ON OPTIMAL SHORT RECURRENCES FOR GENERATING ORTHOGONAL KRYLOV SUBSPACE BASES

J. LIESEN† AND Z. STRAKOŠ‡

Abstract. We analyze necessary and sufficient conditions on a nonsingular matrix A, such that for *any* initial vector r_0, an orthogonal basis of the Krylov subspaces $\mathcal{K}_n(A, r_0)$ is generated by a short recurrence. Orthogonality here is meant with respect to some unspecified positive definite inner product. This question is closely related to the question of existence of *optimal* Krylov subspace solvers for linear algebraic systems, where optimal means the smallest possible error in the norm induced by the given inner product. The conditions on A we deal with were first derived and characterized more than 20 years ago by Faber and Manteuffel (SIAM J. Numer. Anal., 21 (1984), pp. 352–362). Their main theorem is often quoted and appears to be widely known. Its details and underlying concepts, however, are quite intricate, with some subtleties not covered in the literature we are aware of. Our paper aims to present and clarify the existing important results in the context of the Faber-Manteuffel Theorem. Furthermore, we review attempts to find an easier proof of the theorem, and explain what remains to be done in order to complete that task.

Key words. Krylov subspace methods, orthogonal bases, short recurrences, conjugate gradient-like methods.

AMS subject classifications. 65F10, 65F25, 15A57

1. Introduction.

Krylov subspace methods are powerful and widely used iterative methods for solving large and sparse linear algebraic systems, singular value and eigenvalue problems. They are based on subspaces spanned by an initial vector r_0 and vectors formed by repeated multiplication of r_0 by the given square matrix A. The use of these so-called Krylov subspaces,

$$\mathcal{K}_n(A, r_0) \equiv \text{span}\{r_0, Ar_0, \ldots, A^{n-1}r_0\}\,, \quad n = 1, 2, \ldots\,,$$

in iterative methods for linear algebraic systems is counted among the "Top 10 algorithmic ideas of the 20th century" [11, 12].

It is immediately clear that the basis of the Krylov subspace $\mathcal{K}_n(A, r_0)$, formed by the vectors $r_0, Ar_0 \ldots, A^{n-1}r_0$ is generally severely ill conditioned (recall the power method), and hence infeasible for use in actual computations. A well-conditioned, at best *orthogonal* basis is required in order to prevent loss of information due to repeated matrix-vector multiplication performed in finite precision arithmetic. For efficiency reasons, it is desirable to generate such a basis with a *short recurrence*, meaning that in each iteration step only a few of the latest basis vectors are required to generate the new basis vector. In this paper we discuss when (in exact arithmetic) orthogonal Krylov subspace bases can be generated by short recurrences. More precisely, given a Hermitian positive definite (HPD) matrix B, consider the corresponding B-inner product $\langle \cdot, \cdot \rangle_B$. We analyze necessary and sufficient conditions on a nonsingular matrix A, such that for *any* initial vector r_0, a B-orthogonal basis of $\mathcal{K}_n(A, r_0)$, $n = 1, 2, \ldots$, is generated by an *optimal short recurrence.* The precise definition of what we consider an optimal short recurrence is given in Section 2.

†Institute of Mathematics, Technical University of Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany (liesen@math.tu-berlin.de). The work of this author was supported by the Emmy Noether-Programm of the Deutsche Forschungsgemeinschaft.

‡Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod Vod. věží 2, 182 07 Prague, Czech Republic (strakos@cs.cas.cz). The work of this author was supported by the National Program of Research 'Information Society' under project 1ET400300415 and by the Institutional Research Plan AV0Z10300504.

The conditions on A we deal with in this paper are not new. In fact, they were derived and characterized more than 20 years ago by Faber and Manteuffel [15] (see also Greenbaum's book [21, Chapter 6] for a summary and some discussion of their results). Their work answered a question posed by Golub (SIGNUM Newsletter, vol. 16, no. 4, 1981), namely to construct a three-term *conjugate gradient-like descent method* for general nonsymmetric matrices, or to prove that there can be no such extension. Related questions were investigated by Voevodin and Tyrtyshnikov in [45]; see also [43, 44]. Their results, however, are not of the same strength as those of Faber and Manteuffel. Moreover, the considerations of Voevodin and Tyrtyshnikov are restricted to nonderogatory matrices.

Let us briefly describe what kind of method Golub had in mind (for a more complete description we refer to [15] or [21, Chapter 6]). Suppose that we want to solve a linear system $Ax = b$, where A is a square nonsingular matrix, by an iterative method. Starting from an initial guess x_0, consider the n-th iterate $x_n \in x_0 + \mathrm{span}\{p_0, \ldots, p_{n-1}\}$, $n = 1, 2, \ldots$, where the p_j are certain direction vectors, i.e. $x_n = x_{n-1} + \alpha_{n-1} p_{n-1}$, where α_{n-1} is some (nonzero) scalar coefficient that needs to be determined. This becomes a Krylov subspace method when $\mathrm{span}\{p_0, \ldots, p_{n-1}\} = \mathcal{K}_n(A, r_0)$, $n = 1, 2, \ldots$, where $r_0 = b - Ax_0$ is the initial residual. The n-th error of this method is given by $x - x_n \in x - x_0 + \mathrm{span}\{p_0, \ldots, p_{n-1}\}$. We speak of a conjugate gradient-like descent method, when the error is minimized in some given inner product norm, $\| \cdot \|_B = \langle \cdot, \cdot \rangle_B^{1/2}$, where B is a given HPD matrix, that is *independent* of the initial residual (see, e.g., [4] for a framework of methods whose norms depend on the initial residual). The condition that $\|x - x_n\|_B$ is minimal is equivalent to the error $x - x_n$ being B-orthogonal to the subspace spanned by the direction vectors. By construction, the n-th error can be written as

$$x - x_n \;=\; (x - x_{n-1}) - \alpha_{n-1} p_{n-1} \,.$$

Using this relation, the orthogonality conditions $\langle x - x_n, p_j \rangle_B = 0$, for $j = 0, \ldots, n-1$, translate into

$$\langle x - x_{n-1}, p_j \rangle_B - \alpha_{n-1} \langle p_{n-1}, p_j \rangle_B \;=\; 0 \,, \quad \text{for } \; j = 0, \ldots, n-1 \,.$$

This set of n equations is satisfied if and only if

$$\alpha_{n-1} \;=\; \frac{\langle x - x_{n-1}, p_{n-1} \rangle_B}{\langle p_{n-1}, p_{n-1} \rangle_B} \,, \quad \textit{and} \quad \langle p_{n-1}, p_j \rangle_B \;=\; 0 \;\text{ for }\; j = 0, \ldots, n-2 \,,$$

i.e. the direction vectors $p_0, \ldots, p_{n-1}$ must form a B-orthogonal set. If B is wisely chosen, α_{n-1} can be computed even though $x - x_{n-1}$ is unknown (note that x is unknown); see [3] or [43] for frameworks containing numerous different methods. In the conjugate gradient method [24], A is HPD, $B = A$, and the A-orthogonal, or "conjugate" set of direction vectors is generated by orthogonalizing (with respect to the A-inner product) the Krylov sequence $r_0, Ar_0, \ldots, A^{n-1} r_0$ by means of what we call an optimal three-term recurrence. More generally, the existence of optimal short recurrences is closely related to the existence of efficient implementations of conjugate gradient-like descent methods, and hence to the question asked by Golub.

The answer to Golub's question given by Faber and Manteuffel [15], known as the Faber-Manteuffel Theorem, is that for most non-Hermitian matrices A there exists no extension of the conjugate gradient method based on a single short (let alone three-term) recurrence. More precisely, apart from rather special classes of matrices

specified below, for general non-Hermitian matrices A there exists no optimal short recurrence for generating B-orthogonal Krylov subspace bases (for any given HPD matrix B). This fundamental message of the Faber-Manteuffel Theorem is often quoted and appears to be widely known. The details and underlying concepts, however, are quite intricate, with some subtleties not covered in the literature we are aware of, including [15], [21, Chapter 6], and the recent paper [36].

In [36], three different matrix properties are studied in the context of optimal short recurrences: *A admits for the given B an optimal $(s+2)$-term recurrence*, *A is B-normal(s)*, and *A is reducible for the given B to $(s+2)$-band Hessenberg form*. Using this approach as a starting point, we clarify some inaccuracies, give more rigorous definitions of all three properties (see Definitions 2.4, 2.6, and 2.11, resp.), and review the known relations between them; see Fig. 2.2. We hereby present and clarify the existing important results in the context of the Faber-Manteuffel Theorem. We prove strengthened versions of the *sufficiency* of the B-normal(s) property of the matrix A for the other two properties; see the arrows labelled "Theorem 2.9" and "Theorem 2.13" in Fig. 2.2. Furthermore, we investigate the B-normal(s) property of A, give a new equivalent characterization, improve the bound on the degree of the minimal polynomial of A in terms of s for B-normal(s) matrices, and provide several examples.

