & O \end{bmatrix},$$

which leads to

$$A^{\mathrm{D}} AA^\dagger \begin{bmatrix} U_C & U_W & Y \end{bmatrix} = A^{\mathrm{D}} \begin{bmatrix} U_C & U_W & Y \end{bmatrix} \begin{bmatrix} I & O & O \\ O & I & O \\ O & O & O \end{bmatrix} = \quad \cdots$$

$$\begin{bmatrix} U_C C^{-1} & O & O \end{bmatrix} \begin{bmatrix} I & O & O \\ O & I & O \\ O & O & O \end{bmatrix} = \begin{bmatrix} U_C C^{-1} & O & O \end{bmatrix} = A^{\mathrm{D}} \begin{bmatrix} U_C & U_W & Y \end{bmatrix}.$$

Similarly, $A^\dagger A = P_{\mathcal{R}(A^*)}$ implies

$$A^\dagger A \begin{bmatrix} V & Z \end{bmatrix} = \begin{bmatrix} U_C & U_V & O \end{bmatrix}$$

and

$$A^\dagger AA^{\mathrm{D}} \begin{bmatrix} U_C & U_V & Z \end{bmatrix} = A^\dagger A \begin{bmatrix} U_C & U_V & Z \end{bmatrix} \begin{bmatrix} C^{-1} & O & O \\ O & O & O \\ O & O & O \end{bmatrix} = \quad \cdots$$

$$\begin{bmatrix} U_C & U_V & O \end{bmatrix} \begin{bmatrix} C^{-1} & O & O \\ O & O & O \\ O & O & O \end{bmatrix} = \begin{bmatrix} U_C C^{-1} & O & O \end{bmatrix} = A^{\mathrm{D}} \begin{bmatrix} U_C & U_V & Z \end{bmatrix},$$

which completes the proof. □

Using these orthonormal bases, A can be written as

$$A \begin{bmatrix} U_C & U_V & Z \end{bmatrix} = \begin{bmatrix} U_C & U_W & Y \end{bmatrix} \begin{bmatrix} C & O & O \\ O & M & O \\ O & O & O \end{bmatrix}, \tag{4.27}$$

where $C \in \mathbb{C}^{r\times r}$ and $M \in \mathbb{C}^{s\times s}$ are nonsingular, $r = \operatorname{rank}(A^d)$ and $s = \operatorname{rank}(A) - r$. The zero blocks in the last column are a simple consequence of $\operatorname{span}\{Z\} = \mathcal{N}(A)$. That the last row is zero can be seen from the adjoint equation together with $\operatorname{span}\{Y\} = \mathcal{N}(A^*)$. Since $\mathcal{R}(A^d)$ is A-invariant, there holds $AU_C = U_C C$, which shows that the $(2,1)$ block is zero. Similarly, the A-invariance of $\mathcal{N}(A^d)$ implies $AU_V = U_W M + YB$ and the $(1,2)$ block has to be zero, since U_W and Y are orthogonal to U_C. The representation (4.27) can be regarded as a generalization of (v) in Theorem 4.4.1.

Proposition 4.4.9 provides the last equality in

Proposition 4.4.11. *If $\mathcal{R}(A^d) \perp \mathcal{N}(A^d)$, then for any right hand side $\boldsymbol{b}$ and any initial guess $\boldsymbol{x}_0$ the Lth MR correction $\boldsymbol{c}_L^{\mathrm{MR}}$ with respect to $\mathcal{K}_L(A, \boldsymbol{r}_0)$, $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$ satisfies*

$$P_{\mathcal{R}(A^d)}\boldsymbol{c}_L^{\mathrm{MR}} = A^{\mathrm{D}}A\boldsymbol{c}_L^{\mathrm{MR}} = A^{\mathrm{D}}\boldsymbol{r}_0 = P_{\mathcal{R}(A^d)}A^{\dagger}\boldsymbol{r}_0.$$

Proof. The Lth MR correction minimizes

$$\|\boldsymbol{r}_0 - A\boldsymbol{c}\| = \|P_{\mathcal{R}(A^d)}\boldsymbol{r}_0 - AP_{\mathcal{R}(A^d)}\boldsymbol{c}\| + \|P_{\mathcal{N}(A^d)}\boldsymbol{r}_0 - AP_{\mathcal{N}(A^d)}\boldsymbol{c}\|$$

over all $\boldsymbol{c} \in \mathcal{K}_L(A, \boldsymbol{r}_0)$. Due to the decomposition (4.18) and the orthogonality condition this can be done independently for the components in the range and nullspace of A^d. But $\|P_{\mathcal{R}(A^d)}\boldsymbol{r}_0 - AP_{\mathcal{R}(A^d)}\boldsymbol{c}\| = 0$ for $P_{\mathcal{R}(A^d)}\boldsymbol{c} = A^{\mathrm{D}}\boldsymbol{r}_0 \in \mathcal{K}_L(A, \boldsymbol{r}_0) \cap \mathcal{R}(A^d)$. □

The identities in (4.25) can be proved under much weaker conditions.

Proposition 4.4.12. *There holds*

(i) $A^{\mathrm{D}} = A^{\dagger}AA^{\mathrm{D}} \Leftrightarrow \mathcal{R}(A^d) \subseteq \mathcal{R}(A^*)$,

(ii) $A^{\mathrm{D}} = A^{\mathrm{D}}AA^{\dagger} \Leftrightarrow \mathcal{N}(A^*) \subseteq \mathcal{N}(A^d)$.

Proof. (i) is an immediate consequence of $\mathcal{R}(A^{\mathrm{D}}) = \mathcal{R}(A^d)$ and $A^{\dagger}A = P_{\mathcal{R}(A^*)}$. Similarly, $\mathcal{R}(I - AA^{\dagger}) = \mathcal{N}(A^*) \subseteq \mathcal{N}(A^d)$ yields $A^{\mathrm{D}}(I - AA^{\dagger}) = O$ and vice versa. The latter is a reformulation of the left term in (ii). □

Since $\mathcal{R}(A^{\mathrm{D}}) = \mathcal{R}(A^d) \subseteq \mathcal{R}(A) = \mathcal{R}(AA^{\dagger})$ and $\mathcal{R}(I - A^{\dagger}A) = \mathcal{N}(A) \subseteq \mathcal{N}(A^d) = \mathcal{N}(A^{\mathrm{D}})$ for any A, there holds

$$A^{\mathrm{D}} = A^{\mathrm{D}}A^{\dagger}A = AA^{\dagger}A^{\mathrm{D}}.$$

Together with (i) of the above proposition, we conclude that $\mathcal{R}(A^d) \subseteq \mathcal{R}(A^*)$ implies $AA^{\dagger}A^{\mathrm{D}} = A^{\dagger}AA^{\mathrm{D}}$, i.e., A commutes with its pseudoinverse on $\mathcal{R}(A^d)$ (which is a weakened version of (ii) in Theorem 4.4.1). Similarly, for any vector $\boldsymbol{v} \in \mathcal{H}$, the difference $AA^{\dagger}\boldsymbol{v} - A^{\dagger}A\boldsymbol{v}$ lies in $\mathcal{N}(A^d)$ if $\mathcal{N}(A^*) \subseteq \mathcal{N}(A^d)$.

An example for a matrix A which satisfies both conditions in Proposition 4.4.12, i. e.,

$$\mathcal{R}(A^d) \subseteq \mathcal{R}(A^*) \qquad \text{and} \qquad \mathcal{N}(A^*) \subseteq \mathcal{N}(A^d), \tag{4.28}$$

but violates the orthogonality assumption in Propositions 4.4.9 and 4.4.11,

$$\mathcal{R}(A^d) \perp \mathcal{N}(A^d), \qquad d = \mathrm{index}(A), \tag{4.29}$$

is not obvious. This is impossible in three dimensions, as may be verified by a short exercise (but see also below). We give here general construction guidelines and illustrate them by a concrete example. Before we start some observations should be mentioned.

First, (4.29) implies (4.28). This follows from a combination of Propositions 4.4.9 and 4.4.11 or may be seen from the bases constructed in the poof of Proposition 4.4.9.

Now, assume that (4.28) is satisfied and that actually equality holds in one relation, e. g. $\mathcal{N}(A^*) = \mathcal{N}(A^d)$. Then $\dim \mathcal{N}(A^d) = \dim \mathcal{N}(A^*) = \dim \mathcal{N}(A)$, and, since $\mathcal{N}(A) \subseteq \mathcal{N}(A^d)$, this implies $\mathcal{N}(A) = \mathcal{N}(A^d)$, which means that $d = \operatorname{index}(A) \leq 1$. But then using the second inclusion we get

$$\mathcal{N}(A^*)^\perp = \mathcal{R}(A) \subseteq \mathcal{R}(A^*) \perp \mathcal{N}(A) = \mathcal{N}(A^*),$$

which establishes (4.29) for $d = 1$. For $\mathcal{R}(A^d) = \mathcal{R}(A^*)$ we conclude similarly $d \leq 1$ and

$$\mathcal{R}(A)^\perp = \mathcal{N}(A^*) \subseteq \mathcal{N}(A) \perp \mathcal{R}(A^*) = \mathcal{R}(A).$$

In either case the remaining inclusion must be an equality. In other words, if one of the inclusions in (4.28) becomes an equality, then there holds equality in both relations, index of A is one and (4.29) is satisfied.

Remark 4.4.13. In view of (4.7), we start with the decomposition $TA = T \left[\begin{smallmatrix} C & O \\ O & N \end{smallmatrix}\right]$, where $C \in \mathbb{C}^{r\times r}$ is nonsingular, $r = \operatorname{rank}(A^d)$, $d = \operatorname{index}(A)$, and $N \in \mathbb{C}^{(n-r)\times(n-r)}$ is nilpotent of index d. We make no further initial assumption on C and N. At first we choose the nullspace of A^*. Let $Y = \begin{bmatrix} \boldsymbol{y}_1 & \dots & \boldsymbol{y}_t \end{bmatrix}$ denote an orthonormal basis of $\mathcal{N}(A^*)$. From its orthogonal complement $\mathcal{R}(A)$, we choose an orthonormal basis U of an r-dimensional subspace which becomes the range of A^d. We want to construct the basis T such that $T = \begin{bmatrix} U & X & Y \end{bmatrix}$, where $\begin{bmatrix} X & Y \end{bmatrix}$ forms a basis of $\mathcal{N}(A^d)$. This construction ensures that the condition $\mathcal{N}(A^*) \subset \mathcal{N}(A^d)$ is satisfied (the inclusion has to be proper, since otherwise (4.29) holds).

For later use, we note that, given Y and U as above, one can choose a basis $V = \begin{bmatrix} \boldsymbol{v}_1 & \dots & \boldsymbol{v}_s \end{bmatrix}$ with $s = n - r - t$ such that $\begin{bmatrix} U & V \end{bmatrix}$ is an orthonormal basis of $\mathcal{R}(A)$ and thus, $\begin{bmatrix} U & V & Y \end{bmatrix}$ is an orthonormal basis of $\mathcal{H}$, i. e. $\begin{bmatrix} U & V & Y \end{bmatrix}^* \begin{bmatrix} U & V & Y \end{bmatrix} = I$.

An arbitrary vector $\boldsymbol{y} \in \mathcal{N}(A^*)$ is of the form $\boldsymbol{y} = Y\boldsymbol{g}$ with $\boldsymbol{g} \in \mathbb{C}^t$. Using $U^*Y = O$ we obtain[2]

$$\boldsymbol{0} = T^*A^*\boldsymbol{y} = (AT)^*\boldsymbol{y} = \begin{bmatrix} C^{\mathrm{H}} & O \\ O & N^{\mathrm{H}} \end{bmatrix} \begin{bmatrix} U^* \\ X^* \\ Y^* \end{bmatrix} Y\boldsymbol{g} = \begin{bmatrix} C^{\mathrm{H}} & O \\ O & N^{\mathrm{H}} \end{bmatrix} \begin{bmatrix} \boldsymbol{0} \\ X^*Y\boldsymbol{g} \\ Y^*Y\boldsymbol{g} \end{bmatrix}.$$

Since $\boldsymbol{g}$ is arbitrary, we conclude that the t columns of $\begin{bmatrix} X & Y \end{bmatrix}^* Y$ lie in the nullspace of N^{H}. This null space must have the same dimension t as the nullspace of A^*. Therefore, if $S = \begin{bmatrix} S_X \\ S_Y \end{bmatrix}$ denotes a basis of $\mathcal{N}(N^{\mathrm{H}})$, the freedom in the choice of N and Y is restricted by the compatibility condition $S_Y = Y^*Y$. Additionally, we get the ts equations $S_X = X^*Y$ for the columns of X.

[2] With regard to the notation with * and $^{\mathrm{H}}$, we recall our distinction between quantities belonging to the general Hilbert space and their coordinates with respect to a specific basis, cf. Remark 4.2.7.

Note that these restrictions in the choice of Y and N are not essential: If we choose Y orthonormal as above, this condition reduces to $S_Y = I$. This implies that the last t components of the vectors in $\mathcal{N}(N^{\mathrm{H}})$ can vary arbitrarily. The Jordan canonical form J_N of $N = T_N J_N {T_N}^{-1}$ must contain exactly t Jordan blocks, i. e., there are exactly t columns and rows with all zero components. Thus it is possible to choose T_N as a permutation such that the last t rows of N are zero. In other words, the Jordan structure of N (and A) can be chosen arbitrarily. This defines the number t, and now we choose t orthonormal vectors, which form the columns of Y.

Since $\mathcal{N}(A) \subset \mathcal{N}(A^d) = \operatorname{span}\{X, Y\}$, we can write any vector in the nullspace of A as

$$\boldsymbol{z} = \begin{bmatrix} X & Y \end{bmatrix} \begin{bmatrix} \boldsymbol{g}_X \\ \boldsymbol{g}_Y \end{bmatrix}, \quad \text{where } \boldsymbol{g} = \begin{bmatrix} \boldsymbol{g}_X \\ \boldsymbol{g}_Y \end{bmatrix} \text{ is a zero eigenvector of } N.$$

Let $\{\boldsymbol{g}_1, \ldots, \boldsymbol{g}_t\}$ denote a basis of $\mathcal{N}(N)$, whereas

$$\boldsymbol{g}_j = \begin{bmatrix} \boldsymbol{g}_X^{(j)} \\ \boldsymbol{g}_Y^{(j)} \end{bmatrix} \quad \text{for } j = 1, \ldots, t.$$

The condition $\mathcal{R}(A^d) \subset \mathcal{R}(A^*) = \mathcal{N}(A)^\perp$ implies that the vectors of the basis U are orthogonal to any $\boldsymbol{z} \in \mathcal{N}(A)$, i. e., $\boldsymbol{0} = U^* \begin{bmatrix} X & Y \end{bmatrix} \boldsymbol{g}_j = \begin{bmatrix} U^*X & O \end{bmatrix} \boldsymbol{g}_j = U^* X \boldsymbol{g}_X^{(j)}$ for $j = 1, \ldots, t$. This yields another set of tr equations for X. If we write the vectors $\boldsymbol{g}_X^{(j)}$ as columns in a matrix $G \in \mathbb{C}^{s \times t}$, a short representation of these equations is $U^* X G = O$.

Altogether, we have a system of $tr + ts$ linear equations for the s unknown vectors in $X = \begin{bmatrix} \boldsymbol{x}_1 & \ldots & \boldsymbol{x}_s \end{bmatrix}$. Setting $X = Y S_X^{\mathrm{H}}$, we conclude that the system is solvable. Obviously, this solution violates a third condition on the basis X, which arises from $\mathcal{R}(A^d) \not\perp \mathcal{N}(A^d)$, namely $U^*X \neq O$. If there exists a matrix $F \in \mathbb{C}^{t \times s}$ such that

$$GF = I_s \tag{4.30}$$

any solution X will violate this condition. We will see next, that also the reverse implication holds true.

Note first that (4.30) implies that F is an equation solving generalized inverse of G (since there holds $GFG = F$). Further, by [4, Lemma 1.2], equation (4.30) is equivalent to rank $G = s$.

Remember that $\begin{bmatrix} U & V & Y \end{bmatrix}$ is an orthonormal basis of $\mathcal{H}$. We can express any solution X in terms of this basis. Let $X = Y S_X^{\mathrm{H}} + V + UM$ with $M \in \mathbb{C}^{r \times s}$. Then $S_X = X^*Y$ is satisfied and condition $U^*X \neq O$ is equivalent to $M \neq O$. The matrix equation $U^*XG = O$ reduces to $MG = O$. The set of all solutions of the latter equation is given by

$$\left\{ H(I - GG^{\bar{}}) \mid H \in \mathbb{C}^{r \times s} \right\}.$$

(cf. [14, Corollary 6.3.2]). Thus, a nonzero solution of $MG = O$ exists if and only if $GG^{\bar{}} \neq I$ for all equation solving inverses of G or, equivalently, if and only if rank $G \neq s$.

In fact, this is another condition for N: It must be possible to choose a basis $\tilde{G} = \begin{bmatrix} \boldsymbol{g}_1 & \dots & \boldsymbol{g}_t \end{bmatrix}$ of $\mathcal{N}(N)$ such that the $s \times t$ upper block of $\tilde{G}$ has a rank different from s.

If $s = 1$, the matrix G degenerates to a row vector $\boldsymbol{g}$ and (4.30) holds if we set $F = \boldsymbol{g}/(\boldsymbol{g}, \boldsymbol{g})_2$. This can be done since $\boldsymbol{g} \neq \boldsymbol{0}$, which follows from the fact that $\mathcal{N}(A)$ has to be a proper subset of $\mathcal{N}(A^d)$. Consequently, if $s = 1$, it is not possible to construct a matrix A with the desired properties.

On the other hand, $s > 1$ means that $\dim \mathcal{N}(A^d) \geq \dim \mathcal{N}(A) + 2$. In other words, N must at least consist of a three-dimensional Jordan block or of two two-dimensional Jordan blocks. This explains, why the smallest possible example is four-dimensional.

Evidently, if $t < s$, the rank of $G \in \mathbb{C}^{s \times t}$ is always smaller than s. The example given below is also of this type. It remains an open question whether there exists an example with $t \geq s$. The condition $t < s$ can be formulated more comprehensibly as follows: The nullspace of A^d has to be large in comparison with the nullspace of A. More precisely, the dimension of $\mathcal{N}(A^d)$ has to exceed twice the dimension of the nullspace of A.

Example 4.4.14. We now construct an example $A \in \mathbb{R}^{4 \times 4}$ as described above. For the nilpotent part of A we choose the 3×3 Jordan block

$$N = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

The regular component is simply $C = \begin{bmatrix} 1 \end{bmatrix}$. This implies $\operatorname{index}(A) = 3$ and $r = \operatorname{rank}(A^3) = 1$. Obviously, $\mathcal{N}(A^*)$ is spanned by the last unit vector, in particular; $t = 1$ and consequently $s = n - r - t = 2$. Thus we have $S_X^{\mathrm{H}} = \begin{bmatrix} 0 & 0 \end{bmatrix}$ and $S_Y = \begin{bmatrix} 1 \end{bmatrix}$. The latter means that we can choose any vector of unit norm as the basis Y. Let

$$Y = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad U = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \text{and} \qquad V = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The nullspace of N is spanned by the first unit vector, which yields in $G = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. A nonzero M with $MG = O$ is given by $M = \begin{bmatrix} 0 & 1 \end{bmatrix}$ and thus

$$X = YS_X^{\mathrm{H}} + V + UM = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Putting all together, we obtain the matrix

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

which has the desired properties.

4.5 Stability and Continuity of the Krylov Solution

Consider the operator $K^{\mathrm{MR}} : \mathcal{H} \to \mathcal{H}$, which assigns to an initial residual $\boldsymbol{r}$ its Lth MR Krylov correction with respect to a fixed linear operator A:

$$K^{\mathrm{MR}}\boldsymbol{r} = \boldsymbol{c}_L^{\mathrm{MR}} = A^{\mathcal{K}_L(A,\boldsymbol{r})}\boldsymbol{r}.$$

For simplicity we assume that $\mathcal{H}$ is finite dimensional. Thus K^{MR} is well defined on $\mathcal{H}$. Since L, the degree of the minimal polynomial of $\boldsymbol{r}$ with respect to A, is not continuous as function of $\boldsymbol{r}$, we cannot expect that K^{MR} is a continuous operator. On the other hand, for regular A it is known that $K^{\mathrm{MR}} = A^{-1}$ is linear. More generally, from Theorem 4.3.5 we conclude that $K^{\mathrm{MR}} = A^{\mathrm{D}}$ on $\mathcal{R}(A^d)$, i. e., K^{MR} is continuous on $\mathcal{R}(A^{\mathrm{index}(A)})$.

Example 4.5.1. Consider the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

and the initial residual $\boldsymbol{r} = \begin{bmatrix} 0 & 1 & \varepsilon \end{bmatrix}^\top$. For $\varepsilon = 0$ we have $L = \mathrm{index}(\boldsymbol{r}, A) = 2$ and

$$H_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad \mathcal{K}_2(A, \boldsymbol{r}) = \mathrm{span}\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \},$$

i. e., a singular breakdown in step 2. The previous correction is $\boldsymbol{c}_{L-1}^{\mathrm{MR}} = \boldsymbol{c}_1^{\mathrm{MR}} = \boldsymbol{0}$, which is also the subspace inverse solution, i. e., the MR approximation in the Krylov space $\mathcal{K}_2(A, \boldsymbol{r})$ with minimal norm, so that $\boldsymbol{c}_L^{\mathrm{MR}} = \boldsymbol{c}_2^{\mathrm{MR}} = \boldsymbol{0}$.

If $\varepsilon \neq 0$, in exact arithmetic we get $L = \mathrm{index}(\boldsymbol{r}, A) = 3$ and $\mathcal{K}_3(A, \boldsymbol{r}) = \mathcal{H} = \mathbb{R}^3$. Thus, the subspace inverse coincides with the Moore-Penrose inverse and yields

$$\boldsymbol{c}_L^{\mathrm{MR}} = \boldsymbol{c}_3^{\mathrm{MR}} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ \varepsilon \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

with residual norm $|\varepsilon|$. The previous Krylov approximation $\boldsymbol{c}_{L-1}^{\mathrm{MR}}$ with the same residual is $\boldsymbol{c}_2^{\mathrm{MR}} = \begin{bmatrix} -1/\varepsilon^2 & 0 & 1 \end{bmatrix}^\top$. Note further that $\eta_{3,2} = \|(I - P_{\mathcal{V}_2})A^2\boldsymbol{r}\| = O(\varepsilon^2)$.

This shows, that K^{MR} may be non-continuous in exact arithmetic and the computation may become unstable in presence of roundoff errors.

Due to the growing computational and storage requirement for larger correction spaces in practice a restart strategy is often used: After a fixed number m of steps in algorithm (3.14) the calculated basis of $\mathcal{V}_m$ is discarded and the procedure is restarted with the current iterate $\boldsymbol{x}_m$ as the initial guess (cf. [60]).

In terms of polynomials and Krylov subspaces this means, that we have formed $\boldsymbol{r}_m = \boldsymbol{b} - A\boldsymbol{x}_m = \boldsymbol{r}_0 - Aq_{m-1}(A)\boldsymbol{r}_0 = p_m(A)\boldsymbol{r}_0$ with respect to $\mathcal{K}_m(A, \boldsymbol{r}_0)$ and then continue the MR approximation process with $\mathcal{C}_j = \mathcal{K}_j(A, \boldsymbol{r}_m)$, $j = 1, 2, \ldots$. With regard to our results about the role of the index one may ask, when $\mathrm{index}(\boldsymbol{r}_m, A) = 0$ holds true. Unfortunately, this is only possible if $\mathrm{index}(\boldsymbol{r}_0, A) = 0$ since the residual polynomial satisfies the normalization condition $p_m(0) = 1$ (cf. [64]). On the other hand, if the restart index m is smaller then the termination index L (which is the typical case in practice), we will never be annoyed with a singular breakdown. This seems to be good news, but we can not expect that the residuals of the restarted method converge to zero unless $\boldsymbol{r}_0 \in \mathcal{R}(A^{\mathrm{index}(A)})$.

To add insult to injury, due to round-off errors we may end up with a nonzero index even if the initial residual satisfies $\boldsymbol{r}_0 \in \mathcal{R}(A^{\mathrm{index}(A)})$ or, equivalently, $\mathrm{index}(\boldsymbol{r}_0, A) = 0$.

This problems can be resolved if a projection on $\mathcal{R}(A^{\mathrm{index}(A)})$ is known. For $\mathrm{index}(A) = 1$ this was also observed by Brown and Walker (cf. [7, p. 44]). An application of this kind of stabilizing Krylov methods is given in Section 5 in context of Markov chains.

4.6 Krylov Subspace Methods for Drazin Inverse Solutions

We have already seen in Section 4.3 that Krylov subspaces are closely related to the Drazin inverse of A. Thus, an obvious idea is to use MR and OR methods with Krylov subspaces to approximate the Drazin inverse solution of (3.1). The basic idea is to force all computations into $\mathcal{R}(A^d)$. Corollary 4.2.12 shows that this can be done by multiplying all relevant quantities by A^q, $q = \mathrm{index}(\boldsymbol{r}_0, A)$. Of course, this is only possible if q itself or at least an upper bound for it, such as $d = \mathrm{index}(A)$, is explicitly known.

The simplest case arises for $d = 1$ (for example, if A is symmetric or skew-symmetric), then one additional multiplication with A suffices. For $\boldsymbol{b} \in \mathcal{R}(A)$ even this is not necessary (cf. Proposition 4.4.5). If the linear system is symmetric (but inconsistent) the analysis of the resulting method benefits additionally from the fact $\mathcal{R}(A) \perp \mathcal{N}(A)$ (cf. [26]).

Unfortunately, in general a simple bound for q is not available. Clearly, for operator equations of finite rank, namely for $n \times n$ matrices, an overall bound for q or d is always given by the dimension n. But the use of this knowledge is limited by the computational effort for large systems and since the method becomes ill-conditioned: The condition number of A^k increases of one power per each multiplication with A.

To describe the general approach in details, suppose we know an upper bound of q. Since $\mathcal{R}(A^d)$ is A-invariant, without lost of generality we can base our considerations on q itself. Instead of solving (3.1) directly we consider the system

$$A^q A\boldsymbol{x} = A^q\boldsymbol{b}. \tag{4.31}$$

Given an initial guess $\boldsymbol{x}_0$ the corresponding residual is now $A^q\boldsymbol{b} - A^qA\boldsymbol{x}_0 = A^q\boldsymbol{r}_0$. The correction spaces have to be chosen such that $\mathcal{C}_m \subseteq \mathcal{R}(A^d)$. This can be done by

$$\mathcal{C}_m := \operatorname{span}\{A^q\boldsymbol{r}_0, \ldots, A^{q+m-1}\boldsymbol{r}_0\} = \mathcal{K}_m(A, A^q\boldsymbol{r}_0).$$

The corresponding approximation and residual spaces are

$$\begin{aligned}\mathcal{W}_m &= A^{q+1}\mathcal{C}_m = \mathcal{K}_m(A, A^{2q+1}\boldsymbol{r}_0) \qquad \text{and}\\ \mathcal{V}_{m+1} &= \operatorname{span}\{A^q\boldsymbol{r}_0\} + \mathcal{W}_m = \operatorname{span}\{A^q\boldsymbol{r}_0, A^{2q+1}\boldsymbol{r}_0, \ldots, A^{2q+m}\boldsymbol{r}_0\}.\end{aligned}$$

Since the index of $A^q\boldsymbol{r}_0$ with respect to A is zero, a singular breakdown can never occur. In the same manner as (4.16) one can prove that

$$\begin{aligned}L - q &= \min\{m \,:\, \mathcal{K}_m(A, A^q\boldsymbol{r}_0) = \mathcal{K}_{m+1}(A, A^q\boldsymbol{r}_0)\}\\ &= \min\{m \,:\, A^{\mathrm{D}}\boldsymbol{r}_0 \in \mathcal{K}_m(A, A^q\boldsymbol{r}_0)\},\end{aligned}$$

where L is the degree of the minimal polynomial of $\boldsymbol{r}_0$ with respect to A.

The Gram-Schmidt-Arnoldi algorithm supplies the decomposition

$$A^{q+1}C_m = V_{m+1}\tilde{H}_m,$$

where $V_{m+1} = \begin{bmatrix}\boldsymbol{v}_1 \ldots \boldsymbol{v}_m\end{bmatrix}$ is an orthonormal basis of $\mathcal{V}_m$ with $\boldsymbol{v}_1 = A^q\boldsymbol{r}_0/\beta$ ($\beta := \|A^q\boldsymbol{r}_0\|$). Since we solve (4.31) the MR correction $\boldsymbol{c}_m^{\mathrm{MR}} = C_m\boldsymbol{y}_m^{\mathrm{MR}}$ have to minimize

$$\begin{aligned}\|A^q\boldsymbol{r}_m^{\mathrm{MR}}\| = \|A^q\boldsymbol{r}_0 - A^{q+1}\boldsymbol{c}_m^{\mathrm{MR}}\| &= \cdots\\ \cdots = \|V_{m+1}\beta\boldsymbol{u}_1^{(m+1)} - A^{q+1}C_m\boldsymbol{y}_m^{\mathrm{MR}}\| &= \|\beta\boldsymbol{u}_1^{(m+1)} - \tilde{H}_m\boldsymbol{y}_m^{\mathrm{MR}}\|,\end{aligned}$$

which can be solved by the usual QR factorization. $\boldsymbol{x}_m^{\mathrm{MR}} = \boldsymbol{x}_0 + \boldsymbol{c}_m^{\mathrm{MR}}$ and $\boldsymbol{c}_m^{\mathrm{MR}} \in \mathcal{R}(A^d)$ implies that

$$\begin{aligned}(I - AA^{\mathrm{D}})\boldsymbol{x}_m^{\mathrm{MR}} &= (I - AA^{\mathrm{D}})\boldsymbol{x}_0 \qquad \text{and}\\ (I - AA^{\mathrm{D}})\boldsymbol{r}_m^{\mathrm{MR}} &= (I - AA^{\mathrm{D}})\boldsymbol{r}_0 = (I - AA^{\mathrm{D}})\boldsymbol{b} - A(I - AA^{\mathrm{D}})\boldsymbol{x}_0.\end{aligned}$$

For $m = L - q$ we know that $\boldsymbol{c}_{L-q} = A^{\mathrm{D}}\boldsymbol{r}_0$ is a suitable correction for which

$$\boldsymbol{r}_{L-q} = \boldsymbol{r}_0 - A\boldsymbol{c}_{L-q} = (I - AA^{\mathrm{D}})\boldsymbol{r}_0$$

and hence $A^q\boldsymbol{r}_{L-q} = \boldsymbol{0}$. The corresponding iterate is

$$\boldsymbol{x}_{L-q} = \boldsymbol{x}_0 + A^{\mathrm{D}}\boldsymbol{r}_0 = A^{\mathrm{D}}\boldsymbol{b} + (I - AA^{\mathrm{D}})\boldsymbol{x}_0.$$

Methods of this type are proposed by Sidi, namely DGCR [61], DBi-CG [62] and D-GMRES [63]. Here, the "D" in front of the abbreviated method name stands for "Drazin".