While the *sufficiency* of the B-normal(s) property of A for the two other properties can be shown by a rather straightforward and completely algebraic proof, the proof of *necessity* of this property given by Faber and Manteuffel [15] is based on a clever, highly nontrivial construction. The proof uses a continuity argument to extend certain "easy" cases to "difficult" cases, where the "difficult" cases form a set of measure zero in the space of all cases. So far, this proof has not even been reproduced in any of the numerous surveys and books on iterative methods. It is unknown if a simpler proof of the necessity part can be found. In view of the fundamental nature of the Faber-Manteuffel Theorem, such proof would be a welcome addition to the existing literature. It would lead to a better understanding of the theorem by enlightening some (possibly unexpected) relationships, and it would also be more suitable for classroom teaching. In this paper, we discuss possible strategies for finding such a proof.

We point out that the theory developed by Faber and Manteuffel, as well as most results in this paper, only apply to what we call optimal short recurrences. Given the negative implications of the Faber-Manteuffel Theorem, several attempts have been made for finding other types of short recurrences for generating orthogonal Krylov subspace bases. For completeness, we briefly review these attempts as well.

The paper is organized as follows. In Section 2, we define the concept of optimal short recurrences. We then prove sufficient conditions for their existence, which strengthen previous results (Section 2.1). Next, we consider the necessary conditions, and suggest possible approaches for proving the necessity part in a simpler way (Section 2.2). In Section 3, we characterize the B-normal(s) property of A. In Section 4, we review results on other types of short recurrences for generating orthogonal Krylov subspace bases. We end with concluding remarks in Section 5.

REMARK 1.1. *Throughout the paper we consider exact arithmetic and for simplicity, with an application to linear algebraic solvers in mind, we make the following assumptions: The matrix A is nonsingular and an element of $\mathbb{F}^{N\times N}$, where either $\mathbb{F}=\mathbb{R}$ or $\mathbb{F}=\mathbb{C}$. If $\mathbb{F}=\mathbb{R}$, i.e. A is a real (nonsingular) matrix, we consider only real HPD matrices $B\in\mathbb{R}^{N\times N}$, and real initial vectors $r_0\in\mathbb{R}^N$. In this case state-*

ments like "for any r_0" should be read "for any $r_0 \in \mathbb{R}^N$", the resulting polynomials have real coefficients, and the conjugate transpose, denoted by v^* for a vector v and M^* for a matrix M, coincides with the transpose.

2. Optimal short recurrences. Let A be a given nonsingular $N \times N$ matrix, and let r_0 be any nonzero vector of length N. It is well known that the Krylov subspaces $\mathcal{K}_n(A, r_0)$, $n = 1, 2, \ldots$, form a nested sequence of subspaces of increasing dimension that eventually become invariant under A. Hence there exists an index

$$d \;=\; d(A, r_0) \;\equiv\; \dim \mathcal{K}_N(A, r_0)\,, \tag{2.1}$$

often called the *grade of r_0 with respect to A*, for which

$$\mathcal{K}_1(A, r_0) \;\subset\; \ldots \;\subset\; \mathcal{K}_d(A, r_0) \;=\; \mathcal{K}_{d+1}(A, r_0) \;=\; \ldots \;=\; \mathcal{K}_N(A, r_0)\,.$$

For a given vector r_0, d is equal to the degree of the minimal polynomial of r_0 with respect to A, see [18, Chapter VII, § 1, § 2 and § 8]. Clearly, $d \leq d_{\min}(A)$, where $d_{\min}(A)$ denotes the degree of the minimal polynomial of A, and there always exists an r_0 such that $d = d(A, r_0) = d_{\min}(A)$, see, e.g., [18, Chapter VII, § 2, Theorem 2] or [1, Section 3].

For any $N \times N$ HPD matrix B, the functional $\langle \cdot, \cdot \rangle_B$, defined for vectors x and y of length N by $\langle x, y \rangle_B \equiv y^* B x$, is a positive definite inner product. For a nonsingular matrix A, a given HPD matrix B, and a vector r_0 of grade d, consider a B-orthogonal basis $\widehat{v}_1, \ldots, \widehat{v}_d$,

$$\mathrm{span}\,\{\widehat{v}_1, \ldots, \widehat{v}_n\} = \mathcal{K}_n(A, r_0)\,, \qquad n = 1, \ldots, d\,, \tag{2.2}$$

$$\langle \widehat{v}_j, \widehat{v}_k \rangle_B = 0\,, \qquad j \neq k\,,\;\; j, k \;=\; 1, \ldots, d\,. \tag{2.3}$$

Such a set of basis vectors is generated by the Arnoldi recurrence [2]

$$\widehat{v}_1 = r_0\,, \tag{2.4}$$

$$\widehat{v}_{n+1} = A\widehat{v}_n \;-\; \sum_{m=1}^{n} \widehat{h}_{m,n} \widehat{v}_m\,, \qquad \widehat{h}_{m,n} \;=\; \frac{\langle A\widehat{v}_n, \widehat{v}_m \rangle_B}{\langle \widehat{v}_m, \widehat{v}_m \rangle_B}\,, \tag{2.5}$$

$$n = 1, \ldots, d-1\,, \qquad d = \dim \mathcal{K}_N(A, r_0)\,, \tag{2.6}$$

stated here with the classical Gram-Schmidt orthogonalization. Its basic properties can be found, e.g., in [40, Section 6.3]. Note that we have skipped normalization of the basis vectors for notational convenience. It is easy to see that the matrix representation of (2.4)–(2.6) is given by

$$\widehat{v}_1 = r_0\,, \tag{2.7}$$

$$A \underbrace{[\widehat{v}_1, \ldots, \widehat{v}_{d-1}]}_{\equiv \widehat{V}_{d-1}} = \underbrace{[\widehat{v}_1, \ldots, \widehat{v}_d]}_{\equiv \widehat{V}_d} \underbrace{\begin{bmatrix} \widehat{h}_{1,1} & \cdots & \widehat{h}_{1,d-1} \\ 1 & \ddots & \vdots \\ & \ddots & \widehat{h}_{d-1,d-1} \\ & & 1 \end{bmatrix}}_{\equiv \widehat{H}_{d,d-1}}\,, \tag{2.8}$$

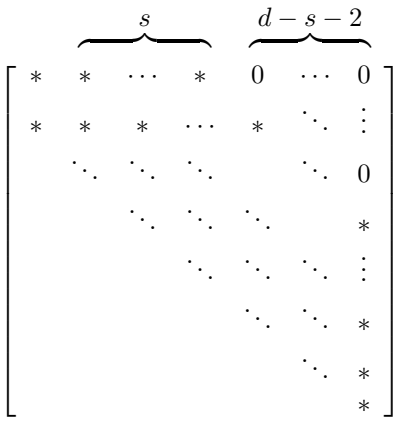

FIG. 2.1. *The band structure of an* $(s+2)$*-band Hessenberg matrix of size* $d\times(d-1)$*. All entries above its* s*-th superdiagonal are zero, and at least one entry in its* s*-superdiagonal is nonzero.*

$$\widehat{V}_d^* B \widehat{V}_d \text{ is diagonal}, \quad d = \dim \mathcal{K}_N(A, r_0)\,. \tag{2.9}$$

We point out that the whole basis $\widehat{v}_1, \ldots, \widehat{v}_d$ is generated in $d-1$ steps of (2.5), which yields a (non-square) matrix $\widehat{H}_{d,d-1}$ of size $d\times(d-1)$. Hence (2.8) does not represent a *reduction* of A to Hessenberg form (see (2.16) below for such a reduction).

Any other basis $v_1, \ldots, v_d$ satisfying (2.2)–(2.3) can be obtained by scaling the columns of $\widehat{V}_d$ by some $d \times d$ nonsingular diagonal matrix S_d,

$$V_d \;\equiv\; [v_1, \ldots, v_d] \;=\; \widehat{V}_d S_d\,.$$

The matrix V_d then satisfies the identity

$$AV_{d-1} \;=\; V_d\, H_{d,d-1}\,, \quad \text{where } H_{d,d-1} \;=\; S_d^{-1} \widehat{H}_{d,d-1} S_{d-1}\,, \tag{2.10}$$

and S_{d-1} is the $(d-1)\times(d-1)$ leading principal submatrix of S_d. Clearly, with any such scaling, the nonzero pattern of $H_{d,d-1}$ is uniquely defined and identical to the nonzero pattern of $\widehat{H}_{d,d-1}$. In the following we will be mostly interested in this pattern, particularly in the upper bandwidth of $\widehat{H}_{d,d-1}$.

DEFINITION 2.1. *An unreduced upper Hessenberg matrix is called* $(s+2)$*-band Hessenberg, when its* s*-th superdiagonal contains at least one nonzero entry, and all its entries above its* s*-th superdiagonal are zero.*

REMARK 2.2. Extension of this definition to the case $s = 0$ (2-band Hessenberg), where the 0-th superdiagonal means the diagonal, includes a lower bidiagonal matrix with all entries on its subdiagonal nonzero, and at least one nonzero entry on its diagonal.

REMARK 2.3. The square matrices A and B are both of the same size $N \times N$. In order to simplify the statements, in the rest of the paper we will not repeat this fact again.