5 Continuous-Time Markov Chains

The investigation of Markov chains with large finite state space often requires the solution of large sparse singular linear systems. Problems of this type arise, e. g., in the analysis of telecommunication systems (cf. [31]) and computers and computer networks (cf. [21]), or ranking of web pages in the search engine Google (cf. [42, 43]). Modeling techniques for this problems include queuing networks (cf. [8]), stochastic Petri nets (cf. [3, 2]), Markovian process algebras (cf. [10, 9]) and stochastic automata networks (cf. [11]). There are several software tools for a computer-based specification of Markov models, for example [37, 36, 65, 48].

All kinds of iterative methods for solving linear systems has been proposed to compute stationary probability distributions (see, e. g., [54, 67, 51, 12]). The multilevel-kind aggregation/disaggregation methods (cf. [69, 49, 13]) and preconditioned MR Krylov methods (cf. [58, 59, 29]) seem to be the most promising numerical algorithms. Preconditioners can often benefit from the special structure of the Markov model or the underlying state space (cf. [15, 16, 17]). Krylov subspace methods are also suitable for computing transient solutions of continuous-time Markov chains (cf. [55, 67]).

In this section we give a short introduction into Markov chains. We concentrate ourselves on the continuous-time homogeneous Markov chains with finite state space and explain how the essential properties may be derived in a purely analytical and matrix algebraic context.

5.1 Definitions and Motivations from Stochastics

We regard a continuous-time stochastic process $X(t)$ with finite or countable state space $\mathcal{S}$. To be precise, $\{X(t) \mid t \in \mathcal{T}\}$ is a family of random variables on a probability space $(\Omega, \mathcal{F}, \Pr\{\cdot\})$ with values in $\mathcal{S}$. The sample space Ω is arbitrary, $\mathcal{F}$ is a sigma algebra whose elements are subsets of Ω and sometimes called events. $\Pr\{\cdot\} : \mathcal{F} \to [0,1]$ is a probability measure. For each fixed $t \in \mathcal{T}$ the mapping $\omega \mapsto X(t,\omega) = X(t)$ is a measurable function, i. e., the pre-images $\{\omega \,|\, X(t,\omega) = i\} = \{X(t) = i\}$ are measurable sets for all $i \in \mathcal{S}$. Thus, we can measure probabilities such as $\Pr\{X(t) = i\}$ or $\Pr\{X(s) = x(s)\}$ (with a deterministic function $x : \mathcal{T} \to \mathcal{S}$).

We denote the *states* by positive integers, $\mathcal{S} = \{1, 2, \dots\}$. We will refer to the parameter set $\mathcal{T} = [0, \infty)$ as *time*, an element $s \in \mathcal{T}$ is a *moment* or *time instant* and for the moments $s, s+t \in \mathcal{T}$ we refer to t as the *period* between them.

In other words, $X(t)$ is a random function. For every fixed $\omega \in \Omega$ the (deterministic) function $t \mapsto X(t,\omega) = X(t)$ is a realization of the stochastic process (also called a *sample path* or *trajectory*). Under relatively weak regularity conditions one may

conclude, that the sample paths of a stochastic process are piecewise constant functions (step functions).[1] For simplicity we will *assume* here all stochastic processes to be *jump processes*, such that the realizations are right continuous step functions.

The jumps are often called *events* or *transitions* of the process. We say that an event happens at time $s \in \mathcal{T}$ if the considered realization makes a jump at s. The event is said to be the transition from state i to state j if $X(s) = j$ and $X(t) = i$ for $t < s$ in the neighborhood of s. The probability of being in state j at time s is $\Pr\{X(s) = j\}$. Similarly, $\Pr\{X(t) = i,\ X(s) = j\}$ is the joint probability that the process is in i at time instant t and in j at s. Moreover, we will use conditional probabilities, defined as usual by $\Pr\{X(s) = j \mid X(t) = i\} = \Pr\{X(t) = i,\ X(s) = j\} / \Pr\{X(t) = i\}$. For $t < s$ this may be interpreted as the probability that the process stays in j after time s given that it was in state i at the previous moment t.

As common in Markov chain literature, we use row vectors for distributions and denote them by (bold) small Greek letters. The *state space distribution* or *state distribution* of the process at time t is the vector

$$\boldsymbol{\pi}(t) := \big[\Pr\{X(t) = i\}\big]_{i \in \mathcal{S}} =: \big[\pi_i(t)\big]_{i \in \mathcal{S}}\,.$$

The ith entry $\pi_i(t)$ of $\boldsymbol{\pi}(t)$ is the probability, that the process is in state i at the time t. For each fixed $t \in \mathcal{T}$ the vector $\boldsymbol{\pi}(t)$ forms a probability distribution over the state space $\mathcal{S}$. Thus it satisfies the normalization conditions

$$0 \le \pi_i(t) \le 1 \qquad \text{and} \qquad \sum_{i \in \mathcal{S}} \pi_i(t) = 1,$$

or, denoting by $\boldsymbol{1} := \big[1 \ldots 1\big]^\top$ the column vector of suitable dimension with all entries equals 1,

$$\boldsymbol{\pi}(t) \ge \boldsymbol{0}^\top \qquad \text{and} \qquad \boldsymbol{\pi}(t)\,\boldsymbol{1} = 1. \tag{5.1}$$

Here and henceforth inequalities between vectors or matrices are interpreted as inequalities in each component.

Comparing $\boldsymbol{\pi}(t)$ with the sample paths of a process, we should note the different view-points regarding a stochastic process: At each point of the time scale we have a probability distribution on the state space which describes the stochastic behavior of the process at this moment. If we take for each moment $t \in \mathcal{T}$ a realization of these state distributions we get a deterministic function (or sequence) which is a realization of the entire process.

Example 5.1.1 (Homogeneous Poisson process). The *Poisson process with intensity* $\lambda > 0$ is a stochastic process $X(t)$ with state space $\mathcal{S} = \{0, 1, 2, \ldots\}$ having stochastically

[1] This is not true in general, even not for Markov processes. Many Markov chain books (cf. [19, 28, 27]) are in great parts devoted to the construction of such "pathological" examples. However, all those example seems to require a countable infinite state space.

independent increments distributed as

$$\Pr\{X(t) - X(s) = n\} = \frac{[\lambda(t-s)]^n}{n!}\, e^{-\lambda(t-s)}$$

for $n = 0, 1, 2, \ldots$ and $s < t$. Evidently we have $X(s) \leq X(t)$ for $s \leq t$ and a natural normalization is the assumption that the process starts at $X(0) = 0$. The process is called *homogeneous* or *stationary* since the distribution of the increments depends only on the difference $t - s$.

The sample paths of $X(t)$ are non-decreasing step functions and the heights of their jumps are equal to 1 (which implies that the Poisson process is a so called *simple point process*). For each state n let T_n denote a real valued random variable distributed as $\Pr\{X(t) = n\}$, i. e., $\Pr\{T_n \leq t\} = \Pr\{X(t) = n\}$. These *jump times* form an increasing sequence $(T_n)_{n\geq 0}$ of random variables, i. e.,

$$T_0 := 0 < T_1 < T_2 < \cdots < T_n < \ldots, \qquad \lim_{t\to\infty} T_n = +\infty.$$

The times $Y_n = T_{n+1} - T_n$ between subsequent jumps are independent, identically distributed random variables with the *exponential distribution*

$$\Pr\{Y_n \leq t\} = 1 - e^{-\lambda t}.$$

If we consider a time interval $[s, t]$ in which exactly one jump occurs at the moment $s + T$ then the random time T is uniformly distributed over this interval, i. e.,

$$\Pr\{T \leq h | X(t) = n, X(s) = n - 1\} = \frac{h}{t - s}.$$

Therefore, the Poisson process can be seen as the most random stochastic process. It is in some sense the best model for an unordered and rather unpredictable sequence of events in continuous time.

Example 5.1.2 (Discrete time homogeneous Markov chain). Let $\hat{P} = \left[\hat{p}_{i,j}\right]_{i,j=1}^{n}$ denote an $n \times n$ row stochastic matrix, i. e., $\hat{P} \geq 0$ and $\hat{P}\,\mathbf{1} = \mathbf{1}$ (cf. Definition 5.2.1).

We define a discrete time stochastic process by the following recursive construction: Under the condition, that the process is in state i in the ℓth step, the distribution in the next step $\ell + 1$ is given by the ith row of $\hat{P}$. In other words, $\hat{p}_{i,j}$ is the conditioned probability of reaching state j in the next step if the chain stays in state i, in formulas:

$$\Pr\{X_{\ell+1} = j \mid X_\ell = i\} = \hat{p}_{i,j}.$$

The entries $\hat{p}_{i,j}$ are called transition probabilities and $\hat{P}$ is the (one-step) transition probability matrix. The process is homogeneous since the transition probabilities do not depend on ℓ.

The probability, that the process is in state i after ℓ steps is given by the ith entry of the vector

$$\boldsymbol{\pi}_\ell := \boldsymbol{\pi}_0 \hat{P}^\ell, \qquad \ell \geq 0,$$

where $\boldsymbol{\pi}_0$ denotes a given initial distribution vector. The entries of $\hat{P}^\ell$ are the ℓ-step transition probabilities. The sequence of random variables $(X_\ell)_{\ell \geq 0}$, where X_ℓ has the probability distribution $\boldsymbol{\pi}_\ell$, uniquely defines a stochastic process with state space $\mathcal{S} = \{1, 2, \ldots, n\}$.

Both processes in the examples above are memoryless in the sense, that the state after a jump depends only on the preceding state right before the jump. This is characteristic for Markov processes.

Definition 5.1.3 (Markov Property)**.** Let $X(t)$ be a (continuous-time) stochastic process with finite or countable state space $\mathcal{S}$. The process is said to be *Markovian* if one of the following equivalent conditions holds:

- For all $s, t \geq 0$ and nonnegative integers $i, j, x(u)$ there holds

$$\begin{aligned} &\Pr\{X(s+t) = j \mid X(s) = i,\ X(u) = x(u),\ 0 \leq u \leq s\} \\ &\qquad = \quad \Pr\{X(s+t) = j \mid X(s) = i\}. \end{aligned}$$

- For any combination of states $i, j, i_1, \ldots, i_k \in \mathcal{S}$ and times $0 \leq s_1 < \cdots < s_k < s$, $t > 0$ there holds

$$\begin{aligned} &\Pr\{X(s+t) = j \mid X(s) = i,\ X(s_1) = i_1, \ldots, X(s_k) = i_k\} \\ &\qquad = \quad \Pr\{X(s+t) = j \mid X(s) = i\}. \end{aligned} \tag{5.2}$$

- For all $s, t, h \in \mathcal{T}$ with $0 \leq s < t < h$ and $i, j, k \in \mathcal{S}$ there holds

$$\begin{aligned} &\Pr\{X(s) = i, X(h) = k \mid X(t) = j\} \\ &\qquad = \quad \Pr\{X(s) = i \mid X(t) = j\} \Pr\{X(h) = k \mid X(t) = j\}. \end{aligned}$$

A stochastic process with finite or countable state space satisfying one of these properties is called a *Continuous-Time Markov Chain* (CTMC).

The Markov chain is said to be *homogeneous* (time-stationary) if the right-hand side $\Pr\{X(s+t) = j \mid X(s) = i\}$ of the equations above is independent of s.

The first two conditions express that the next possible event depends only on the directly preceding state and is independent from all other previous states. The third characterization may be paraphrased as: Given the present $X(t)$, the past $X(s)$ and the future $X(h)$ of the process are stochastically independent (cf. [6, Theorem 4.1]).

For a homogeneous CTMC with finite state space we collect the *transition probabilities*

$$p_{i,j}(t) := \Pr\{X(s+t) = j \mid X(s) = i\} = \Pr\{X(t) = j \mid X(0) = i\}, \qquad i, j \in \mathcal{S},$$

in a matrix

$$P(t) = [p_{i,j}(t)]_{i,j\in\mathcal{S}},$$

which is called the *transition function* of $X(t)$. For each $i \in \mathcal{S}$ the row $\left[p_{i,j}(t)\right]_{j\in\mathcal{S}}$ constitute a probability distribution over $\mathcal{S}$ (this follows from the Markov property and the conservation law of probability). Thus it satisfies the normalization condition

$$\sum_{j\in\mathcal{S}} p_{i,j}(t) = 1,$$

moreover, since the entries are probabilities, they lie between 0 and 1. In matrix form we have

$$P(t) \geq O, \qquad P(t)\mathbf{1} = \mathbf{1}. \tag{5.3}$$

This identifies $P(t)$ as a stochastic matrix (cf. Definition 5.2.1). It is sometimes useful to deal with transition functions, which do not satisfy this condition in full strength, for example, if we derive a process from another process by excluding some states. Even then the transition function satisfies $P(t) \geq O$ and $P(t)\mathbf{1} \leq \mathbf{1}$. The Markov chain (or rather the transition function) is said to be *honest* if there holds (5.3) and is said to be *dishonest* otherwise. Note that it is simply possible to make a dishonest Markov chain honest by adding an auxiliary state (cf. [1, Proposition 1.1]).

An often useful fact is, that a CTMC is associated with a family of discrete time Markov chains: For each fixed $\tau > 0$ consider the stochastic matrix $P(\tau)$, which generates a discrete time Markov chain as described in Example 5.1.2. This chain is called the *τ-skeleton* of $X(t)$.

The initial distribution $\boldsymbol{\pi}(0) = \left[\pi_i(0)\right]_{i\in\mathcal{S}}$, where $\pi_i(0) = \Pr\{X(0) = i\}$, together with the transition function completely describes the stochastic behavior of a CTMC: Using the law of total probability one easily deduces

$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)P(t). \tag{5.4}$$

Thus, the state distribution of the process for any time t is determined by the matrix function $P(t)$ and the initial guess $\boldsymbol{\pi}(0)$. Note that (5.3) guarantees that the normalization (5.1) is satisfied for all $t > 0$ whenever it holds for $t = 0$.

An important subclass of Markov chains are irreducible CTMCs: The Markov chain is called *irreducible*, if there is a positive probability to reach each state starting in any other state. Equivalently, this can be expressed in terms of the transition probabilities $P(t) = \left[p_{i,j}(t)\right]_{i,j\in\mathcal{S}}$: The Markov chain is irreducible if and only if for any states i, j, $i \neq j$, there exists an $t > 0$ such that $p_{i,j}(t) > 0$.

If a (nontrivial) row vector $\tilde{\boldsymbol{\pi}} \gneq \boldsymbol{0}^\top$ is such that

$$\tilde{\boldsymbol{\pi}} P(t) = \tilde{\boldsymbol{\pi}} \qquad \text{for all } t \in \mathcal{T}, \tag{5.5}$$

it is called *invariant* or *stationary*. A *stationary distribution* additionally satisfies the normalization condition (5.1). If the state space is finite, any given stationary vector $\tilde{\boldsymbol{\pi}}$ can be easily normalized by $\hat{\boldsymbol{\pi}} := \tilde{\boldsymbol{\pi}}/(\tilde{\boldsymbol{\pi}}\boldsymbol{1})$. The process $X(t)$ exhibits time-constant state distributions if the initial distribution is stationary, i. e., $\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0) = \hat{\boldsymbol{\pi}}$. In the analysis of a Markov chain the question of existence and uniqueness of a stationary distribution naturally arise.

At first glance, the description of a CTMC as given in (5.4) appears not as satisfactory as for discrete time Markov chains (cf. Example 5.1.2), where a constant matrix suffices to specify the process. Instead, we have to deal with a matrix of functions. A deeper study of the transition function is therefore required.

As a consequence of the Markov property, the transition function satisfies the *Chapman-Kolmogorov equation*

$$P(t+s) = P(t)P(s),$$

which shows that the functions forming $P(t)$ are tightly coupled with each other.

In applications of Markov theory an urgent question is the long term behavior of the process. The Markov chain is called *ergodic* if there exists a *limit distribution*

$$\boldsymbol{\pi} := \lim_{t\to\infty} \boldsymbol{\pi}(t) = \boldsymbol{\pi}(0) \lim_{t\to\infty} P(t), \tag{5.6}$$

which is independent from the initial distribution $\boldsymbol{\pi}(0)$ and, additionally, is strictly positive, i. e., $\boldsymbol{\pi} > \boldsymbol{0}^\top$. Since the limit in (5.6) is independent of $\boldsymbol{\pi}(0)$ we may substitute the columns of the identity matrix one after the other for $\boldsymbol{\pi}(0)$ on the right hand side and conclude that $\lim_{t\to\infty} P(t)$ has all rows equal to $\boldsymbol{\pi}$, i. e.,

$$\lim_{t\to\infty} P(t) = \boldsymbol{1}\boldsymbol{\pi} =: \Pi.$$

A simple calculation using the Chapman-Kolmogorov equation shows, that the limit distribution satisfies $\boldsymbol{\pi} P(t) = \boldsymbol{\pi}$ for all $t \in \mathcal{T}$. Thus, if the limit distribution exists, it is stationary (cf. (5.5)).

Two other properties of the transition function not yet mentioned are

$$\lim_{t\to 0} P(t) = P(0) = I.$$

The rightmost equality means, that a transition cannot take place in zero time. The continuity is justified by our assumption, that the realizations are right continuous functions. A Markov chain respectively the transition function is called *standard* if $P(t)$ is right continuous at zero.

The further study of Markov chains can be based on the properties of the transition function only without any reference to the stochastic origin of the problem. Before we continue with it, we recapitulate some background from matrix analysis.

5.2 Stochastic Matrices and Singular M-Matrices

We summarize here some well known concepts from matrix analysis and quote the main results in a form appropriate for our subsequent considerations.

Definition 5.2.1. A real matrix is called *substochastic* if the entries are nonnegative and each row sum is bounded by 1, in formulas:

$$A = \left[a_{i,j}\right]_{i,j=1}^{n} \in \mathbb{R}^{n\times n}, \qquad 0 \le a_{i,j} \le 1, \qquad \sum_{j=1}^{n} a_{i,j} \le 1 \qquad \text{for all } i \text{ and } j.$$

If equality holds in the last equation we call A (row) *stochastic*.

In matrix notation, the real matrix A is (sub)stochastic if and only if

$$A \ge O \qquad \text{and} \qquad A\mathbf{1} = \mathbf{1} \quad (A\mathbf{1} \le \mathbf{1}). \tag{5.7}$$

With regard to the notation we remind that inequalities between matrices are interpreted as inequalities in each component and $\mathbf{1}$ denotes the column vector of suitable dimension containing 1 in each component. Moreover, we denote by $\rho(A)$ the spectral radius of A.

Stochastic and substochastic matrices have many algebraic properties, which often stem from the fact that these matrices are in particular nonnegative and we can apply the Perron-Frobenius theory. We start, however, with a very basic result concerning the location of eigenvalues.

Proposition 5.2.2. *If $A \in \mathbb{R}^{n\times n}$ is substochastic there holds $\rho(A) \le 1$. Equality holds if A is a stochastic matrix and then $\rho(A) = 1$ is an eigenvalue of A and, moreover, there exists an nonnegative left eigenvector corresponding to* 1.

Proof. If $A\mathbf{1} = \mathbf{1}$, then obviously 1 is an eigenvalue of A. The main assertion now follows from (5.7) and Geršgorin's Theorem (cf. [39, Theorem 6.1.1]). The existence of a nonnegative left eigenvector of a stochastic matrix follows from [39, Theorem 8.3.1] applied to $A^\top$. □

The main results of the Perron-Frobenius theory can be stated for irreducible matrices.

Definition 5.2.3. A matrix $Q \in \mathbb{C}^{n\times n}$, $n > 1$, is called *reducible* if there is some permutation matrix T and some integer r, $0 < r < n$, such that

$$T^\top Q T = \begin{bmatrix} A & B \\ O & C \end{bmatrix}$$

with $A \in \mathbb{C}^{r\times r}$, $B \in \mathbb{C}^{r\times(n-r)}$, $C \in \mathbb{C}^{(n-r)\times(n-r)}$ and a zero matrix $O \in \mathbb{C}^{(n-r)\times r}$. The matrix Q is called *irreducible* if it is not reducible (cf. [39, p. 360]).

We shall need the following characterization of irreducibility.

Theorem 5.2.4. *A matrix $Q \in \mathbb{C}^{n\times n}$ is irreducible if and only if the following condition is satisfied:*

For every pair of distinct indices i, j there is a sequence of distinct integers $k_1 = i, k_2, k_3, \ldots, k_{m-1}, k_m = j$, such that all of the matrix entries $q_{k_1,k_2}, q_{k_2,k_3}, \ldots, q_{k_{m-1},k_m}$ are nonzero.

Proof. See [39, Definition 6.2.7 and Theorem 6.2.24]. □

Directly from the definition we have the following corollaries.

Corollary 5.2.5. *Let $A, B \in \mathbb{C}^{n\times n}$ with $A = \left[a_{i,j}\right]_{i,j=1}^{n}$ and $B = \left[b_{i,j}\right]_{i,j=1}^{n}$. If there holds for all i and j, $i \neq j$, that*

$$a_{i,j} = 0 \qquad \Longleftrightarrow \qquad b_{i,j} = 0,$$

then A is irreducible if and only if B is irreducible.

Corollary 5.2.6. *Let $A \in \mathbb{C}^{n\times n}$. Then A is irreducible if and only if $A^\top$ is irreducible.*

The main result about irreducible stochastic matrices guarantees the existence and uniqueness of a positive left eigenvector corresponding to the spectral radius.

Theorem 5.2.7. *Suppose that $P \in \mathbb{R}^{n\times n}$ is stochastic and irreducible. Then 1 is an algebraically (and hence geometrically) simple eigenvalue of P and there exists a positive left eigenvector $\boldsymbol{\pi}$ corresponding to the eigenvalue 1, i. e., $\boldsymbol{\pi}P = \boldsymbol{\pi}$, $\boldsymbol{\pi} > \boldsymbol{0}^\top$.*

Proof. Note that the spectral radius of a stochastic matrix equals 1 (see Proposition 5.2.2) and apply [39, Theorem 8.4.4, page 508] to $P^\top$. □

If P is not irreducible, still the following holds:

Theorem 5.2.8. *The stochastic matrix P possesses a positive left eigenvalue corresponding to $\rho(P) = 1$ if and only if*

$$TPT^\top = \begin{bmatrix} P_{1,1} & O & \ldots & O \\ O & P_{2,2} & \ldots & O \\ \vdots & & \ddots & \vdots \\ O & \ldots & O & P_{m,m} \end{bmatrix}, \tag{5.8}$$

where $T \in \mathbb{R}^{n\times n}$ is a permutation matrix and each diagonal block $P_{j,j}$ is irreducible with $\rho(P_{j,j}) = 1$ (i. e., the diagonal blocks are itself stochastic).

Proof. Noting the fact, that a row stochastic matrix has a positive right eigenvector corresponding to its spectral radius (namely $\boldsymbol{1}$), the assertion follows from [5, Theorem 2.3.14]. □

We continue with some some definitions and results concerning M-matrices.

Definition 5.2.9. A matrix $T \in \mathbb{R}^{n\times n}$ is said to be *semiconvergent* whenever

$$\lim_{k\to\infty} T^k \text{ exists.}$$

Proposition 5.2.10. *Assume that* $P = \left[p_{i,j}\right]_{i,j=1}^{n}$ *is a stochastic matrix and consider the conditions*

(a) P *is semiconvergent;*

(b) P *has only one eigenvalue of maximum modulus (that is,* 1 *is the only eigenvalue with modulus* 1*);*

(c) $p_{i,i} > 0$ *for all* $i = 1, \ldots, n$.

Then the following implication chain holds:

$$\text{(c)} \Longrightarrow \text{(b)} \Longrightarrow \text{(a)}.$$

Proof. See, e. g., [41, Hauptsatz 3.10, p. 101]. □

Definition 5.2.11. Any matrix $A \in \mathbb{R}^{n\times n}$ of the form

$$A = \sigma I - B, \qquad \sigma > 0, \qquad B \geq O, \tag{5.9}$$

for which $\sigma \geq \rho(B)$, is called an M-*matrix*. A is a *singular* M-*matrix* if $\sigma = \rho(B)$.

Definition 5.2.12. An M-matrix A is said to have *"property c"* if the splitting (5.9) can be chosen such that the matrix B/σ is semiconvergent.

Proposition 5.2.13. *An* M-*matrix* A *has "property c" if and only if* $\operatorname{index}(A) \leq 1$.

Proof. See [5, Lemma 4.11, p. 153]. □

5.3 Semigroups of Stochastic Matrices

Let $P(t) \in \mathbb{R}^{n\times n}$, $t \geq 0$ denote a family of matrices which satisfies

$$P(0) = I \qquad \text{and} \tag{5.10}$$

$$P(s+t) = P(s)P(t) \qquad \text{for all } s, t \geq 0. \tag{5.11}$$

Thus, $\{P(t) \mid t \geq 0\}$ is a *semigroup* with identity element I and it is a subsemigroup of the multiplicative semigroup of real matrices. Equation (5.11) is sometimes referred to as semigroup property. We will claim various additional properties on $P(t)$, such as non-negativity, (a special kind of) boundedness, or continuity:

$$P(t) \geq O \qquad \text{for all } t \geq 0, \tag{5.12}$$

$$P(t)\mathbf{1} \leq \mathbf{1} \qquad \text{for all } t \geq 0, \tag{5.13}$$

$$\lim_{h \to 0} P(t+h) = P(t) \qquad \text{for } t, h \geq 0.$$

It can be proved that a semigroup is continuous if and only if it is continuous at the origin:

$$\lim_{h \to 0} P(h) = I. \tag{5.14}$$

The conditions (5.12) and (5.13) clearly imply that $P(t)$ is a substochastic matrix for any $t \geq 0$ (and thus the entries of $P(t)$ are bounded by 1). The semigroup is said to be *stochastic* if $P(t)\mathbf{1} = \mathbf{1}$ for all $t \geq 0$ (i. e., if $P(t)$ is a stochastic matrix) and *substochastic* if at least (5.13) holds.[2] In memory of their role in the Markov chain realm, a semigroup satisfying (5.10)–(5.13) is called a *transition semigroup*. If it is (right) continuous, $P(t)$ is called a *standard* semigroup.

Semigroups of substochastic matrices have many interesting properties.

Proposition 5.3.1. *Let $P(t) = \big[p_{i,j}(t)\big]_{i,j=1}^{n}$ denote a transition semigroup. Then:*

(i) *the components of the vector $P(t)\mathbf{1}$ are nonincreasing functions of t;*

(ii) *if $P(t)$ is a stochastic matrix for some $t > 0$, it is stochastic for all $t \geq 0$.*

For any fixed $i = 1, \ldots, n$ there holds:

(iii) $0 \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} p_{i,j}(t) \leq 1 - p_{i,i}(t)$ *for all $t \geq 0$;*

(iv) $\lim_{h \to 0} p_{i,i}(h) = 1$ *implies* (5.14).

If $P(t)$ is standard we have additionally:

(v) $p_{i,i}(t) > 0$ *for all $t \geq 0$;*

(vi) *if $p_{i,i}(t) = 1$ for some $t > 0$, then $p_{i,i}(t) = 1$ for all $t \geq 0$ (moreover, by* (iii)*, $p_{i,j}(t) = 0$ for all $j \neq i$ and $t \geq 0$);*

[2] Some authors call the semigroup "honest" or "dishonest" likewise the Markov chain.

(vii) $|p_{i,j}(t+h) - p_{i,j}(t)| \leq 1 - p_{i,i}(|h|)$, *for* $t \geq 0$ *such that* $-h < t$, *that is the entry functions of* $P(t)$ *are uniformly continuous.*

Proof. The above results can be found, e. g., in Anderson's book (cf. [1, Propositions 1.1.2 and 1.1.3]). □

The main result is about differentiability of the semigroup, which follows from the uniform continuity.

Theorem 5.3.2. *The standard transition semigroup* $P(t)$ *possesses a finite derivative on* $[0, \infty)$*, which satisfies*

$$P'(t) = \frac{\mathrm{d}}{\mathrm{d}t} P(t) = QP(t) = P(t)Q, \tag{5.15}$$

where $Q := P'(0)$ *is the derivative at the origin. The unique solution of this system of differential equations with initial condition* $P(0) = I$ *is given by:*

$$P(t) = e^{tQ} = \sum_{k=0}^{\infty} \frac{t^k}{k!} Q^k. \tag{5.16}$$

Definition 5.3.3. The matrix $Q = P'(0)$ is called the *infinitesimal generator* of the semigroup $P(t)$. The differential equations $P'(t) = P(t)Q$ and $P'(t) = QP(t)$ in (5.15) are known as *Kolmogorov forward* and *Kolmogorov backward equation*, respectively.