Let the matrix $\widehat{H}_{d,d-1}$ be $(s+2)$-band Hessenberg, cf. Fig. 2.1. Then for each $n = 1, \ldots, d-1$, (2.5) reduces to

$$\widehat{v}_{n+1} \;=\; A\widehat{v}_n \;-\; \sum_{m=\max\{n-s,1\}}^{n} \widehat{h}_{m,n}\widehat{v}_m\,, \qquad \widehat{h}_{m,n} \;=\; \frac{\langle A\widehat{v}_n, \widehat{v}_m\rangle_B}{\langle \widehat{v}_m, \widehat{v}_m\rangle_B}\,, \tag{2.11}$$

and the B-orthogonal Krylov subspace basis $\widehat{v}_1, \ldots, \widehat{v}_d$ is then generated by an $(s+2)$-*term recurrence*. Since precisely the last $s+1$ basis vectors $\widehat{v}_n, \ldots, \widehat{v}_{n-s}$ are required to determine $\widehat{v}_{n+1}$ (and not just any collection of $s+1$ previous basis vectors), and only one matrix-vector multiplication with A is performed, we call an $(s+2)$-term recurrence of the form (2.11) an *optimal* $(s+2)$*-term recurrence*. We stress that in the following we will be only concerned with this type of recurrence. There exist other short recurrences which are not of the form (2.11). For example, in order to generate $\widehat{v}_{n+1}$, the *isometric Arnoldi algorithm* [20] subtracts from $A\widehat{v}_n$ not only a linear combination of previously generated orthogonal basis vectors, but also a linear combination of some of them multiplied by A. A brief review of results on various types of short recurrences for generating orthogonal Krylov subspace bases is, for completeness, given in Section 4.

Optimal $(s+2)$-term recurrences of the form (2.11) are highly desirable, since they conveniently limit work and storage requirements for generating the B-orthogonal basis vectors. Given an HPD matrix B and a small s, it is therefore essential to understand for which matrices A, (2.7)–(2.9) leads for *any* initial vector r_0 with $d \geq s+2$ to a matrix $\widehat{H}_{d,d-1}$ which is at most $(s+2)$-band Hessenberg.

DEFINITION 2.4. *Let A be a nonsingular matrix with minimal polynomial degree $d_{\min}(A)$. Let B be an HPD matrix, and let s be a nonnegative integer, $s+2 \leq d_{\min}(A)$.*

1. *If for an initial vector r_0 the matrix $\widehat{H}_{d,d-1}$ in (2.7)–(2.9) is $(s+2)$-band Hessenberg, then we say that A admits for the given B and r_0 an optimal $(s+2)$-term recurrence.*
2. *If A admits for the given B and any initial vector r_0 an optimal recurrence of length at most $s+2$, while it admits for the given B and at least one r_0 an optimal $(s+2)$-term recurrence, then we say that A admits for the given B an optimal $(s+2)$-term recurrence.*

It is appropriate to comment on some subtleties of this definition. First, the definition intentionally distinguishes a property that holds for the given A, B, s and a particular *given* r_0 (item (1)) from a property that holds for the given A, B, s and *all* r_0 (item (2)). This distinction has not been made, to our knowledge, in the previous literature, which led to some ambiguities and inaccuracies. Second, consistent with Remark 2.2, s is assumed nonnegative; there can be no 0- or 1-term recurrences. Third, no recurrence of the form (2.4)–(2.6) can produce more than $d_{\min}(A)$ linearly independent vectors. Therefore it is meaningless to consider $s+2 > d_{\min}(A)$.

In practice, $d_{\min}(A)$ is usually very large. Given B, we are interested in conditions on A, so that it admits an optimal $(s+2)$-term recurrence with $s \ll d_{\min}(A)$.

2.1. Sufficient conditions. Consider a nonsingular matrix A, an HPD matrix B, and a nonnegative integer s with $s+2 \leq d_{\min}(A)$. Our goal is to derive sufficient conditions on A so that it admits for the given B an optimal $(s+2)$-term recurrence.

REMARK 2.5. If $s+2 = d_{\min}(A)$, then A *can* admit an optimal recurrence of length at most $s+2$. It *does* admit an optimal $(s+2)$-term recurrence, if there exists

an initial vector r_0 with $d = s+2$, such that the upper right element $\widehat{h}_{1,d-1}$ of $\widehat{H}_{d,d-1}$ is nonzero. As we will see, this property is nontrivial. Until this point is clarified, we must include the otherwise uninteresting case $s+2 = d_{\min}(A)$ in our considerations. It will be dropped later in this section; see Remark 2.8.

Let r_0 be any initial vector with $d \geq s+2$. If A admits for the given B and r_0 an optimal $(s+2)$-term recurrence, then the entries $\widehat{h}_{m,n}$ of $\widehat{H}_{d,d-1}$ in (2.7)–(2.9) must satisfy

$$\begin{aligned} (2.12)\quad \widehat{h}_{m,n} &= 0\,, \quad \text{whenever} \quad 1 \leq m < n-s\,, \qquad n = 1,\ldots,d-1\,, \ \text{i.e.,} \\ \widehat{h}_{m,n} &= 0\,, \quad \text{whenever} \quad m+s < n \leq d-1\,, \quad m = 1,\ldots,d\,. \end{aligned}$$

From (2.5) it follows that

$$0 \;=\; \widehat{h}_{m,n} \;=\; \frac{\langle A\widehat{v}_n, \widehat{v}_m\rangle_B}{\langle \widehat{v}_m, \widehat{v}_m\rangle_B}\,,$$

if and only if

$$(2.13)\quad 0 = \langle A\widehat{v}_n, \widehat{v}_m\rangle_B \;=\; \widehat{v}_m^* B A\widehat{v}_n \;=\; (B^{-1}A^*B\widehat{v}_m)^*\, B\widehat{v}_n \;=\; \langle \widehat{v}_n, A^+\widehat{v}_m\rangle_B\,,$$

where the matrix

$$(2.14)\qquad A^+ \;\equiv\; B^{-1}A^*B$$

is usually called the *B-adjoint of A*.

Now assume that, for the given B, $A^+ = p_s(A)$, where p_s is a polynomial of degree s, and, for clarity, no polynomial with smaller degree and the same property exists. Then

$$A^+\widehat{v}_m \;=\; p_s(A)\,\widehat{v}_m \;\in\; \mathcal{K}_{m+s}(A, \widehat{v}_1)\,.$$

For $n > m+s$, the vector $\widehat{v}_n$ is B-orthogonal to $\mathcal{K}_{m+s}(A, \widehat{v}_1)$ by construction, so that $\langle \widehat{v}_n, A^+\widehat{v}_m\rangle_B = 0$, giving $\widehat{h}_{m,n} = 0$. The condition that $A^+ = p_s(A)$ is worth a formal definition.

Definition 2.6. *Let A be a nonsingular matrix, and let B be an HPD matrix. Suppose that*

$$(2.15)\qquad A^+ \;\equiv\; B^{-1}A^*B \;=\; p_s(A)\,,$$

where p_s is a polynomial of the smallest possible degree s having this property. Then A is called normal of degree s with respect to B, or, shortly, B-normal(s).

The term *B-normal(s)* appears to be standard in this context; cf., e.g., [3, Section 2]. We emphasize that in our definition the B-normal(s) property of A refers to the given HPD matrix B, and s is uniquely determined. In particular, contrary to the usage of this term in the previous literature, here if A is B-normal(s), then A *not* B-normal(t) for any $t \neq s$. In Section 3 below we show, in addition to other things, that not only s, but also the polynomial p_s of the smallest possible degree for which $A^+ = p_s(A)$ is *uniquely* determined (see Theorem 3.1). Here we get the following result.

Lemma 2.7. *Let A be a nonsingular matrix with minimal polynomial degree $d_{\min}(A)$. Let B be an HPD matrix, and let s be a nonnegative integer, $s+2 \leq d_{\min}(A)$.*

If A is B-normal(s), then for any r_0 with $d \geq s+2$ the corresponding matrix $\widehat{H}_{d,d-1}$ is $(s+2)$-band Hessenberg.

Proof. Let r_0 be any initial vector of grade $d \geq s+2$. Using (2.15) and (2.7)–(2.9), there exist a nonzero scalar ζ and a vector $w \in \mathcal{K}_s(A, \widehat{v}_1)$, such that $A^+\widehat{v}_1 = p_s(A)\widehat{v}_1 = \zeta \widehat{v}_{s+1} + w$. Using the B-orthogonality of $\widehat{v}_{s+1}$ to $\mathcal{K}_s(A, \widehat{v}_1)$ yields

$$\widehat{h}_{1,s+1} \;=\; \frac{\langle A\widehat{v}_{s+1}, \widehat{v}_1\rangle_B}{\langle \widehat{v}_1, \widehat{v}_1\rangle_B} \;=\; \frac{\langle \widehat{v}_{s+1}, A^+\widehat{v}_1\rangle_B}{\langle \widehat{v}_1, \widehat{v}_1\rangle_B} \;=\; \overline{\zeta}\, \frac{\langle \widehat{v}_{s+1}, \widehat{v}_{s+1}\rangle_B}{\langle \widehat{v}_1, \widehat{v}_1\rangle_B} \;\neq\; 0\,,$$

and the s-th superdiagonal of $\widehat{H}_{d,d-1}$ is nonzero. Since (2.12) and the considerations following it show that $\widehat{H}_{d,d-1}$ is at most $(s+2)$-band Hessenberg, the proof is finished. □

REMARK 2.8. *After Definition 2.4 we have already noted that the question of whether A admits for the given B an optimal $(s+2)$-term recurrence is meaningless when $s+2 > d_{\min}(A)$.*

Furthermore, if $s+2 = d_{\min}(A)$, then, given any HPD matrix B, (2.12) is trivially satisfied for any r_0, i.e., A admits for any B an optimal recurrence of length at most $s+2$. If A is, moreover, B-normal(s), then A admits for the given B and any r_0 with $d{=}s+2$ an optimal $(s+2)$-term recurrence (in particular, the element $\widehat{h}_{1,d-1}$ of $\widehat{H}_{d,d-1}$ is nonzero, cf. the proof of Lemma 2.7). Therefore A admits for the given B an optimal $(s+2)$-term recurrence.

In the theorems throughout the rest of this section we will therefore exclude the special and in practice uninteresting case $s+2 = d_{\min}(A)$, and consider $s+2 < d_{\min}(A)$.

The discussion of the sufficient conditions on A can be summarized as follows.

THEOREM 2.9. *Let A be a nonsingular matrix with minimal polynomial degree $d_{\min}(A)$. Let B be an HPD matrix, and let s be a nonnegative integer, $s+2 < d_{\min}(A)$. If A is B-normal(s), then A admits for the given B and any initial vector r_0 an optimal recurrence of length at most $s+2$, while for any r_0 of grade with respect to A at least $s+2$, it admits an optimal $(s+2)$-term recurrence. Therefore, A being B-normal(s) represents a sufficient condition for A to admit for the given B an optimal $(s+2)$-term recurrence.*

Note that Theorem 2.9 is stronger than the sufficiency result of Faber and Manteuffel in [15]: In our notation, they show that if A is B-normal(s), then for any r_0 the corresponding optimal recurrence has *at most* $s+2$ terms (cf. [15, pp. 355–357]; also cf. the sufficiency proof of [21, Theorem 6.1.1, p. 99]). Theorem 2.9, on the other hand, states that for any r_0 with $d \geq s+2$, the corresponding optimal recurrence has *exactly* $s+2$ terms.

2.2. Necessary conditions. In our notation, Faber and Manteuffel consider a given B and $s+2 < d_{\min}(A)$. They show that if A admits for the given B and any initial vector r_0 an optimal recurrence of length at most $s+2$, then A is normal of degree at most s with respect to B (cf. [15, p. 359–361], or [21, Theorem 6.1.1]). Considering Definition 2.4 and Theorem 2.9, we can make the statement of Faber and Manteuffel a little stronger. Suppose that A admits for the given B an optimal $(s+2)$-term recurrence. Then by [15], A must be normal of degree at most s with respect to B. If A is B-normal(t) for some $t < s$, then by Theorem 2.9, A admits for

the given B an optimal $(t+2)$-term recurrence, which is a contradiction. Hence A must be B-normal(s), and s is the smallest nonnegative integer with this property.

THEOREM 2.10. *Let A be a nonsingular matrix with minimal polynomial degree $d_{\min}(A)$. Let B be an HPD matrix, and let s be a nonnegative integer, $s+2 < d_{\min}(A)$. If A admits for the given B an optimal $(s+2)$-term recurrence, then A is B-normal(s).*

As described in the Introduction, the proof of Theorem 2.10 given by Faber and Manteuffel in [15] is based on a clever, highly nontrivial construction. Finding an easier proof (that possibly avoids continuity and topological arguments) is an interesting research problem. Some work in this direction has recently been done in [36], where the authors extend earlier ideas of Voevodin and Tyrtyshnikov [44, 45]. It turns out, however, that the exposition in [36] is not fully accurate, and some claims made there are incorrect. We will now summarize and clarify the approach from [36].

Consider A and B as above, and let r_0 be any initial vector of grade d with respect to A, see (2.1). By construction, $A\widehat{v}_d \in \mathcal{K}_d(A, \widehat{v}_1)$, and

$$A\widehat{v}_d \;=\; \sum_{m=1}^{d} \widehat{h}_{m,d}\, \widehat{v}_m\,, \quad \text{where} \quad \widehat{h}_{m,d} \;=\; \frac{\langle A\widehat{v}_d, \widehat{v}_m\rangle_B}{\langle \widehat{v}_m, \widehat{v}_m\rangle_B}\,, \quad m = 1, \ldots, d\,.$$

This relation can be used to extend the matrix equation (2.8),

$$A \underbrace{[\widehat{v}_1, \ldots, \widehat{v}_d]}_{\equiv \widehat{V}_d} \;=\; \underbrace{[\widehat{v}_1, \ldots, \widehat{v}_d]}_{\equiv \widehat{V}_d} \underbrace{\begin{bmatrix} \widehat{h}_{1,1} & \cdots & \widehat{h}_{1,d-1} & \widehat{h}_{1,d} \\ 1 & \ddots & \vdots & \vdots \\ & \ddots & \widehat{h}_{d-1,d-1} & \widehat{h}_{d-1,d} \\ & & 1 & \widehat{h}_{d,d} \end{bmatrix}}_{\equiv \widehat{H}_d}\,. \tag{2.16}$$

In (2.16), the matrix A is orthogonally *reduced* to upper Hessenberg form. Analogously to Definition 2.4 we get:

DEFINITION 2.11. *Let A be a nonsingular matrix with minimal polynomial degree $d_{\min}(A)$. Let B be an HPD matrix, and let s be a nonnegative integer, $s+2 \le d_{\min}(A)$.*

(1) If for an initial vector r_0 the matrix $\widehat{H}_d$ in (2.16) is $(s+2)$-band Hessenberg, then we say that A is reducible for the given B and r_0 to $(s+2)$-band Hessenberg form.

(2) If A is reducible for the given B and any initial vector r_0 to at most $(s+2)$-band Hessenberg form, while it is reducible for the given B and at least one r_0 to $(s+2)$-band Hessenberg form, then we say that A is reducible for the given B to $(s+2)$-band Hessenberg form.

This definition is more rigorous than the one given in [36, p. 2151].

REMARK 2.12. *Note that it is possible to extend Definition 2.11 to the case $s+1 = d_{\min}(A)$ (in contrast to Definition 2.4, where this case is meaningless). However, for clarity we have chosen to exclude this special case, and to unify the assumptions in both definitions. In order to be consistent with our main focus, see Remark 2.8, we will not further consider the special case $s+2 = d_{\min}(A)$, and continue with the assumption $s+2 < d_{\min}(A)$.*

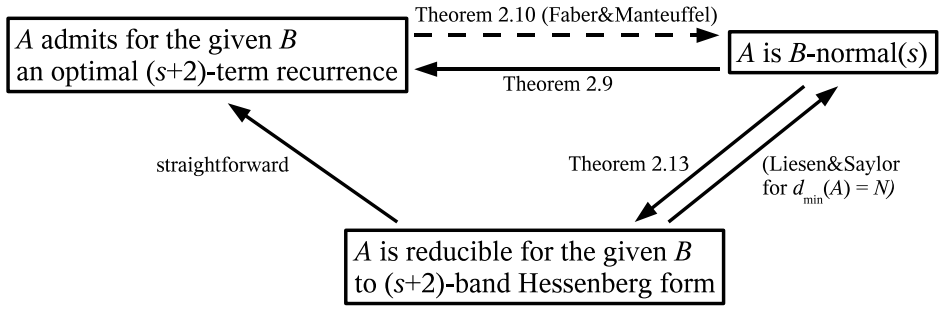

FIG. 2.2. *Implications between the different properties of a nonsingular matrix A for a given HPD matrix B and a nonnegative integer s, $s+2 < d_{\min}(A)$. The numbers of the theorems in this paper are shown at the respective arrows. Solid arrows indicate known linear algebra based proofs, and the dashed arrow indicates the proof of Theorem 2.10 given by Faber and Manteuffel [15].*

Clearly, if $s+2 < d_{\min}(A)$, and A is reducible for the given B to $(s+2)$-band Hessenberg form, then A admits for the given B an optimal $(s+2)$-term recurrence. If, on the other hand, A admits for the given B an optimal $(s+2)$-term recurrence, then, without further analysis, we can say nothing about the possible zero elements in the last column of $\widehat{H}_d$. In particular, one cannot immediately conclude that A is reducible for the given B to $(s+2)$-band Hessenberg form. A simple modification of the proof of Theorem 2.9, however, gives a strengthened version of results published before (cf. [36, Section 3], and [25, Theorem 5], where the author considers the case $B = I$).

THEOREM 2.13. *Let A be a nonsingular matrix with minimal polynomial degree $d_{\min}(A)$. Let B be an HPD matrix, and let s be a nonnegative integer, $s+2 < d_{\min}(A)$. If A is B-normal(s), then A is reducible for the given B and any initial vector r_0 to at most $(s+2)$-band Hessenberg form, while for any r_0 of grade with respect to A at least $s+2$, it is reducible to $(s+2)$-band Hessenberg form. Therefore, A being B-normal(s) represents a sufficient condition for A to be reducible for the given B to $(s+2)$-band Hessenberg form.*

An overview of the known implications between the three different matrix properties studied in this paper is shown in Fig. 2.2. Theorems 2.9 and 2.10 comprise what is known as the Faber-Manteuffel Theorem [15]. Using the necessity part (Theorem 2.10) as well as Theorem 2.13 yields the following important equivalence, which in particular shows that A being B-normal(s) also represents the necessary condition for reducibility of A to $(s+2)$-band Hessenberg form.