Proof. Let $S(\tau) := \int_0^\tau P(t)\,\mathrm{d}t$, where the integral is defined for each scalar component. Our first objective is to show that $S(\tau)$ is invertible for sufficiently small $\tau > 0$. This is certainly true if $\|I - \tau^{-1}S(\tau)\| < 1$ in some matrix norm $\|\cdot\|$. One may accept the claim that this holds for any matrix norm since $P(t)$ is uniformly continuous (cf. (vii) in Proposition 5.3.1). A more detailed and elementary proof is given here: Let $c \in (\frac{1}{2}, 1)$ be fixed. Then we choose an $\tau > 0$ sufficiently small so that $p_{i,i}(t) > c$ for all i and $0 < t < \tau$. This is possible since $p_{i,i}(0) = 1$ and since $p_{i,i}(t)$ is uniformly continuous with respect to t (and even so, with respect to i, however, this is trivial, since there are only finitely many i's). Take as $\|\cdot\|$ the maximum row sum norm. Then we have

$$\begin{aligned} \|I - \tau^{-1}S(\tau)\| &= \max_{i=1}^{n} \sum_{j=1}^{n} \left| \delta_{i,j} - \frac{1}{\tau} \int_0^\tau p_{i,j}(t)\,\mathrm{d}t \right| \\ &= \max_{i=1}^{n} \left(\left| 1 - \frac{1}{\tau} \int_0^\tau p_{i,i}(t)\,\mathrm{d}t \right| + \frac{1}{\tau} \sum_{j \neq i} \int_0^\tau p_{i,j}(t)\,\mathrm{d}t \right), \end{aligned}$$

where we may omit the absolute value in the left term, since $\frac{1}{\tau}\int_0^\tau p_{i,i}(t)\,\mathrm{d}t \leq 1$. The term on the right can be estimated using (iii) of Proposition 5.3.1 as

$$\frac{1}{\tau} \int_0^\tau \sum_{j \neq i} p_{i,j}(t)\,\mathrm{d}t \;\leq\; \frac{1}{\tau} \int_0^\tau (1 - p_{i,i}(t))\,\mathrm{d}t \;=\; 1 - \frac{1}{\tau} \int_0^\tau p_{i,i}(t)\,\mathrm{d}t.$$

Thus we get

$$\|I - \tau^{-1}S(\tau)\| \;\leq\; \max_{i=1}^{n} 2\left(1 - \frac{1}{\tau}\int_0^\tau p_{i,i}(t)\,\mathrm{d}t\right) \;\leq\; \max_{i=1}^{n} 2\,(1-c) < 1$$

The two rightmost inequalities follow from the choice of τ and c: On the one hand, $p_{i,i}(t) > c$ for all $t \in (0,\tau)$ implies $\int_0^\tau p_{i,i}(t)\,\mathrm{d}t \geq \tau c$, on the other hand, $1 - c < \frac{1}{2}$ since $c \in (\frac{1}{2}, 1)$. Now, using (5.11), we write

$$\frac{1}{h}\,(P(h) - I)\int_0^\tau P(t)\,\mathrm{d}t = \frac{1}{h}\left(\int_0^\tau P(t+h)\,\mathrm{d}t - \int_0^\tau P(t)\,\mathrm{d}t\right).$$

We manipulate the integrals, the first by substituting $s = t+h$, the second by splitting at h, and then reorganize the sum to obtain integration intervals of length h. Due to the invertibility of $S(\tau)$ we finally get

$$\frac{1}{h}\,(P(h) - I) = \left(\frac{1}{h}\int_\tau^{\tau+h} P(t)\,\mathrm{d}t - \frac{1}{h}\int_0^h P(t)\,\mathrm{d}t\right)\left(\int_0^\tau P(t)\,\mathrm{d}t\right)^{-1}.$$

Letting $h \to 0$ in each component shows that $P'(0)$ exists and all entries are finite, namely

$$Q := \frac{\mathrm{d}}{\mathrm{d}t}P(t)\bigg|_{t=0} = (P(\tau) - I)\left(\int_0^\tau P(t)\,\mathrm{d}t\right)^{-1}. \tag{5.17}$$

Using (5.11) again, we write

$$\frac{P(t+h) - P(t)}{h} = P(t)\frac{P(h) - I}{h} = \frac{P(h) - I}{h}P(t).$$

Taking (in each component) limits $h \to 0$ yields the differential equations (5.15). Existence and uniqueness of the solution (5.16) are common result on systems of linear differential equations (cf. [40, p. 432]). □

Remark 5.3.4. Semigroups may be studied in a more abstract setting, where $P(t)$ is a family of bounded operators on a Banach space $\mathcal{S}$. This is done, e. g., by Pazy in [53], where the above proof of the existence and finiteness of the derivative $Q = P'(0)$ was found (cf. [53, Theorem 1.1.2]). We restated it here mainly because it yields the remarkable formula (5.17).

This formula provides an expression for Q in terms of a parameter τ, that has to be chosen sufficiently small. In terms of the Markov process, "sufficiently small" means that the probability of leaving the current state within any time interval smaller then τ is exceeded by the probability of staying in the current state, regardless which one the current state is. Roughly speaking, the process makes probably no jump within a period of length τ.

In Markov chain literature the results of Theorem 5.3.2 are commonly proved in a setting which covers the case of countable many states too and allows infinite values for the derivatives (cf. [1, 6, 19, 28, 70]).

The usual proof can be sketched as follows: Show that $q_i := \lim_{t\to 0}(1 - p_{i,i}(t))/t$ exists, but may be infinite. Then show that $q_{i,j} := \lim_{t\to 0} p_{i,j}(t))/t$ for $i \neq j$ exists and is always finite. Assume that all q_i are finite (or justify this by an auxiliary argument like "finite state space") and prove the remaining assertions.

The differentiability of $p_{i,i}(t)$ at 0 is proved by regarding the subadditive function $h(t) := -\log(p_{i,i}(t))$, which satisfies $\lim_{t\to 0} h(t) = 0$. A common lemma on such functions says that the limit $q_i = \lim_{t\to 0} h(t)/t$ exists and coincides with $\sup_{t>0} h(t)/t$, but it may be $+\infty$ (for details see [6, Theorem A.1.11 and Theorem 8.2.1] or [70, Lemma 8.3.1 and Proposition 8.3.2]). This proof does not require that the matrices $P(t)$ are neither stochastic nor substochastic (cf. [1, Lemma 1.2.1 and Proposition 1.2.2]).

The existence and finiteness of the derivative $p'_{i,j}(0)$ for all $i \neq j$ is usually proved using a "stochastic argument": One considers the discrete time Markov chain generated by the matrix $P(\tau)$ with an appropriate chosen $\tau > 0$ (the τ-skeleton). Details may be found in [70, Proposition 8.3.3] or [6, Theorem 8.2.1]. Chung [19, Theorem II.2.6] additionally remarks that the stochastic background is merely for sake of motivation. Freedman [28] shows that the proof even holds for a substochastic semigroup by previously constructing a discrete time Markov chain from a substochastic matrix.

Comparing the proofs one also observes, that some details are apparently similar. For example, the choice of an t such that $p_{i,i}(t) > c > \frac{1}{2}$ appears in the same manner in the beginning of Todorovic's proof ([70, p. 208], see also [6, p. 334]).

Next, we state some properties of the derivative $Q = P'(0)$.

Corollary 5.3.5. *Denote by $Q = \left[q_{i,j}\right]_{i,j=1}^{n}$ the infinitesimal generator of a standard transition semigroup $P(t)$. There holds*

(i) $Q\mathbf{1} \leq \mathbf{0}$, *where equality holds if $P(t)$ is stochastic;*

(ii) $q_i := -q_{i,i} \geq \sum_{\substack{j=1 \\ j\neq i}}^{n} q_{i,j}$ *with equality holding if $P(t)$ is stochastic;*

(iii) $q_{i,j} \geq 0$ *for all $i \neq j$ and hence $q_i \geq 0$;*

(iv) $p_{i,i}(t) \geq e^{-q_i t} \geq 1 - q_i t$ *for all $t \geq 0$;*

(v) $q_i = 0$ *if and only if $p_{i,i}(t) = 1$ for all $t \geq 0$.*

Proof. According to Proposition 5.3.1, (i), the components of $P(t)\mathbf{1}$ are nonincreasing functions of t. Thus, their derivatives are non-positive. Letting $t = 0$ we get by (5.15)

$$\mathbf{0} \geq \frac{\mathrm{d}}{\mathrm{d}t} P(t)\mathbf{1}\bigg|_{t=0} = P(0)Q\mathbf{1} = Q\mathbf{1} \geq QP(t)\mathbf{1} = \frac{\mathrm{d}}{\mathrm{d}t} P(t)\mathbf{1},$$

where we used (5.12) and again (5.15) for the rightmost relations. Therefrom we derive (i) and (ii).

A closer look at the limits constituting the derivative $P'(0) = Q = \left[q_{i,j}\right]_{i,j=1}^{n}$ yields

$$q_{i,j} = \lim_{t\to 0}(p_{i,j}(t))/t \;\geq\; 0 \qquad \text{for } i \neq j \;\text{ and} \tag{5.18}$$

$$q_{i,i} = \lim_{t\to 0}(p_{i,i}(t) - 1)/t \;\leq\; 0, \tag{5.19}$$

which proves (iii).

For a proof of the remaining parts we refer to [1, Proposition 1.2.2]. □

The special structure of Q is remarkable since it implies a representation in terms of a (sub)stochastic matrix.

Proposition 5.3.6. *The infinitesimal generator Q of a continuous (sub)stochastic semigroup $P(t)$ can be written as*

$$Q = \lambda(\hat{P} - I) \tag{5.20}$$

with a sufficiently large $\lambda \geq 0$ and a (sub)stochastic matrix $\hat{P}$. The parameter λ may be chosen such that the diagonal entries of $\hat{P}$ are positive.

Proof. The sign structure of Q suggest to rescale it as $\frac{1}{\lambda}Q$ such that the diagonal entries are smaller (or equal) then one and then to define

$$\hat{P} := I + \frac{1}{\lambda}Q \tag{5.21}$$

to obtain a nonnegative matrix $\hat{P} = \left[\hat{p}_{i,j}\right]_{i,j=1}^{n}$. Thus we have to chose $\lambda \geq \hat{\lambda} := \max_{i=1}^{n} q_i$ and by $\lambda > \hat{\lambda}$ we achieve $\hat{p}_{i,i} > 0$. It remains to show that $\hat{P}$ is (sub)stochastic. Using $Q\mathbf{1} \leq \mathbf{0}$ we compute

$$\hat{P}\mathbf{1} = \mathbf{1} + \frac{1}{\lambda}Q\mathbf{1} \leq \mathbf{1},$$

where equality holds if and only if the semigroup is stochastic (cf. (i) in Corollary 5.3.5). □

In view of Proposition 5.2.2 and Definition 5.2.11, a reformulation of Proposition 5.3.6 is the statement that Q is a negative M-matrix. If we restrict to stochastic semigroups, we can state a little more.

Proposition 5.3.7. *If $Q \in \mathbb{R}^{n\times n}$ is the infinitesimal generator of a stochastic semigroup then $-Q$ is a singular* M*-matrix with "property c". In particular there holds*

$$\operatorname{index}(Q) = 1.$$

Proof. With $B = \lambda\hat{P}$ we have $-Q = \lambda I - B$ and $\rho(B) \leq \lambda$ where equality holds if $\hat{P}$ (and likewise the semigroup) is stochastic. Hence Q is a negative singular M-matrix. In view of Definition 5.2.12 and Proposition 5.2.13, it remains to show that $B/\lambda = \hat{P}$ is semiconvergent. Since we can choose λ in (5.20) such that $\hat{p}_{i,i} > 0$ this follows from Proposition 5.2.10. □

It should be noted, that the properties of Q stated in Propositions 5.3.6 and 5.3.7 are only consequences of the sign structure and row sum feature established in Corollary 5.3.5. We show next that these relatively minor conditions conversely imply that Q is an infinitesimal generator.

Proposition 5.3.8. *Any matrix* $Q = \left[q_{i,j}\right]_{i,j=1}^{n}$ *which satisfies*

$$q_{i,j} \geq 0 \quad \text{for } i \neq j, \text{ and} \quad -q_{i,i} = \sum_{\substack{j=1 \\ j\neq i}}^{n} q_{i,j}$$

is the derivative at 0 *of some standard stochastic semigroup* $P(t)$.

Proof. We show that $P(t) = e^{tQ}$ has the desired properties.

Obviously, $P(0) = I$. The semigroup property follows from the well-known fact that the "exponential addition theorem" holds true for the matrix exponential if the summands commute with each other (cf. [40, Theorem 6.2.38]).

That $P(t)$ is stochastic can be seen from the Taylor series representation. We have

$$P(t)\mathbf{1} = \sum_{k=0}^{\infty} \frac{t^k}{k!} Q^k \mathbf{1} = I\mathbf{1} + \left(\sum_{k=1}^{\infty} \frac{t^k}{k!} Q^{k-1} \right) Q\mathbf{1} = \mathbf{1}$$

since $Q\mathbf{1} = \mathbf{0}$. For sufficiently small t, we have $P(t) = I + tQ + \cdots \geq O$. Then by the semigroup property we have $P(\ell t) = P(t)^{\ell} \geq O$ and thus $P(t) \geq O$ for all $t \geq 0$.

It is well known, that the matrix exponential $t \mapsto e^{tQ}$ is a smooth map (see [40, Theorem 6.2.34]). Thus, $P(t)$ is continuous and satisfies $P'(0) = Q$. □

Noting that each M-matrix has nonpositive off-diagonal entries, we may summarize our results in the following form:

Theorem 5.3.9. *$P(t)$ is a continuous stochastic semigroup if and only if $P(t) = e^{-tA}$ with an* M*-matrix A satisfying $A\mathbf{1} = \mathbf{0}$.*

In particular, the matrix exponential transforms every negative M-matrix with zero row sum into a stochastic matrix, a notable fact on its own.

Corollary 5.3.10. *If $Q \in \mathbb{R}^n$ has nonnegative off-diagonal entries (e. g., if $-Q$ is an* M*-matrix) then e^Q is nonnegative. If there holds additionally $Q\mathbf{1} = \mathbf{0}$, then e^Q is a stochastic matrix.*

The next result justifies the concordance in naming of irreducibility.

Proposition 5.3.11. *The continuous transition semigroup $P(t) = e^{tQ}$ is irreducible, in formulas*

$$\forall\, i, j \in \mathcal{S}, i \neq j : \exists\, t > 0 : p_{i,j}(t) > 0, \tag{5.22}$$

if and only if the infinitesimal generator Q is irreducible.

Proof. Assume (5.22) but $Q \in \mathbb{R}^{n\times n}$ is reducible. Given a permutation matrix T, we have the identity $T^\top P(t) T = e^{tT^\top QT}$. Thus, we may assume without loss of generality that

$$Q = \begin{bmatrix} A & B \\ O & C \end{bmatrix}$$

with $A \in \mathbb{R}^{r\times r}$, $B \in \mathbb{R}^{r\times(n-r)}$ $C \in \mathbb{R}^{(n-r)\times(n-r)}$ and a zero matrix $O \in \mathbb{R}^{(n-r)\times r}$ for an integer r, $0 < r < n$ (cf. Definition 5.2.3). We rewrite $P(t) = e^{tQ}$ as a power series

$$P(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} Q^k = I + tQ + \frac{t^2}{2} Q^2 + \cdots .$$

A short calculation shows, that with Q also $Q^2, Q^3, \ldots$ has a $(n-r)\times r$ zero block in the lower left corner. Thus, also $P(t)$ possesses this zero block for all $t \geq 0$, i. e., $p_{i,j}(t) = 0$ for all $t \geq 0$, $r < i \leq n$ and $1 \leq j \leq r$, which contradicts the irreducibility of $P(t)$.