THEOREM 2.14. *Let A be a nonsingular matrix with minimal polynomial degree $d_{\min}(A)$. Let B be an HPD matrix, and let s be a nonnegative integer, $s+2 < d_{\min}(A)$. Then the following three assertions are equivalent:*

(1) A admits for the given B an optimal $(s+2)$-term recurrence.

(2) A is B-normal(s).

(3) A is reducible for the given B to $(s+2)$-band Hessenberg form.

Proof. By Theorem 2.13, (2) implies (3). As explained above, the implication from (3) to (1) is straightforward. Finally, the equivalence is closed by Theorem 2.10, which shows that (1) implies (2). □

Apart from a direct attempt of finding a simpler proof of Theorem 2.10 (the

necessity part of the Faber-Manteuffel Theorem [15]), one might possibly consider the following approach. If a nonsingular and *nonderogatory* matrix A (i.e. $d_{\min}(A) = N$) is reducible for the given HPD matrix B to $(s+2)$-band Hessenberg form, then A is B-normal(s), see [36]. (Note that on the contrary to the claim in [36, p. 2154], the proof of this implication given there does not apply to general nonsingular matrices.) A somewhat related result has appeared earlier in [45]; the more widely cited paper [44] only contains statements of theorems without proofs. If an extension of this result from nonderogatory to general nonsingular matrices is found, then an alternative, and possibly simpler proof of the Faber-Manteuffel Theorem might be completed by proving (in an elementary way) the missing implication from "A admits for the given B an optimal $(s+2)$-term recurrence" to "A is reducible for the given B to $(s+2)$-band Hessenberg form", see Fig 2.2. Summarizing, in this approach the equivalence of the assertions from Theorem 2.14 would result from the implications: (2) $\Rightarrow$ (1) (Theorem 2.9), (1) $\Rightarrow$ (3) (needs to be proved), (3) $\Rightarrow$ (2) (needs to be extended to nonderogatory A).

3. Equivalent characterizations. In this section we study the property that A is B-normal(s). We start with a general characterization.

THEOREM 3.1. *Let A be a nonsingular matrix, and let B be an HPD matrix. Then the following two assertions are equivalent:*

(1) A is B-normal(s).

(2 a) A is diagonalizable with the eigendecomposition $A = W\Lambda W^{-1}$ (without loss of generality we consider the eigenvalues and eigenvectors of A ordered so that equal eigenvalues form a single diagonal block in Λ),

and

b) using the eigenvector matrix W of A, the matrix B^{-1} has the decomposition $B^{-1} = WDW^*$, where D is an HPD block diagonal matrix with block sizes corresponding to those of Λ,

and

c) there exists a polynomial p_s of degree s such that $p_s(\Lambda) = \Lambda^*$, and s is the smallest degree of all polynomials with this property. The polynomial p_s is uniquely determined.

Proof. Suppose that A is B-normal(s), and let p_s be a polynomial of smallest possible degree giving $A^+ = p_s(A)$. Then, an elementary computation shows that

$$B^{-1/2}A^*B^{1/2} \;=\; p_s(B^{1/2}AB^{-1/2})\,,$$

so that $B^{-1/2}A^*B^{1/2}$ commutes with its adjoint, and so it is normal, and hence unitarily diagonalizable,

$$B^{-1/2}A^*B^{1/2} \;=\; U\Lambda^*U^*\,, \tag{3.1}$$

where $U^*U = UU^* = I$. Then from (3.1),

$$A \;=\; (B^{-1/2}U)\Lambda(B^{-1/2}U)^{-1} \;=\; W\Lambda W^{-1}\,, \quad W \equiv B^{-1/2}U\,. \tag{3.2}$$

Consequently, A is diagonalizable and condition (2a) is satisfied, where, without loss of generality, the diagonal elements of Λ and the columns of U and W are correspondingly ordered. Using (3.1) and the eigendecomposition of A in (3.2),

$$\begin{aligned}(B^{-1/2}U)\,\Lambda^*\,(B^{-1/2}U)^{-1} &= B^{-1/2}\,(B^{-1/2}A^*B^{1/2})\,B^{1/2} \;=\; A^+ \;=\; p_s(A)\\ &= (B^{-1/2}U)\,p_s(\Lambda)\,(B^{-1/2}U)^{-1}\,,\end{aligned}$$

which proves $p_s(\Lambda) = \Lambda^*$. If there exists a polynomial p_t with degree $t < s$ and $p_t(\Lambda) = \Lambda^*$, then we get $A^+ = p_t(A)$, which contradicts the minimality assumption on s. We now show that p_s is uniquely determined. From $p_s(\Lambda) = \Lambda^*$ we get $s \leq d_{\min}(A) - 1$, since the interpolating polynomial of degree $d_{\min}(A) - 1$, that has value $\overline{\lambda}$ at each eigenvalue λ of A, has the desired property. If there exist polynomials $p_s^{(1)}$ and $p_s^{(2)}$ with $p_s^{(1)}(\Lambda) = p_s^{(2)}(\Lambda) = \Lambda^*$, then their difference $p_s^{(1)} - p_s^{(2)}$ is a polynomial of degree at most s having $d_{\min}(A)$ distinct zeros. Since $s \leq d_{\min}(A) - 1$, we must have $p_s^{(1)} - p_s^{(2)} = 0$. We thus have shown that condition (2c) holds. For condition (2b) realize that

$$\begin{aligned} B^{-1}\,(W^{-*}\Lambda^* W^*)\,B = B^{-1}A^*B \;&=\; A^+ \;=\; p_s(A) \;=\; W p_s(\Lambda) W^{-1} \\ &= W\Lambda^* W^{-1}\,, \end{aligned}$$

giving $\Lambda^* = (W^*BW)\Lambda^*(W^*BW)^{-1}$. Clearly, the columns of the matrix W^*BW represent eigenvectors of the diagonal matrix Λ^*. Consequently, a column of W^*BW can have nonzero entries only in the positions corresponding to the block part of Λ^* determined by the related eigenvalue. If all eigenvalues of Λ^* (and so all eigenvalues of A) are simple, then W^*BW must be diagonal. In general, $W^*BW \equiv D^{-1}$ is an HPD block diagonal matrix, where the block sizes correspond to the multiplicities of the individual eigenvalues, which proves condition (2b).

On the other hand, assume that conditions (2a)–(2c) hold. From $A = W\Lambda W^{-1}$ and $p_s(\Lambda) = \Lambda^*$ we receive

$$A^+ = B^{-1}A^*B \;=\; (B^{-1}W^{-*})\,\Lambda^*\,(W^*B) \;=\; (B^{-1}W^{-*})\,p_s(\Lambda)\,(W^*B)\,.$$

Substituting $B = (WDW^*)^{-1}$ gives $A^+ = W(Dp_s(\Lambda)D^{-1})W^{-1} = Wp_s(\Lambda)W^{-1} = p_s(A)$. The minimality of s for which this holds is given by construction. ∎

Theorem 3.1 gives conditions on A *and* B such that A is B-normal(s). Now consider a nonsingular *diagonalizable* matrix $A = W\Lambda W^{-1}$, where we use the block ordering of the eigenvalues of A on the diagonal of Λ as in condition (2a). Then we define the class of matrices satisfying condition (2b),

$$\mathcal{B} \;\equiv\; \left\{ (WDW^*)^{-1}\,,\; \begin{array}{l} D \text{ is an HPD block diagonal matrix} \\ \text{with the sizes of its blocks} \\ \text{corresponding to the blocks of } \Lambda \end{array} \right\}. \tag{3.3}$$

As shown in the proof of Theorem 3.1, a polynomial of minimal degree s satisfying condition (2c) is nothing but the (unique) interpolating polynomial $\mathcal{L}$ satisfying $\mathcal{L}(\lambda_j) = \overline{\lambda}_j$ for all eigenvalues λ_j of A. Clearly, $p_s \equiv \mathcal{L}$ is of degree $s \leq d_{\min}(A) - 1$, and s is uniquely determined by this construction. Note that for every $B \in \mathcal{B}$, $A^+ = B^{-1}A^*B = W\Lambda^*W^{-1}$. We summarize these consequences of Theorem 3.1 in the following corollary.

COROLLARY 3.2. *Let A be a nonsingular matrix, and let B be an HPD matrix, such that A is B-normal(s). Then, using the notation of Theorem 3.1, $A = W\Lambda W^{-1}$, $s \leq d_{\min}(A) - 1$ is uniquely determined by the location of the eigenvalues of A, $B \in \mathcal{B}$ as defined in (3.3), and $A^+ = W\Lambda^*W^{-1}$.*

We have seen that s is determined by the location of the eigenvalues of A. The case $s = d_{\min}(A) - 1$ is in the context of the existence of optimal $(s+2)$-term recurrences meaningless, and we have simplified the statements of the theorems in Section 2 by

excluding also the case $s = d_{\min}(A) - 2$, see Remark 2.8. In practice we are interested only in s for which $s \ll d_{\min}(A)$. The following result, first obtained in a different, but mathematically equivalent formulation by Faber and Manteuffel [15, Lemma 3] (also cf. [21, Theorem 6.1.3, p. 101]), characterizes the two smallest possible values of s, namely $s = 0$ and $s = 1$.

THEOREM 3.3. *Let A be a nonsingular matrix.*
(1) There exists an HPD matrix B for which A is B-normal(0) if and only if $A = \alpha I$ for some nonzero $\alpha \in \mathbb{C}$.
(2) There exists an HPD matrix B for which A is B-normal(1) if and only if A is diagonalizable with $d_{\min}(A) \geq 2$ and A has collinear eigenvalues (i.e. all eigenvalues lie on a single straight line in the complex plane).

Proof. If the matrix A is B-normal(0) for some given HPD matrix B, then from Theorem 3.1, A must be diagonalizable, and there exists a polynomial p_0 of degree zero, say $p_0(z) \equiv \overline{\alpha}$ for some nonzero $\alpha \in \mathbb{C}$, that satisfies $p_0(\lambda_j) = \overline{\alpha} = \overline{\lambda}_j$ for all eigenvalues λ_j of A. Clearly, $A = \alpha I$. The other implication is a straightforward consequence of Theorem 3.1. Note that $A = \alpha I$ is B-normal(0) for *any* HPD matrix B.

Now suppose that A is B-normal(1) for some given HPD matrix B. From Theorem 3.1, A must be diagonalizable, and there exists a polynomial $p_1(z) \equiv \alpha + \beta z$, $\alpha, \beta \in \mathbb{C}$ with $\beta \neq 0$, such that $p_1(\lambda_j) = \overline{\lambda}_j$ for all eigenvalues $\lambda_1, \ldots, \lambda_m$ of A. Since one is the minimal degree of a polynomial with this property, A must have at least two distinct eigenvalues and $d_{\min}(A) \geq 2$. If A has exactly two distinct eigenvalues, then they are trivially collinear. Otherwise we determine the coefficient β using any two of the distinct eigenvalues of A, say λ_1 and λ_2,

$$\beta \;=\; \frac{\overline{\lambda}_2 - \overline{\lambda}_1}{\lambda_2 - \lambda_1}\,.$$

Clearly, $|\beta| = 1$ and we write for convenience $\beta = e^{\iota(2\varphi)}$, $\varphi \in [0, \pi)$. The coefficient β and therefore the angle φ are uniquely determined independent of the choice of the (distinct) eigenvalues above. We will now rotate the complex plane by the angle φ and show that after this rotation all rotated eigenvalues $e^{\iota\varphi}\lambda_j$, $j = 1, \ldots, m$, are located on a single line parallel to the real axis, which proves that λ_j, $j = 1, \ldots, m$, are located on the inversely rotated line. Indeed, using

$$\alpha + e^{2\iota\varphi}\lambda_j \;=\; \overline{\lambda}_j\,,$$

we easily get

$$2\iota\, Im(e^{\iota\varphi}\lambda_j) \;=\; e^{\iota\varphi}\lambda_j - e^{-\iota\varphi}\overline{\lambda}_j \;=\; -e^{-\iota\varphi}\alpha\,,$$

i.e., the imaginary part of $e^{\iota\varphi}\lambda_j$ is a constant independent of the index j.

Conversely, suppose that the distinct eigenvalues $\lambda_1, \ldots, \lambda_m$, where $m \geq 2$, of the diagonalizable and nonsingular matrix A are collinear. Then there exist $\omega \in \mathbb{C}$ and $\varphi \in [0, \pi)$ such that $\lambda_j = \omega + \varrho_j e^{\iota\varphi}$ for some $\varrho_j \in \mathbb{R}$, $j = 1, \ldots, m$. An easy computation shows that the degree one polynomial

$$p_1(z) \;\equiv\; (\overline{\omega} - e^{-2\iota\varphi}\omega) + e^{-2\iota\varphi} z$$

satisfies $p_1(\lambda_j) = \overline{\lambda}_j$ for $j = 1, \ldots, m$. Since $m \geq 2$, the same property cannot hold for any polynomial of degree zero. Consequently, A is B-normal(1) for any HPD matrix $B \in \mathcal{B}$, see (3.3). ∎

We now return to the question of the existence of the optimal $(s+2)$-term recurrences and assume that $s+2 < d_{\min}(A)$. Consider $s=1$ and a nonsingular matrix A with $3 < d_{\min}(A)$. Theorems 2.9 and 2.10 show that A admits for a given HPD B an optimal 3-term recurrence if and only if A is B-normal(1). Theorem 3.3 then shows that there exists an HPD matrix B for which A is B-normal(1), and, consequently, A admits an optimal 3-term recurrence, if and only if A is diagonalizable with collinear eigenvalues. Well known classes of diagonalizable matrices with collinear eigenvalues are the Hermitian and skew-Hermitian matrices. Note that these matrices are unitarily diagonalizable, which results, with the choice $D=I$ in (3.3), in $B=I$.

Interesting examples of matrices A and $B \neq I$ for which A is B-normal(1) are given in the context of saddle point problems in [17, 9], with a generalization presented in [34]. Here

$$A \;=\; \begin{bmatrix} A_1 & A_2^T \\ -A_2 & A_3 \end{bmatrix} \;\in\; \mathbb{R}^{(m+k)\times(m+k)}\,, \tag{3.4}$$

where $A_1 \in \mathbb{R}^{m\times m}$ is symmetric positive definite, $A_2 \in \mathbb{R}^{k\times m}$ has full rank $k \leq m$, and $A_3 \in \mathbb{R}^{k\times k}$ is symmetric positive semidefinite (possibly zero). An elementary computation shows that $A^T \neq p_1(A)$ for any polynomial p_1 of degree one, i.e. A is *not* I-normal(1). On the other hand, the symmetric matrix

$$B \;=\; \begin{bmatrix} A_1 - \gamma I & A_2^T \\ A_2 & \gamma I - A_3 \end{bmatrix}\,, \quad \gamma \equiv \frac{1}{2}\lambda_{\min}(A_1)\,, \tag{3.5}$$

satisfies $BA = A^T B$. As shown in [9] for $A_3 = 0$, and in [34] for a symmetric positive semidefinite A_3, the matrix B is positive definite when

$$\lambda_{\min}(A_1) \;>\; 4\,(\,\lambda_{\max}(A_3) + \lambda_{\max}(A_2 A_1^{-1} A_2^T)\,)\,.$$

If this condition is satisfied, then $A^+ = B^{-1}A^T B = A$, and A is B-normal(1).

We will now prove a new bound on the degree of the minimal polynomial $d_{\min}(A)$ in terms of s for general B-normal(s) matrices. Consider a nonsingular *diagonalizable* matrix A. As shown above, A is normal of degree s with respect to *any* HPD matrix $B \in \mathcal{B}$, see (3.3), where s is the smallest degree of a polynomial p_s for which $p_s(\Lambda) = \Lambda^*$, see condition (2c) in Theorem 3.1. Equivalently, s is determined as the smallest degree of a polynomial p_s such that the eigenvalues of A are roots of the harmonic polynomial[1] $p_s(z) - \overline{z}$. From Theorem 3.3, $s=1$ if and only if $d_{\min}(A) \geq 2$ and the eigenvalues of A are collinear. If $d_{\min}(A) > 2$ and the eigenvalues of A are *not* collinear, which is the case we are interested in here, then s must be larger than one. As shown by Khavinson and Świątek using techniques of complex dynamics, a harmonic polynomial $p_s(z) - \overline{z}$ with $s > 1$ may have *at most* $3s-2$ roots [31, Theorem 1] (see [28] for an elementary proof of this result for $s=2$). Recently, Geyer [19] has shown that for all $s > 1$ this bound on the maximal number of roots is sharp (see also [36, Example 3.7] for the case $s=3$).

The result of [31, Theorem 1] has the following fundamental consequence: Consider a nonsingular diagonalizable matrix A with eigenvalues that are not collinear. Then A cannot be B-normal(1) for any HPD matrix B. Consider, in addition, an integer $s \geq 2$. If $d_{\min}(A) > 3s-2$, there exists no polynomial p_t of degree $t \leq s$

[1] A harmonic polynomial is a function of the form $p(z) + \overline{q(z)}$, where p and q are polynomials, see, e.g., [41, Section 1.1.7].

that satisfies $p_t(\Lambda) = \Lambda^*$, since the eigenvalues of such an A cannot be roots of *any* harmonic polynomial $p_t(z) - \overline{z}$ with $t \leq s$. By Theorem 3.1, such a matrix A cannot be normal of degree $t \leq s$ with respect to *any* given B. An alternative way to state this consequence of [31, Theorem 1] is that if a matrix A is normal of degree $s > 1$ with respect to some given B, then $d_{\min}(A) \leq 3s - 2$. Using this fact together with Theorem 2.10 yields the following important result, which improves previous results in [15, Lemma 4] and [21, Theorem 6.1.2, p. 100] (there the bound on $d_{\min}(A)$ is s^2 instead of $3s - 2$).

THEOREM 3.4. *Let A be a nonsingular matrix with eigenvalues that are not collinear, and let s be a positive integer greater than one. If the degree of the minimal polynomial of A satisfies $d_{\min}(A) > 3s - 2$, then there exists no HPD matrix B for which A admits an optimal recurrence of length at most $s + 2$.*

In practice $s \ll d_{\min}(A)$. Consequently, except for the diagonalizable matrices having collinear eigenvalues, there exists no practically interesting matrix A (with sufficiently large $d_{\min}(A)$) and no HPD matrix B such that A admits for B an optimal $(s + 2)$-term recurrence with a small s.