Assume now, that Q is irreducible but $p_{i,j}(t) = 0$ for all $t > 0$ and some states i and j. Using Theorem 5.2.4 there exists a sequence $k_1 = i, k_2, k_3, \ldots, k_{m-1}, k_m = j$, such that $q_{k_i,k_{i+1}} \neq 0$, $i = 1, \ldots, m-1$. Since $p_{i,j}(t) = 0$ for all $t > 0$ we get from the Kolmogorov backward equation

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} p_{i,j}(t) = \sum_{\ell=1}^{n} q_{i,\ell} p_{\ell,j}(t).$$

All terms in the sum on the right hand side are nonnegative and thus has to vanish if the sum is zero. For $\ell = k_2$ we know that $q_{i,\ell} = q_{k_1,k_2}$ is nonzero, which implies $p_{k_2,j}(t) = 0$ for all $t > 0$. Writing

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} p_{k_2,j}(t) = \sum_{\ell=1}^{n} q_{k_2,\ell} p_{\ell,j}(t).$$

we conclude by a similar argument that $p_{k_3,j}(t) = 0$ for all $t > 0$. By an obvious induction we get

$$p_{k_1,j}(t) = 0,\ p_{k_2,j}(t) = 0,\ \ldots\ , p_{k_{m-1},j}(t) = 0,\ p_{k_m,j}(t) = 0$$

and, noting $k_m = j$, this yields in $p_{j,j}(t) = 0$ for all $t > 0$. This contradicts to the continuity assumption on $P(t)$ since $\lim_{t\to 0} p_{j,j}(t) = 0$ but $p_{j,j}(0) = 1$. □

Note that, the proof above does not use the fact that $P(t)$ is stochastic or substochastic. We only require $P(t) \geq 0$, which is guaranteed if Q has nonnegative off-diagonal entries.

The first part of the proof shows: If Q is reducible, then $P(t)$ is a reducible matrix for every $t \geq 0$. Conversely, if there exists an $\tau > 0$ for which $P(\tau)$ is an irreducible matrix, then Q is irreducible. On the other hand, if $P(t)$ is a reducible matrix for every $t \geq 0$, then the semigroup $P(t)$ is not irreducible, or, conversely, if the semigroup $P(t)$ is irreducible, then there exists an $\tau > 0$ such that $P(\tau)$ is an irreducible matrix. Taking all together, we have

Corollary 5.3.12. *The following statements are equivalent:*

(i) *$P(t)$ is an irreducible semigroup,*

(ii) *Q is irreducible,*

(iii) *there exists $\tau > 0$ such that $P(\tau)$ is an irreducible matrix.*

Actually, there holds a lot more then stated in (iii), namely that $P(\tau)$ is an irreducible matrix for every $\tau > 0$ if the semigroup (or its generator) is irreducible. Moreover, it turns out that the matrices $P(\tau)$ are positive. To comprehend this further refinement we need Lévy's Dichotomy.

Lemma 5.3.13 (Lévy Dichotomy)**.** *For a standard transition semigroup, there holds: Either $p_{i,j}(t) > 0$ for all $t > 0$, or $p_{i,j}(t) = 0$ for all $t > 0$.*

Proof. Note that the entries $p_{i,j}(t)$ are analytic functions of t. In view of (v) in Proposition 5.3.1 we have $p_{i,i}(t) > 0$ for all $t \geq 0$ and all $i = 1, \ldots, n$. Thus, it remains to prove the assertion for the off-diagonal entries. Let $i \neq j$ be fixed. Assume $p_{i,j}(t) = 0$ for some $t > 0$. Using the semigroup property (5.11) we calculate

$$0 = p_{i,j}(t) = \sum_{k=1}^{n} p_{i,k}\left(\tfrac{t}{m}\right) p_{k,j}\left(\tfrac{m-1}{m}t\right) \geq p_{i,j}\left(\tfrac{t}{m}\right) p_{j,j}\left(\tfrac{m-1}{m}t\right).$$

Since the diagonal entries are positive, we conclude $p_{i,j}\left(\frac{t}{m}\right) = 0$ for any integer $m > 0$. That is, set set of zeros of $p_{i,j}(s)$ has a limit point (namely 0). Then, as an analytic function, $p_{i,j}(s)$ has to be identical zero, i. e., $p_{i,j}(s) = 0$ for all $s > 0$. □

Remark 5.3.14. The idea for the proof above was found in [45, p. 263]. Chung gives a proof of Lévy's Dichotomy for measurable transition functions (cf. [19, Theorems II.1.5]).

Corollary 5.3.15. *If the standard transition semigroup $P(t)$ is irreducible, i. e., it satisfies* (5.22)*, then there holds $P(t) > O$ for all $t > 0$.*

We summarize our results in the next theorem.

Theorem 5.3.16. *Let $P(t) = e^{tQ}$ denote a standard transition semigroup. The matrices $P(t)$ are positive for all $t > 0$ if and only if Q is irreducible.*

The following lemma will be used later to establish the uniqueness of invariant (stationary) vectors. It states, that the nullspace of Q is one-dimensional if the semigroup is irreducible. Since we know already, that $\operatorname{index}(Q) = 1$, this is equivalent to say that the rank of the generator is $n - 1$.

Lemma 5.3.17. *If $P(t) = e^{tQ}$ is an irreducible stochastic semigroup, then there holds $\dim \mathcal{N}(Q) = 1$ and there exists a positive left eigenvector of the eigenvalue 0.*

Proof. From Proposition 5.3.6 we have the representation of Q in terms of a stochastic matrix $\hat{P} = I + \frac{1}{\lambda}Q$ with some $\lambda > 0$. According to Corollary 5.2.5 there holds that $\hat{P}$ is irreducible if and only if Q is irreducible. Moreover, ϑ is an eigenvalue of Q if and only if $1 + \vartheta/\lambda$ is an eigenvalue of $\hat{P}$ with same multiplicity and corresponding to the same eigenvector.

Theorem 5.2.7 implies that 1 is a simple eigenvalue of $\hat{P}$ with a corresponding positive left eigenvector $\tilde{\boldsymbol{\pi}} > \boldsymbol{0}^\top$. Therefore 0 is a simple eigenvalue of Q corresponding to the eigenvector $\tilde{\boldsymbol{\pi}}$. In other word, $\mathcal{N}(Q)$ is one-dimensional and spanned by $\tilde{\boldsymbol{\pi}}$. □

5.4 Invariant Vectors and Asymptotic Behavior

We now investigate existence and uniqueness of invariant vectors of a semigroup. Given an initial guess $\boldsymbol{\pi}(0)$ a semigroup defines by

$$\boldsymbol{\pi}(t) := \boldsymbol{\pi}(0)P(t) \tag{5.23}$$

a vector valued function. We denote by $\mathbb{R}^n_+$ the set of vectors with nonnegative components. For $\boldsymbol{\pi}(0) \in \mathbb{R}^n_+$ and a substochastic semigroup $P(t)$ we have $\boldsymbol{\pi}(\cdot) : [0, \infty) \to \mathbb{R}^n_+$. Moreover, if the semigroup is standard, $\boldsymbol{\pi}(t)$ is a continuous function of t. We first state the representation of $\boldsymbol{\pi}(t)$ in terms of the infinitesimal generator, which follows immediately from the Kolmogorov backward equation.

Proposition 5.4.1. *The function $\boldsymbol{\pi}(t)$ defined in (5.24) is the uniquely determined solution of the initial-value problem*

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\pi}(t) = \boldsymbol{\pi}(t)\,Q, \quad t \geq 0, \quad \boldsymbol{\pi}(0) \text{ given.} \tag{5.24}$$

The solution is given by

$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)\,e^{tQ}, \quad t \geq 0. \tag{5.25}$$

This solution is nonnegative, if the semigroup is substochastic and $\boldsymbol{\pi}(0) \in \mathbb{R}^n_+$.

Equation (5.24) is called the *balance equation.*[3] A nontrivial vector $\tilde{\boldsymbol{\pi}}$ is *invariant* with respect to the semigroup $P(t)$, if $\tilde{\boldsymbol{\pi}} = \tilde{\boldsymbol{\pi}} P(t)$ for all $t \geq 0$.

Proposition 5.4.2. *$\tilde{\boldsymbol{\pi}} \in \mathbb{R}^n$ is invariant with respect to $P(t)$ if and only if it satisfies the homogeneous balance equation*

$$\tilde{\boldsymbol{\pi}} Q = \boldsymbol{0}^\top. \tag{5.26}$$

If there holds $\boldsymbol{\pi}(t) = \tilde{\boldsymbol{\pi}}$ for some $t \geq 0$, then it holds for all t, i. e., $\boldsymbol{\pi}(t)$ is constant.

Proof. Suppose, $\tilde{\boldsymbol{\pi}}$ is invariant and set $\boldsymbol{\pi}(0) = \tilde{\boldsymbol{\pi}}$. Then, (5.23) reads as $\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)P(t) = \tilde{\boldsymbol{\pi}}$ for all $t \geq 0$, which implies that the derivative of $\boldsymbol{\pi}(t)$ is zero. Using (5.24) we get

$$\boldsymbol{0}^\top = \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\pi}(t) = \boldsymbol{\pi}(t)Q = \tilde{\boldsymbol{\pi}} Q.$$

Assume now, there holds $\tilde{\boldsymbol{\pi}} Q = \boldsymbol{0}^\top$ for some vector $\tilde{\boldsymbol{\pi}}$. Then, using the Taylor series representation of $P(t)$, we conclude

$$\tilde{\boldsymbol{\pi}} P(t) = \tilde{\boldsymbol{\pi}} \sum_{k=0}^{\infty} \frac{t^k}{k!} Q^k = \tilde{\boldsymbol{\pi}} I + \tilde{\boldsymbol{\pi}} Q \left(\sum_{k=1}^{\infty} \frac{t^k}{k!} Q^{k-1} \right) = \tilde{\boldsymbol{\pi}}.$$

To prove the final remark in the assertion, let $\boldsymbol{\pi}(s) = \tilde{\boldsymbol{\pi}}$ be invariant for some $s > 0$. Since e^{tQ} is invertible (cf. [40, Theorem 6.2.38]), the definition of an invariant vector can be rewritten as $\tilde{\boldsymbol{\pi}} = \tilde{\boldsymbol{\pi}} e^{-tQ}$. Then for all $t \geq 0$ we have

$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)\, e^{tQ} = \tilde{\boldsymbol{\pi}}\, e^{(t-s)Q} = \tilde{\boldsymbol{\pi}}.$$

□

Proposition 5.4.3. *The continuous stochastic semigroup $P(t) = e^{tQ}$ possesses a nonnegative invariant vector $\hat{\boldsymbol{\pi}} \in \mathbb{R}^n_+$.*

Proof. We use the representation $Q = \lambda(\hat{P} - I)$ with $\lambda > 0$ and a stochastic matrix $\hat{P}$ (cf. Proposition 5.3.6). By Proposition 5.2.2, there exists a nonnegative left eigenvector $\hat{\boldsymbol{\pi}}$ of $\hat{P}$ corresponding to the eigenvalue 1. Clearly, $\hat{\boldsymbol{\pi}}$ satisfies (5.26) and thus is an invariant vector of $P(t)$. □

The next result is an application the Perron-Frobenius Theorem for irreducible matrices.

Proposition 5.4.4. *Let Q denote the infinitesimal generator of an irreducible continuous stochastic semigroup $P(t)$. Then there exists a uniquely determined invariant vector $\hat{\boldsymbol{\pi}}$ subject to $\hat{\boldsymbol{\pi}}\boldsymbol{1} = 1$. Moreover this invariant vector is positive, that is $\boldsymbol{\pi} > \boldsymbol{0}^\top$, and satisfies the homogeneous balance equation $\hat{\boldsymbol{\pi}} Q = \boldsymbol{0}^\top$.*

[3] In the realm of Brownian motion (5.24) is known as *master equation.*

Proof. Proposition 5.4.2 states that $\hat{\boldsymbol{\pi}}$ is invariant if and only if it lies in the nullspace of Q. By Lemma 5.3.17 this nullspace is one-dimensional and spanned by a positive eigenvector vector $\tilde{\boldsymbol{\pi}}$. Thus, setting $\hat{\boldsymbol{\pi}} := \tilde{\boldsymbol{\pi}}/(\tilde{\boldsymbol{\pi}}\mathbf{1})$ completes the prove. □

We show next, that also the converse holds true, that is, the existence and uniqueness of a positive invariant vector implies the irreducibility of the semigroup (or Q, respectively).

Proposition 5.4.5. *If the stochastic semigroup $P(t) = e^{tQ}$ possesses an invariant vector $\hat{\boldsymbol{\pi}}$, which is positive, $\hat{\boldsymbol{\pi}} > \mathbf{0}^\top$, and uniquely determined subject to $\hat{\boldsymbol{\pi}}\mathbf{1} = 1$, then the semigroup (or rather Q) is irreducible.*

Proof. The existence of an invariant $\hat{\boldsymbol{\pi}} > \mathbf{0}^\top$ implies that the stochastic matrix $\hat{P}$ in (5.20) has a positive left eigenvector corresponding to the eigenvalue $\rho(\hat{P}) = 1$. Then, by Theorem 5.2.8, $\hat{P}$ has the block diagonal form (5.8) with irreducible diagonal blocks.

If $\hat{\boldsymbol{\pi}}$ is uniquely determined up to a scalar factor, there exists only one diagonal block and thus $\hat{P}$ is irreducible. Then, by Corollary 5.2.5, also Q is irreducible and Proposition 5.3.11 provides that the semigroup $P(t)$ is irreducible. □

To finish our investigation of stochastic semigroups we show, that the invariant vector of an irreducible semigroup coincides with the limit $\lim_{t\to\infty} \boldsymbol{\pi}(t)$.