Finally, we remark that even when the matrix A fails to be diagonalizable, it may still admit for *some* HPD matrix B and *some* initial vector r_0 an optimal $(s+2)$-term recurrence with small s. For example, consider the transposed $N \times N$ Jordan block

$$A = \begin{bmatrix} \lambda & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{bmatrix}, \quad \lambda \neq 0\,.$$

For $B = I$, and $r_0 = [1, 0, \ldots, 0]^T \in \mathbb{R}^N$, the matrix representation of the recurrence (2.4)–(2.6) is

$$\widehat{v}_1 = r_0\,,$$

$$A \underbrace{\begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ 0 & \cdots & 0 \end{bmatrix}}_{\equiv \widehat{V}_{N-1}} = \underbrace{\begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}}_{\equiv \widehat{V}_N} \underbrace{\begin{bmatrix} \lambda & & \\ 1 & \ddots & \\ & \ddots & \lambda \\ & & 1 \end{bmatrix}}_{\equiv \widehat{H}_{N,N-1}},$$

$$\widehat{V}_N^T \widehat{V}_N = I\,, \quad N = \dim \mathcal{K}_N(A, r_0)\,.$$

The matrix $\widehat{H}_{N,N-1}$ is 2-band Hessenberg, and hence A admits for $B = I$ and $r_0 = [1, 0, \ldots, 0]^T$ an optimal 2-term recurrence. We stress, however, that the existence of this particular recurrence is only due to the special relationship between A, B and r_0. Since the $N \times N$ Jordan block A is not diagonalizable, there exists no HPD matrix B for which A admits an optimal $(s + 2)$-term recurrence with $s < N - 2$.

4. Other types of short recurrences. Throughout the previous sections we have considered short recurrences of the form (2.11), called optimal $(s + 2)$-term recurrences, where only the latest $s + 1$ (mutually orthogonal) vectors $\widehat{v}_n, \ldots, \widehat{v}_{n-s}$,

and only one matrix-vector multiplication with A are required to generate the next basis vector $\widehat{v}_{n+1}$. Theorem 3.4 shows that such optimal $(s+2)$-term recurrences with $s \ll d_{\min}(A)$ do not exist for most non-Hermitian matrices A.

Motivated by this situation, some attempts have been made to find other types of short recurrences for generating orthogonal Krylov subspace bases. This work has been inspired by the existence of the *isometric Arnoldi algorithm*, originally discovered by Gragg [20][2]. This algorithm has deep connections with Gauss quadrature and orthogonal polynomials on the unit circle, in particular with the classical theory of Szegö [42] (see [20] or [46] for more on these relations, and [10] for a thorough study of implementation details). It allows the generation of orthogonal Krylov subspace bases for a unitary matrix U using two coupled two-term recurrences. When written in the form of a single recurrence (which, as shown in [6], is not advisable from a numerical standpoint), these two recurrences become

$$\sigma_n \widehat{v}_{n+1} \;=\; U\widehat{v}_n - \frac{\gamma_n \sigma_{n-1}}{\gamma_{n-1}} U\widehat{v}_{n-1} + \frac{\gamma_n}{\gamma_{n-1}} \widehat{v}_n \,, \tag{4.1}$$

where the $\widehat{v}_j$ are the orthogonal Krylov subspace basis vectors, and σ_j, γ_j are some scalar coefficients. Clearly, the recurrence (4.1) is not of the form (2.11): Either we have to perform an additional multiplication with U, or we have to store the vector $U\widehat{v}_{n-1}$ in addition to $\widehat{v}_n$ and $\widehat{v}_{n-1}$. Nevertheless, the existence of (4.1) shows that an orthogonal Krylov subspace basis for U can be generated by some form of short recurrence, although U in general is *not* B-normal(s) for any HPD matrix B and any small s. (Note, in particular, that the eigenvalues of U are in general not collinear, and therefore U is not B-normal(1) for any HPD matrix B.) The eigenvectors of U can be chosen orthonormal, $I \in \mathcal{B}$ (see (3.3)), and hence it is natural to investigate the I-normality of U. It turns out, that any unitary matrix U is I-normal(t) for $t = d_{\min}(U) - 1$ [35].

The recurrence (4.1) is a special case of an *$(s+2, t)$-term recurrence* of the form

$$\widehat{v}_{n+1} \;=\; A\widehat{v}_n \;-\; \sum_{m=\max\{n-t,1\}}^{n-1} \widehat{g}_{m,n} A\widehat{v}_m \;-\; \sum_{m=\max\{n-s,1\}}^{n} \widehat{h}_{m,n} \widehat{v}_m \,, \tag{4.2}$$

which has been considered by Barth and Manteuffel in a series of papers [5, 6, 7]. Clearly, the previously considered optimal $(s+2)$-term recurrence of the form (2.11) corresponds to an $(s+2, 0)$-term recurrence of the form (4.2). Hence in the context of $(s+2, t)$-term recurrences, the only cases of additional interest are those with $t > 0$. As shown by Barth and Manteuffel in [6], a recurrence of the form (4.2) for generating B-orthogonal Krylov subspace bases for a nonsingular matrix A exists, if $A^+ = p_s(A)(q_t(A))^{-1}$ for polynomials p_s and q_t of respective degrees s and t. A partial characterization of necessary conditions is given in [7].

In case of a unitary matrix U, we can take $B = I$ and $U^+ = U^* = U^{-1}$, so that $p_s(z) = 1$ and $q_t(z) = z$, i.e. $s = 0$ and $t = 1$, which yields a $(2, 1)$-term recurrence. Written in the form of two coupled two-term recurrences, this is nothing but the above mentioned isometric Arnoldi algorithm. Another example with $t > 0$

[2] Gragg's paper [20] appeared in 1993, but he presented its results already during a visit to Moscow State University in 1981. Subsequently, a Russian version of [20] appeared in the proceedings E. S. Nikolaev, ed., *Numerical Methods in Linear Algebra*, Moscow University Press, Moscow, 1982, pp. 16–32. Algebraically, the isometric Arnoldi algorithm relies on an efficient LR factorization of unitary upper Hessenberg matrices; see [39] for an early description of this idea.

is given by the shifted unitary matrices of the form $A = U + \zeta I$ with a nonzero $\zeta \in \mathbb{C}$ and U unitary. A straightforward calculation shows that any such matrix satisfies $A^* = p_1(A)(q_1(A))^{-1}$, where $p_1(z) = \overline{\zeta}z + (1 - |\zeta|^2)$ and $q_1(z) = z - \zeta$. The sufficiency result of Barth and Manteuffel implies that for shifted unitary matrices there exists a (3,1)-term recurrence of the form (4.2). This generalization of the isometric Arnoldi algorithm has been, prior to the work of Barth and Manteuffel, employed by Jagels and Reichel [29, 30] for constructing a minimal residual method for solving linear systems with shifted unitary matrices. It is easily seen that $(2,1)$-term (resp. $(3,1)$-term) recurrences of the form (4.2) exist for all matrices that are similar to unitary (resp. shifted unitary) matrices; different similarity transformations yield different matrices B, but they do not alter the length of the recurrence.

Beyond these classes of matrices, however, the practical relevance of the $(s+2,t)$-term recurrences is rather limited. It follows from [6, Theorem 3.1], that for a given matrix A with sufficiently large $d_{\min}(A)$, an HPD matrix B exists such that $A^+ = p_s(A)(q_t(A))^{-1}$ with small degrees s and t, if and only if either A is B-normal(1) (and hence $s = 1$, $t = 0$), or A is similar to a (shifted) unitary matrix (and hence $s \in \{0,1\}$, $t = 1$; B in this case being determined by the similarity transformation); see [35] for recent related work.

Barth and Manteuffel [5, 6, 7] have also studied a more general class of matrices that satisfy $A^+ = p_s(A)(q_t(A))^{-1} + R$, where R is a low rank matrix. More recently, Beckermann and Reichel [8] have derived a short recurrence Arnoldi type algorithm for generating orthogonal Krylov subspace bases in case $A^* = A + R$, where R has low rank.

Another line of work has been initiated by the generalized Lanczos algorithm of Elsner and Ikramov [13], which can be considered an extension of the Hermitian Lanczos algorithm [32] to normal matrices, where $AA^* = A^*A$, or, equivalently, $A^* = p(A)$ for some polynomial p (see [22] for numerous additional equivalent definitions). Hence a normal matrix is I-normal(t) for some nonnegative integer t. The generalized Lanczos algorithm of Elsner and Ikramov has been exploited and further developed by Huhtanen, see, e.g., [26, 27], and, more recently, by Faßbender and Ikramov [16].

5. Concluding remarks. In this paper we have aimed at a concise and rigorous discussion of the mathematical concepts and main results concerning optimal short recurrences (as specified in Definition 2.4) for generating orthogonal Krylov subspace bases. Some results in this paper represent strengthened versions of previous results in the literature (in particular Theorems 2.9, 2.10, 2.13, 2.14, and 3.4), and some appear to be new, at least in the form presented here (in particular Theorem 3.1 and Corollary 3.2). An emphasis has been placed on the relationship between the three main matrix properties of interest in this context (see Fig. 2.2), and on possible approaches to finding an easier proof of necessity in the Faber-Manteuffel Theorem. Based on the presentation in this paper, further work in this direction has recently been done by the first author jointly with Faber and Tichý [14].

Throughout the paper we have assumed exact arithmetic. Hence we have treated the computation of an orthogonal Krylov subspace basis from a purely mathematical point of view. In actual implementations of the considered recurrences, two main points, which too often are overlooked, need to be considered:

First, orthogonalization in any short recurrence method is performed with respect to a few of the previous basis vectors only. Global orthogonality between the computed basis vectors is obtained as an implicit mathematical consequence of the explicitly enforced local orthogonality. Such orthogonality properties, derived under the

assumption of exact arithmetic, however, do not hold in finite precision computations. As a consequence, short recurrence methods are *inherently numerically unstable*, regardless of the conditioning of the HPD matrix B that defines the inner product. Any attempt to construct or use short recurrences in practical applications should therefore be accompanied by a thorough numerical stability analysis; for symmetric problems and $s = 1$ see, e.g., the recent survey [38] and the recent book [37].

Second, from a numerical point of view, it is often advisable to implement a *single* short recurrence such as (2.11) or (4.2) in the form of *coupled* short recurrences. For examples we refer to [6], and to [23], where an analysis of the numerical differences between three-term and mathematically equivalent coupled two-term recurrences is given; see, in a different framework, also [33].

Acknowledgments. We thank Petr Tichý for many invaluable comments and suggestions during our work on this paper. We thank Mark Embree, Vance Faber, Tom Manteuffel, Chris Paige, and three anonymous referees for numerous suggestions that helped us to clarify some points and to improve the presentation. Further helpful comments were made by Michele Benzi, Volker Mehrmann, Lothar Reichel, Valeria Simoncini, and Daniel Szyld.

REFERENCES

[1] M. Arioli, V. Pták, and Z. Strakoš, *Krylov sequences of maximal length and convergence of GMRES*, BIT, 38 (1998), pp. 636–643.

[2] W. E. Arnoldi, *The principle of minimized iteration in the solution of the matrix eigenvalue problem*, Quart. Appl. Math., 9 (1951), pp. 17–29.

[3] S. F. Ashby, T. A. Manteuffel, and P. E. Saylor, *A taxonomy for conjugate gradient methods*, SIAM J. Numer. Anal., 27 (1990), pp. 1542–1568.

[4] T. Barth and T. Manteuffel, *Variable metric conjugate gradient methods*, in Advances in Numerical Methods for Large Sparse Sets of Linear Equations, Number 10, Matrix Analysis and Parallel Computing, PCG 94, M. Natori and T. Nodera, eds., Keio University, Yokohama, 1994, pp. 165–188.

[5] ———, *Conjugate gradient algorithms using multiple recursions*, in Proceedings of the AMS-IMS-SIAM Summer Research Conference held at the University of Washington, Seattle, WA, July 9–13, 1995, L. Adams and J. L. Nazareth, eds., Philadelphia, 1996, SIAM, pp. 107–123.

[6] ———, *Multiple recursion conjugate gradient algorithms. I. Sufficient conditions*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 768–796.

[7] ———, *Multiple recursion conjugate gradient algorithms. II. Necessary conditions*, unpublished manuscript, (2000).

[8] B. Beckermann and L. Reichel, *The Arnoldi process and GMRES for nearly symmetric matrices*, tech. report, Université de Lille, submitted, 2006.

[9] M. Benzi and V. Simoncini, *On the eigenvalues of a class of saddle point matrices*, Numer. Math., 103 (2006), pp. 173–196.

[10] B. Bohnhorst, *Beiträge zur numerischen Behandlung des unitären Eigenwertproblems*, PhD thesis, Fakultät für Mathematik, Universität Bielefeld, 1993.

[11] B. A. Cipra, *The best of the 20th century: Editors name top 10 algorithms*, SIAM News, 33 (2000).

[12] J. Dongarra and F. Sullivan, *The top 10 algorithms in the 20th century (Guest editors introduction)*, Comp. Sci. Eng., 2 (2000), pp. 22–23.

[13] L. Elsner and K. D. Ikramov, *On a condensed form for normal matrices under finite sequences of elementary unitary similarities*, Linear Algebra Appl., 254 (1997), pp. 79–98.

[14] V. Faber, J. Liesen, and P. Tichý, *The Faber-Manteuffel Theorem for linear operators*, submitted, 2006.

[15] V. Faber and T. Manteuffel, *Necessary and sufficient conditions for the existence of a conjugate gradient method*, SIAM J. Numer. Anal., 21 (1984), pp. 352–362.

[16] H. Fassbender and K. D. Ikramov, *SYMMLQ-like procedure for $Ax = b$ where A is a special normal matrix*, Calcolo, 43 (2006), pp. 17–37.

[17] B. Fischer, A. Ramage, D. J. Silvester, and A. J. Wathen, *Minimum residual methods for augmented systems*, BIT, 38 (1998), pp. 527–543.

[18] F. R. Gantmacher, *The Theory of Matrices. Vols. 1, 2*, Chelsea Publishing Co., New York, 1959.
[19] L. Geyer, *Sharp bounds for the valence of certain harmonic polynomials*, Tech. Report arXiv:math.CV/0510539, 2005.
[20] W. B. Gragg, *Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle*, J. Comput. Appl. Math., 46 (1993), pp. 183–198.
[21] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, vol. 17 of Frontiers in Applied Mathematics, SIAM, Philadelphia, PA, 1997.
[22] R. Grone, C. R. Johnson, E. M. de Sá, and H. Wolkowicz, *Normal matrices*, Linear Algebra and its Applications, 87 (1987), pp. 213–225.
[23] M. H. Gutknecht and Z. Strakoš, *Accuracy of two three-term and three two-term recurrences for Krylov space solvers*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 213–229.
[24] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Research Nat. Bur. Standards, 49 (1952), pp. 409–436 (1953).
[25] T. Huckle, *The Arnoldi method for normal matrices*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 479–489.
[26] M. Huhtanen, *A Hermitian Lanczos method for normal matrices*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 1092–1108.
[27] ———, *Orthogonal polyanalytic polynomials and normal matrices*, Math. Comp., 72 (2003), pp. 355–373.
[28] K. D. Ikramov, *The matrix adjoint to a normal matrix A as a polynomial in A*, Dokl. Akad. Nauk, 338 (1994), pp. 304–305.
[29] C. Jagels and L. Reichel, *The isometric Arnoldi process and an application to iterative solution of large linear systems*, in Iterative methods in linear algebra (Brussels, 1991), North-Holland, Amsterdam, 1992, pp. 361–369.
[30] C. F. Jagels and L. Reichel, *A fast minimal residual algorithm for shifted unitary matrices*, Numer. Linear Algebra Appl., 1 (1994), pp. 555–570.
[31] D. Khavinson and G. Świątek, *On the number of zeros of certain harmonic polynomials*, Proc. Amer. Math. Soc., 131 (2003), pp. 409–414.
[32] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Research Nat. Bur. Standards, 45 (1950), pp. 255–282.
[33] D. P. Laurie, *Questions related to Gaussian quadrature formulas and two-term recursions*, in Applications and computation of orthogonal polynomials (Oberwolfach, 1998), vol. 131 of Internat. Ser. Numer. Math., Birkhäuser, Basel, 1999, pp. 133–144.
[34] J. Liesen, *A note on the eigenvalues of saddle point matrices*, Technical Report 10-2006, TU Berlin, Institute of Mathematics, 2006.
[35] ———, *When is the adjoint of a matrix a low degree rational function in the matrix?*, Technical Report 23-2006, TU Berlin, Institute of Mathematics, 2006.
[36] J. Liesen and P. E. Saylor, *Orthogonal Hessenberg reduction and orthogonal Krylov subspace bases*, SIAM J. Numer. Anal., 42 (2005), pp. 2148–2158.
[37] G. Meurant, *The Lanczos and Conjugate Gradient Algorithms: From Theory to Finite Precision Computations*, SIAM, Philadelphia, PA, 2006.
[38] G. Meurant and Z. Strakoš, *The Lanczos and conjugate gradient algorithms in finte precision arithmetic*, Acta Numer., 15 (2006), pp. 471–542.
[39] H. Rutishauser, *Bestimmung der Eigenwerte orthogonaler Matrizen*, Numer. Math., 9 (1966), pp. 104–108.
[40] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, PA, second ed., 2003.
[41] T. Sheil-Small, *Complex Polynomials*, vol. 75 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 2002.
[42] G. Szegö, *Orthogonal Polynomials*, American Mathematical Society, New York, 1939. American Mathematical Society Colloquium Publications, v. 23.
[43] V. V. Voevodin, *On methods of conjugate directions*, U.S.S.R. Comput. Maths. Phys., 19 (1979), pp. 228–233.
[44] ———, *The question of non-self-adjoint extension of the conjugate gradients method is closed*, U.S.S.R. Comput. Maths. Phys., 23 (1983), pp. 143–144.
[45] V. V. Voevodin and E. E. Tyrtyshnikov, *On generalization of conjugate direction methods*, in Numerical Methods of Algebra (Chislennye Metody Algebry), Moscow State University Press, Moscow, 1981, pp. 3–9. (English translation provided by E. E. Tyrtyshnikov).
[46] D. S. Watkins, *Some perspectives on the eigenvalue problem*, SIAM Rev., 35 (1993), pp. 430–471.