Theorem 5.4.6. *If the transition semigroup $P(t) = e^{tQ}$ is irreducible, then*

$$\lim_{t\to\infty} P(t) = \mathbf{1}\boldsymbol{\pi},$$

where $\boldsymbol{\pi}$ is the (uniquely determined) solution of

$$\boldsymbol{\pi} Q = \mathbf{0}^\top, \qquad \boldsymbol{\pi}\mathbf{1} = 1.$$

Proof. Since the semigroup is irreducible, for any fixed $\tau > 0$ we know that $P(\tau)$ is a positive matrix. The Perron-Frobenius Theorem (cf. [39, Theorem 8.2.11]) guarantees the existence of

$$\lim_{m\to\infty} P(m\tau) = \lim_{m\to\infty} P(\tau)^m = \mathbf{1}\boldsymbol{\pi}_\tau,$$

where $\boldsymbol{\pi}_\tau$ is the uniquely determined left eigenvector of $P(\tau)$ subject to $\boldsymbol{\pi}_\tau\mathbf{1} = 1$ (moreover there holds $\boldsymbol{\pi}_\tau > \mathbf{0}^\top$). This is true for any $\tau > 0$. The uniformly continuity of $P(t)$ then implies, that the limit is independent of τ and coincides with limit of $P(t)$ as $t \to \infty$. (see, e. g., [1, Lemma 5.1.2]). Writing $\boldsymbol{\pi} = \boldsymbol{\pi}_\tau$ we have

$$\boldsymbol{\pi} P(s+t) = \boldsymbol{\pi} P(s) P(t).$$

Letting $s \to \infty$ and using $\boldsymbol{\pi}\mathbf{1} = 1$ this shows that $\boldsymbol{\pi}$ is invariant with respect to $P(t)$. To finish the proof we apply Proposition 5.4.2. □

Remark 5.4.7. It can be proved by similar arguments, that the limit of $P(t)$ as $t \to \infty$ always exists, even if the semigroup is not irreducible (cf. [1, Theorem 5.1.3]). The limit matrix then no longer has all rows identical, its structure is determined by the *communicating classes* of $P(\tau)$ (cf. [5, Definition 2.3.8]).

Corollary 5.4.8. *Let $P(t) = e^{tQ}$ denote an irreducible transition semigroup and $\boldsymbol{\pi}(0)$ a given nonzero vector satisfying the normalization condition $\boldsymbol{\pi}(0)\mathbf{1} = 1$. Then the vector valued function $\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)P(t)$ converges for $t \to \infty$ to a limit independently of $\boldsymbol{\pi}(0)$, namely*

$$\lim_{t\to\infty} \boldsymbol{\pi}(t) = \boldsymbol{\pi},$$

where $\boldsymbol{\pi}$ is uniquely determined by $\boldsymbol{\pi}Q = \mathbf{0}^\top$, $\boldsymbol{\pi}\mathbf{1} = 1$.

5.5 Stochastic Interpretation of the Results

We now apply the semigroup theory developed in the previous sections to continuous-time Markov chains and give stochastic interpretations of the occurring quantities. The statements in this section are purely informal, but proofs of them can be found in every book on Markov chains mentioned in the references.

Remark 5.5.1 (Transition Rates and Classification of States). The results in Section 5.3 show that the transition semigroup of a CTMC on a finite state space is *characterized* by the infinitesimal generator Q. Since the semigroup in turn characterizes the Markov chain (cf. [1, pp. 3–4]), we have a one-to-one correspondence between the infinitesimal generator and the Markov chain. The numbers forming the matrix Q have a meaningful interpretation in terms of the stochastic process:

The off-diagonal entries $q_{i,j} \geq 0$, $i \neq j$, are the *rates* with which the Markov chain enters j as the next state when leaving the actual state i. The time Y_i, for which the process has stayed before in state i, is exponentially distributed with the parameter $q_i = -q_{i,i}$. The negative diagonal entries q_i of Q are called the *jump rates* associated with state i. The reciprocal value $1/q_i$ is the mean time, that the process stays in state i.

The state i is said to be *stable* if $q_i > 0$, *permanent* or *absorbing* if $q_i = 0$ and *instantaneous* if $q_i = \infty$ (impossible in our setting). Thus, for a permanent state we have $\Pr\{Y_i \leq t\} = 1 - e^{-q_i t} = 0$ for any $t \in \mathcal{T}$, i. e., the probability of a finite sojourn time is zero. This corresponds to the following observation:

If i is a permanent state, then the entire ith row of Q is zero (cf. (ii) and (iii) of Corollary 5.3.5). By (v) of Corollary 5.3.5 the corresponding diagonal entry of $P(t)$ is 1 and, moreover, all off-diagonal entries in the ith row are zero (cf. (vi) in Proposition 5.3.1). That is, the ith row of $P(t)$ coincides with the ith unit vector and is independent form t. This means, that the Markov chain will never leave state i after entering it once.

State i is said to be *conservative* if equality holds in (ii) of Corollary 5.3.5. The new name is introduced, since for infinitely many states this is not true in general, even not if $P(t)$ is stochastic. If all states are conservative, i. e., $Q\mathbf{1} = \mathbf{0}$, then the infinitesimal generator Q is called conservative. The infinitesimal generator of a CTMC with finite state space is conservative.

Remark 5.5.2 (Uniform Markov Chain). Every homogeneous CTMC on a finite state space is *uniformizable*. Consider the stochastic matrix $\hat{P}$ defined in (5.20). As explained in Example 5.1.2, the matrix $\hat{P}$ specifies a discrete time Markov chain $\hat{X}_\ell$. Using the parameter $\lambda > 0$ from (5.20) we construct a homogeneous Poisson process $N(t)$ with intensity λ (cf. Example 5.1.1). Now, we define by

$$X(t) := \hat{X}_{N(t)},$$

an homogeneous CTMC, which is called the *uniform Markov chain* with *clock* $N(t)$ and *subordinated chain* $\{\hat{X}_\ell\}_{\ell\geq 0}$. Note that by definition we have $X(t) = \hat{X}_\ell$ for $T_\ell \leq t < T_{\ell+1}$, where T_ℓ, $\ell \geq 0$, are the jump times of the Poisson process. Observe also, that the continuous-time Markov chain may stay in the current state while the Poisson process makes a jump. This happens precisely if $\hat{X}_\ell = \hat{X}_{\ell+1}$, which is possible since the diagonal elements of $\hat{P}$ can be nonzero.

The process constructed this way coincides with our original Markov chain, since it has the same infinitesimal generator, which is given by (5.20). The transition function has the representation

$$P(t) = \sum_{k=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} \hat{P}^k,$$

as can be seen from $e^{t\lambda(\hat{P}-I)} = e^{-t\lambda} e^{t\lambda\hat{P}}$.

Note further, that due to $\lambda \geq \max_{i=1}^{n} q_i$ we could chose $\lambda = 0$ in (5.20) if and only if $q_i = 0$ for all i, which means that all states are permanent. In view of Corollary 5.3.5, (v), this is equivalent to $P(t) = I$ for all $t \geq 0$. Thus, $\lambda = 0$ corresponds to an entirely constant process, where never happens any transition.

Remark 5.5.3 (The Embedded Chain). We have seen so far two kinds of discrete-time Markov chains connected with a continuous-time chain, namely the τ-skeleton with the transition matrix $P(\tau)$ and the subordinated chain $\hat{X}_\ell$ with the transition matrix $\hat{P}$ corresponding to some equivalent uniform Markov chain.

A third discrete-time Markov chain linked with $P(t)$, the *embedded chain* $(X_k)_{k\geq 0}$. X_k is defined as the kth state, which is visited by the continuous time Markov chain. That is, we just ignore the time spent in each state and regard only the sequence of states. The transition probabilities of the embedded chain are given by

$$p_{i,j} := \Pr\{X_{n+1} = j \mid X_n = i\} = \begin{cases} q_{i,j}/q_i, & i \neq j \\ 0, & i = j \end{cases}.$$

Note that these probabilities are well defined as long as $q_i \neq 0$, i. e., if the Markov chain has no permanent states. The matrix $P = \left[p_{i,j}\right]_{i,j\in\mathbb{S}}$ can be written as

$$P = I - D^{-1}Q, \qquad D = \operatorname{diag}(Q).$$

Comparing the reformulation

$$Q = D(P - I)$$

with (5.20) explains the matrix algebra behind these the things: To define $\hat{P}$ we applied a scaling factor λ to the whole matrix, while P results from an row-wise scaling.

With regard to the discrete time Markov chains defined by $\hat{P}$ and P we remark, that the difference between them is the following: The embedded chain corresponding to P changes its state at each step since the diagonal entries of P are zero. On the other hand, the uniform chain may stay in the same state in some step. This reflects, that the embedded chain regards the underlying CTMC at the random times when a state change occurs, while the uniform chain consists on snap-shots of the CTMC at equidistant time steps small enough not to miss a state change of the process.

Remark 5.5.4 (Irreducible Markov Chains). The infinitesimal generator of an irreducible Markov chain is irreducible.

Recall that a state i is permanent if the ith diagonal element of Q is zero. The sign structure of Q implies then, that the entire row equals to zero, which means that Q is reducible. Then, by Proposition 5.3.11, the Markov chain is not irreducible. Conversely, there holds: An irreducible homogeneous Markov chain with finite state space has no permanent states.

Note that the stochastic matrices $\hat{P}$, $P(\tau)$ and P corresponding to an irreducible Markov chain are primitive. For $\hat{P}$ this holds only for λ large enough: Choose λ in (5.21) such that the diagonal entries of $\hat{P}$ are positive. Then we may apply [39, Lemma 8.5.5].

Finally, a CTMC on a finite state space is ergodic if and only if it is irreducible. This can be seen as follows: Assume, the transition semigroup of an ergodic chain is not irreducible, i. e., $p_{i,j}(t) = 0$ for all $t > 0$ and some states i and j. Since by definition of an ergodic chain the limit distribution $\boldsymbol{\pi} = \left[\pi_1 \ \ldots \ \pi_n\right]$ exists and is independently of the initial distribution, we get $\pi_j = \lim_{t\to\infty} p_{i,j}(t) = 0$, which contradicts the positivity of the limit distribution. The converse implication follows from Corollary 5.4.8.

Remark 5.5.5 (State Distributions). According to Proposition 5.4.1, the state distributions of the Markov chain are given by $\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)\, e^{tQ}$.

A stationary distribution of the CTMC as introduced in (5.5) can be characterized using Proposition 5.4.2 as a nonnegative solution of the homogeneous balance equation $\hat{\boldsymbol{\pi}}Q = 0$ subject to $\hat{\boldsymbol{\pi}}\mathbf{1} = 1$. Moreover, Proposition 5.4.3 guarantees, that such an stationary distribution exists.

An irreducible Markov chain has a uniquely determined stationary distribution, which is positive and coincides with the limit distribution (cf. Corollary 5.4.8).

5.6 Investigation of Markov Chains using Krylov Methods

Krylov subspace methods may be used to compute stationary distributions of a CTMC. This requires the solution of the homogeneous linear equation system

$$\boldsymbol{\pi} Q = \boldsymbol{0}^\top,$$

which may be rewritten using

$$A := -Q^\top, \qquad \boldsymbol{x} := \boldsymbol{\pi}^\top, \qquad \boldsymbol{b} := \boldsymbol{0},$$

such that the notation corresponds to that of Sections 3 and 4.

The treatment of this problem benefits from many characteristics of the generator matrix, which we restate here in summarized form and give links to the relevant results about Krylov subspace methods.

- A is a singular M-matrix with "property c".
- All eigenvalues of A lie in the closed right complex half plain.
- 0 is the only eigenvalue with vanishing real part.
- If A is irreducible, then 0 is an (algebraically and geometrically) simple eigenvalue.
- The linear system (3.1) is consistent.
- The nontrivial solutions of (3.1) are nonnegative and can be normalized such that

$$\boldsymbol{1}^\top \boldsymbol{x} = 1. \tag{5.27}$$

- There holds $\operatorname{index}(A) = 1$ and $\boldsymbol{b} \in \mathcal{R}(A)$.
  Thus, $A^\#$ exists and Theorem 4.4.6 applies. In particular, no singular breakdown occurs. Since $\boldsymbol{b} = \boldsymbol{0}$, there holds

$$\boldsymbol{x}_L^{\mathrm{MR}} = \boldsymbol{x}_L^{\mathrm{OR}} = A^\# \boldsymbol{b} + (I - A^\# A)\boldsymbol{x}_0 = (I - A^\# A)\boldsymbol{x}_0,$$

  i. e., the Krylov solution in the termination step is the projection of $\boldsymbol{x}_0$ onto the nullspace of A along $\mathcal{R}(A)$. This is a nontrivial solution of (3.1) if and only if $\boldsymbol{x}_0 \notin \mathcal{R}(A)$. Since $\boldsymbol{1} \in \mathcal{N}(A^*) \perp \mathcal{R}(A)$ we may choose $\boldsymbol{x}_0$ as a scalar multiple of $\boldsymbol{1}$ in order to get a nontrivial solution. The most popular choice is $\boldsymbol{x}_0 = \frac{1}{n}\boldsymbol{1} \in \mathbb{R}^n$.

- If $\dim \mathcal{N}(A) = 1$ (sufficiently, if A irreducible), there exists a unique solution $\boldsymbol{x}$ of (3.1) subject to the normalization (5.27). Moreover, since $\mathcal{R}(A^{\operatorname{index}(A)}) = \mathcal{R}(A) = \mathcal{N}(A^*)^\perp = \operatorname{span}\{\boldsymbol{1}\}^\perp$, the operation $\boldsymbol{v} - \frac{1}{n}(\boldsymbol{1}, \boldsymbol{v})\boldsymbol{1}$ provides a cheap possibility to project a vector $\boldsymbol{v}$ onto $\mathcal{R}(A^{\operatorname{index}(A)})$. We may apply this projection to the generated orthonormal basis of the Krylov space to avoid stability problems. It is then guaranteed, that the initial residual has index 0 after a restart (cf. Section 4.5).

We may also use Krylov subspace methods to compute approximately transient solutions, i. e., $\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)e^{tQ}$. As above let $A = -Q^\top$ and set $\boldsymbol{r} = \boldsymbol{\pi}(0)^\top$. We look for the MR approximation $\boldsymbol{s}$ for $e^{tA}\boldsymbol{r} = \boldsymbol{\pi}(t)^\top$ from $\mathcal{K}_m(A, \boldsymbol{r})$. The basic idea is to approximate the matrix exponential by $V_m^* e^{tH_m} V_m$ with V_m and H_m from the Arnoldi decomposition. This can be seen as one step of a single-step solver for the ordinary differential system (5.24). Note that the normalization $(\boldsymbol{1}, \boldsymbol{s}) = 1$ implies that the polynomial p_{m-1} corresponding to $\boldsymbol{s}$ is normalized according $p_{m-1}(0) = 1$, i. e., it is a residual polynomial. More details can be found in [67, Chapter 8] and the references therein.

6 Concluding Remarks

We have not studied in detail MR implementations relying on an orthonormal basis W_m of $\mathcal{W}_m$, such as GCR. When $\mathcal{W}_m = \mathcal{W}_{m-1}$ in some step m due to $\mathcal{C}_m \cap \mathcal{N}(A) \neq \{\boldsymbol{0}\}$ we have a singular breakdown. During the orthonormalization process this results in loss of orthogonality in the generated basis W_m. If we can detect this event, we may employ a look-ahead strategy similar to that in Remark 3.4.9 and Algorithm 3.4.11 as long as the sequence of correction spaces does not terminate. If, however, the correction spaces are Krylov subspaces, our results ensure $\mathcal{C}_m \cap \mathcal{N}(A) = \{\boldsymbol{0}\}$ as long as $m < L$. In either case the difficulty is to detect a singular breakdown. This requires the stable implementation of an orthogonal factorization for a matrix not having full column rank. It is not clear if the results of Strakoš et al. (cf. [57, 56]) apply for a singular A.

Krylov subspace methods can not be applied directly, if the matrix $A \in \mathbb{R}^{m\times n}$ of the linear system is not square. An obvious remedy is, to apply a standard iterative method to the normal equation $A^\top A\boldsymbol{x} = A^\top \boldsymbol{b}$. A more sophisticated idea is to introduce a $n\times m$ matrix B in place of $A^\top$ resulting in a (hopefully) better conditioning of the problem. Another possibility is to consider $AB\boldsymbol{z} = \boldsymbol{b}$, $B\boldsymbol{z} = \boldsymbol{x}$, which was first proposed by Zhang and Oyanagi (cf. [72]). In view of the results in Section 4, the matrix B should be determined such that the resulting square matrix is range hermitian. Methods of this type are studied by Ito and Hayami (cf. [46, 47]).

Matrix identities similar to those in Propositions 4.4.9 and 4.4.12 are studied by Cheng and Tian (cf. [18]). Their equations involve the Moore-Penrose and the group inverse and the objective is to characterize range-hermitian and normal matrices. Our conditions (4.28) and (4.29) and their investigation provide a first step toward generalizations of the concept of a range-hermitian matrix. More study is necessary to anticipate the ramifications of this idea.

The study of semigroups of stochastic matrices treated in Sections 5.3 and 5.4 can be extended in several directions. In particular, a more complete theory should involve generalized inverses, namely the group inverse of the infinitesimal generator. For example, there holds $Q^\# = \int_0^\infty (\Pi - P(t))\,\mathrm{d}t$ (cf. [20]), which is connected with the mean first passage times of the corresponding Markov chain.

Bibliography

[1] William J. Anderson. *Continuous-time Markov chains : an applications-oriented approach.* Springer Series in Statistics : Probability and its Applications. Springer, Berlin, Heidelberg, New York, 1991.

[2] Falko Bause, Peter Buchholz, and Peter Kemper. Hierarchically Combined Queueing Petri Nets. In *Proc. 11th Int. Conf. on Analysis and Optimizations of Systems*, Lecture Notes in Control and Information Sciences. Springer, 1994.

[3] Falko Bause and Pieter S. Kritzinger. *Stochastic Petri nets.* Vieweg, Wiesbaden, 1996.

[4] Adi Ben-Israel and Thomas N.E. Greville. *Generalized inverses: Theory and applications.* Pure and Applied Mathematics. John Wiley & Sons, New York etc., 1974.

[5] Abraham Berman and Robert J. Plemmons. *Nonnegative matrices in the mathematical sciences.* Computer Science and Applied Mathematics. Academic Press, New York, San Francisco, London, 1979.

[6] Pierre Brémaud. *Markov chains. Gibbs fields, Monte Carlo simulation, and queues.* Texts in Applied Mathematics. Springer, New York, NY, 1999.

[7] Peter N. Brown and Homer F. Walker. GMRES on (nearly) singular systems. *SIAM J. Matrix Anal. Appl.*, 18(1):37–51, 1997.

[8] Peter Buchholz. A class of hierarchical queueing networks and their analysis. *Queueing Systems*, 1(15):59–80, 1994.

[9] Peter Buchholz. Markovian Process Algebra: Composition and Equivalence. In U. Herzog and M. Rettelbach, editors, *Proc. of the 2nd Workshop on Process Algebras and Performance Modelling*, number 27 in Arbeitsberichte des IMMD, pages 11–30. University of Erlangen, 1994.

[10] Peter Buchholz. On Markovian Process Algebra. Forschungsbericht 500, Universität Dortmund, Fachbereich Informatik, 1994.

[11] Peter Buchholz. Equivalence Relations for Sochastic Automata Networks. In Stewart [68]. Proceedings of the 2nd international workshop on the numerical solution of Markov chains, Raleigh, NC, USA, January 16–18, 1995.

[12] Peter Buchholz. Structured analysis approaches for large Markov chains. *Applied Numerical Mathematics*, 31:375–404, 1999.

[13] Peter Buchholz. Multilevel solutions for structured Markov chains. *SIAM J. Matrix. Anal. Appl.*, 22(2):342–357, 2000.

[14] S. L. Campbell and C. D. Meyer, jr. *Generalized Inverses of Linear Transformations*. Dover Publications, Inc., New York, 1991.

[15] Raymond H. Chan. Iterative Methods for Overflow Queueing Models I. *Numerische Mathematik*, 51:143–180, 1987.

[16] Raymond H. Chan. Iterative Methods for Overflow Queueing Models II. *Numerische Mathematik*, 54:57–78, 1988.

[17] Raymond H. Chan and Wai Ki Ching. Circulant preconditioners for stochastic automata networks. *Numerische Mathematik*, 87(1):35–57, 2000.

[18] Shizhen Cheng and Yongge Tian. Two sets of new characterizations for normal and EP matrices. *Linear Algebra and its Applications*, 375:181–195, 2003.

[19] Kai Lai Chung. *Markov Chains*. Number 104 in Grundlehren der Mathematischen Wissenschaften. Springer, Göttingen, Berlin, Heidelberg, 1960.

[20] Pauline Coolen-Schrijner and Erik A. van Doorn. The deviation matrix of a continuous-time Markov chain. *Probab. Eng. Inf. Sci.*, 16(3):351–366, 2002.

[21] P.J. Courtois. *Decomposability: Queueing and computer system applications*. ACM Monograph Series. Academic Press, a subsidiary of Harcourt Brace Jovanovich, Publishers, New York, San Francisco, London, 1977.

[22] Michael Eiermann and Oliver G. Ernst. Geometric aspects of the theory of Krylov subspace methods. *Acta Numerica*, pages 251–312, 2001.

[23] Michael Eiermann, Oliver G. Ernst, and Olaf Schneider. Analysis of acceleration strategies for restarted minimal residual methods. *J. Comput. Appl. Math.*, 123(1-2):261–292, 2000.

[24] Michael Eiermann, Ivo Marek, and Wilhelm Niethammer. On the solution of singular linear systems of algebraic equations by semiiterative methods. *Numerische Mathematik*, 53:265–283, 1988.

[25] Oliver G. Ernst. *Minimal and Orthogonal Residual Methods and their Generalizations for Solving Linear Operator Equations*. Habilitationsschrift, Fakultät für Mathematik und Informatik der TU Bergakademie Freiberg, 2000.

[26] Bernd Fischer, Martin Hanke, and Marlis Hochbruck. A note on conjugate-gradient type methods for indefinite and/or inconsistent linear systems. *Numerical Algorithms*, 11:181–189, 1996.

[27] David Freedman. *Approximating Countable Markov Chains.* Springer, New York, Heidelberg, Berlin, 1983.

[28] David Freedman. *Markov Chains.* Springer, New York, Heidelberg, Berlin, 1983.

[29] Roland W. Freund and Marlis Hochbruck. On the use of two QMR algorithms for solving singular systems and applications in Markov chain modeling. *Numerical Linear Algebra with Applications*, 1:403–420, 1994.

[30] Gene Golub and Charles F. Van Loan. *Matrix computations.* The Johns Hopkins Univ. Press, Baltimore, MD, 3rd edition, 1996.

[31] Michael Greiner, Manfred R. Jobmann, and Claudia Kluppelberg. Telecommunication traffic, queueing models, and subexponential distributions. *Queueing Systems*, 33(1-3):125–152, 1999.

[32] C. W. Groetsch. *Generalized Inverses of Linear Operators.* Marcel Dekker, Inc., New York, Basel, 1977.

[33] Martin Hanke. *Conjugate Gradient Type Methods for Ill-Posed Problems.* Number 327 in Pitman Research Notes in Mathematics. Longman Scientific & Technical, Harlow, Essex, UK, 1995.

[34] Ken Hayami. On the behavior of the conjugate residual method for singular systems. In Zhong-Ci Shi and Hideo Kawarada, editors, *Proceedings of the Fifth China-Japan Seminar on Numerical Mathematics*, pages 117–126, Beijing/New York, 2002. Science Press. See also: National Institute of Informatics Technical Report NII-2001-002E, 2001.

[35] Ken Hayami and Masaaki Sugihara. On the convergence of GCR(k) method for singular systems. NII Technical Report NII-2004-009E, National Institute of Informatics, 2004.

[36] H. Hermanns and V. Mertsiotakis. A stochastic process algebra based modelling tool. In M. Merabti, M. Carew, and F. Ball, editors, *Performance Engineering of Computer and Telecommunications Systems*, pages 187–201. Springer, 1996.

[37] H. Hermanns, V. Mertsiotakis, and M. Rettelbach. Performance analysis of distributed systems using TIPP – a case study. In Pooley Hillston and King, editors, *UKPEW'94*, pages 131–144. University of Edinburgh, 1994.

[38] Harro Heuser. *Funktionalanalysis.* Mathematische Leitfäden. Teubner, Stuttgart, 3rd edition, 1992.

[39] Roger A. Horn and Charles R. Johnson. *Matrix analysis.* Cambridge University Press, repr. corr. edition, 1990.

[40] Roger A. Horn and Charles R. Johnson. *Topics in matrix analysis.* Cambridge University Press, 1st paperback edition, 1994.

[41] Bertram Huppert. *Angewandte Lineare Algebra.* de Gruyter, Berlin, New York, 1990.

[42] Ilse C. F. Ipsen and Steve Kirkland. Convergence analysis of a pagerank updating algorithm by Langville and Meyer. to appear in SIAM J. Matrix Anal. Appl.

[43] Ilse C. F. Ipsen and Steve Kirkland. Convergence analysis of an improved pagerank algorithm. Technical Report CRSC-TR04-02, North Carolina State University, 2004.

[44] Ilse C. F. Ipsen and Carl D. Meyer. The idea behind Krylov methods. *American Mathematical Monthly*, 105(10):889–899, 1998.

[45] Robert B. Israel, Jeffrey S. Rosenthal, and Jason Z. Wei. Finding generators for Markov chains via empirical transition matrices, with applications to credit ratings. *Mathematical Finance*, 11(2):245–265, 2001.

[46] Tokushi Ito and Ken Hayami. Preconditioned GMRES methods for least squares problems. NII Technical Report NII-2004-006E, National Institute of Informatics, 2004.

[47] Tokushi Ito and Ken Hayami. The solution of least squares problems using the GMRES method. In *Proceedings of the Annual Meeting of the Japan Society for Industrial and Applied Mathematics*, 2005.

[48] Richard L. Klevans and William J. Stewart. From queueing networks to Markov chains: the XMARCA interface. *Perform. Eval.*, 24(1-2):23–45, 1995.

[49] Ivo Marek and Petr Mayer. Convergence analysis of an iterative aggregation/disaggregation method for computing stationary probability vectors of stochastic matrices. *Numerical Linear Algebra with Applications*, 5:253–274, 1998.

[50] Nariyasu Minamide and Kahei Nakamura. A restricted pseudoinverse and its application to constrained minima. *SIAM J. appl. Math.*, 19:167–177, 1970.

[51] Dianne P. O'Leary. Iterative methods for finding the stationary vector for Markov chains. In Carl D. Meyer and Robert J. Plemmons, editors, *Linear Algebra, Markov Chains, and Queueing Models*, volume 48 of *The IMA Volumes in Mathematics and its Applications*, pages 125–133, New York etc., 1993. Springer.

[52] C.C. Paige and M.A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, 12:617–629, 1975.

[53] A. Pazy. *Semigroups of linear operators and applications to partial differential equations.* Applied Mathematical Sciences, 44. Springer, New York etc., 1983.

[54] Bernard Philippe, Yousef Saad, and William Stewart. Numerical methods in Markov chain modeling. *Oper. Res.*, 40(6):1156–1179, 1992.

[55] Bernard Philippe and Roger B. Sidje. Transient solutions of Markov processes by Krylov subspaces. In Stewart [68], pages 95–119. Proceedings of the 2nd international workshop on the numerical solution of Markov chains, Raleigh, NC, USA, January 16–18, 1995.

[56] Miroslav Rozložník, Z. Strakoš, and Miroslav Tůma. Numerical stability of GMRES. *BIT*, 35:309–330, 1995.

[57] Miroslav Rozložník, Zděnek Strakoš, and Miroslav Tůma. On the role of orthogonality in the GMRES method. In *Conference on Current Trends in Theory and Practice of Informatics*, pages 409–416, 1996.

[58] Yousef Saad. Projection methods for the numerical solution of Markov chain models. In Stewart [66], pages 455–471. Proceedings of the 1st International Workshop on the Numerical Solution of Markov Chains 1990.

[59] Yousef Saad. Preconditioned Krylov subspace methods for the numerical solution of Markov chains. In Stewart [68], pages 49–64. Proceedings of the 2nd international workshop on the numerical solution of Markov chains, Raleigh, NC, USA, January 16–18, 1995.

[60] Yousef Saad and Martin H. Schultz. GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems. *SIAM J. Sci. Stat. Comput.*, 7(3):856–869, 1986.

[61] Avram Sidi. A unified approach to Krylov subspace methods for the Drazin-inverse solution of singular nonsymmetric linear systems. *Linear Algebra and its Applications*, 298:99–113, 1999.

[62] Avram Sidi. A Bi-CG type iterative method for Drazin-inverse solution of singular inconsistent nonsymmetric linear systems of arbitrary index. *Electronic Journal of Linear Algebra*, 6:72–94, 2000.

[63] Avram Sidi. DGMRES: A GMRES-type algorithm for Drazin-inverse solution of singular inconsistent nonsymmetric linear systems. *Linear Algebra and its Applications*, 335:189–204, 2001.

[64] L. Smoch. Some results about GMRES in the singular case. *Numerical Algorithms*, 22:193–212, 1999.

[65] William J. Stewart. MARCA: Markov chain analyzer, a software package for Markov modeling. In Stewart [66], pages 37–61. Proceedings of the 1st International Workshop on the Numerical Solution of Markov Chains 1990.

[66] William J. Stewart, editor. *Numerical solution of Markov chains.* Probab., Pure Appl. 8. Marcel Dekker Inc., New York, 1991. Proceedings of the 1st International Workshop on the Numerical Solution of Markov Chains 1990.

[67] William J. Stewart. *Introduction to the numerical solution of Markov chains.* Princeton University Press, 1994.

[68] William J. Stewart, editor. *Computations with Markov chains.* Kluwer Academic Publishers, Boston, MA, 1995. Proceedings of the 2nd international workshop on the numerical solution of Markov chains, Raleigh, NC, USA, January 16–18, 1995.

[69] William J. Stewart and Wei Wu. Numerical experiments with iteration and aggregation for Markov chains. *ORSA Journal on Computing*, 4(3):336–350, 1992.

[70] Petar Todorovic. *An introduction to stochastic processes and their applications.* Springer Series in Statistics : Probability and its Applications. Springer, 1. print. edition, 1992.

[71] Per Åke Wedin. On angles between subspaces of a finite dimensional inner product space. In B. Kagstrom and A. Ruhe, editors, *Matrix Pencils*, Lecture Notes in Mathematics 973, pages 263–285. Springer, Berlin, Heidelberg, New York, 1982.

[72] Shao Liang Zhang and Yoshido Oyanagi. A necessary and sufficient convergence condition of Orthomin(k) methods for least squares problem with weight. *Annals of the Institute of Statistical Mathematics*, 42(4):805–811, 1990.