GLASNIK MATEMATIČKI
Vol. 35(55)(2000), 137 – 148

STABILITY OF POLAR DECOMPOSITIONS

Cornelis V. M. van der Mee[1], André C. M. Ran and Leiba Rodman[2]

Università di Cagliari, Italy, Vrije Universiteit, The Netherlands, The College of William and Mary, USA

Dedicated to the memory of Branko Najman

Abstract. Certain continuity properties of the factors in generalized polar decompositions of real and complex matrices are studied. A complete characterization is given of those generalized polar decompositions that persist under small perturbations in the matrix and in the scalar product. Connections are made with quadratic matrix equations, and with stability properties of certain invariant subspaces.

1. Introduction

Let F be the field of real numbers $\mathbf{R}$ or the field of complex numbers $\mathbf{C}$. Choose a fixed real symmetric (if $F = \mathbf{R}$) or complex Hermitian (if $F = \mathbf{C}$) positive definite $n \times n$ matrix H. Consider the scalar product induced by H by the formula $[x, y] = \langle Hx, y\rangle$, $x, y \in F^n$. Here $\langle \cdot\, , \cdot \rangle$ stands for the usual scalar product in F^n, i.e., $\langle x, y\rangle = \sum_{j=1}^{n} x_j \overline{y_j}$, where x and y are column vectors with components $x_1, \cdots, x_n$ and $y_1, \cdots, y_n$, respectively, and $\overline{y_j} = y_j$ if $F = \mathbf{R}$.

Well-known concepts related to scalar products are defined in an obvious way. Thus, given an $n \times n$ matrix A over F, the H-adjoint A^H is defined by $[Ax, y] = [x, A^H y]$ for all $x, y \in F^n$. In that case $A^H = H^{-1}A^*H$, where A^* denotes the conjugate transpose of A (with $A^* = A^T$, the transpose of A, if $F = \mathbf{R}$). An $n \times n$ matrix A is called *H-selfadjoint* if $A^H = A$. An $n \times n$ matrix

2000 *Mathematics Subject Classification.* 15A23.
Key words and phrases. Polar decompositions, perturbations.
[1] The work of this author was partially supported by INDAM-GNFM and MURST.
[2] The work of this author partially supported by the NSF grant DMS 9800704.
The work of all authors is partially supported by the NATO grant CGR 960700.

U is called H-*unitary* if $[Ux, Uy] = [x, y]$ for all $x, y \in F^n$ (or equivalently, if $U^* HU = H$).

An H-*polar decomposition* of a matrix $X \in F^{n\times n}$ is, by definition, a factorization of the form

$$X = UA, \tag{1.1}$$

where U is H-unitary and A is H-selfadjoint. This definition is more general than the standard definition in that we allow A to be H-selfadjoint (not just H-positive semidefinite) and the scalar product need not be the standard one.

In this paper we study the behaviour of the factors in an H-polar decomposition of a matrix X under perturbations of X and H.

2. Selfadjoint Square Roots and Polar Decompositions

We review some basic results concerning selfadjoint square roots, polar decompositions, and their stability. The results in this section are known, or can be easily obtained from known results.

The following statement is standard when $H = I$, and can be easily reduced to this case. Indeed, A is H-selfadjoint if and only if $H^{\frac{1}{2}} A H^{-\frac{1}{2}}$ is I-selfadjoint; here $H^{\frac{1}{2}}$ is the positive definite square root of H.

Proposition 2.1. *The following statements are equivalent for an H-selfadjoint matrix A:*

(i) *A has the form $A = B^H B$ for some B.*
(ii) *All eigenvalues of A are nonnegative.*
(iii) *There exists an H-selfadjoint square root of A, i.e., a matrix B such that $B = B^H$ and $B^2 = A$.*

If any (and hence all) of the statements (i)-(iii) *holds true, then there is a unique H-selfadjoint square root of A with nonnegative spectrum, denoted $\sqrt[H]{A}$.*

If the statements (i)-(iii) of Proposition 2.1 hold true and A is nonsingular, then $\sqrt[H]{A}$ is given by a functional calculus formula:

$$\sqrt[H]{A} = \frac{1}{2\pi i} \int_\Gamma z^{\frac{1}{2}} (zI - A)^{-1} \, dz$$

where Γ is a simple rectifiable contour in the open right halfplane that contains the eigenvalues of A in its interior and $z^{1/2}$ is the branch of the square root satisfying $1^{1/2} = 1$. We indicate another formula, perhaps less known, for $\sqrt[H]{A}$ that applies also for singular A:

$$\sqrt[H]{A} = \frac{1}{\pi} \int_0^\infty \lambda^{-\frac{1}{2}} (\lambda I + A)^{-1} A \, d\lambda. \tag{2.1}$$

Note that the convergence of the integral is guaranteed (under the statements (i)-(iii) of Proposition 2.1. To verify (2.1), first observe that the right-hand side of (2.1) is clearly H-selfadjoint with only nonnegative eigenvalues. To prove that its square is equal to A, we may assume, using a similarity transformation, that A is diagonal. Then the proof reduces to the scalar case, where the elementary integral

$$\frac{1}{\pi}\int_0^\infty \lambda^{-\frac{1}{2}}(\lambda+q)^{-1}q\,d\lambda = \sqrt{q}, \qquad q \geq 0$$

completes the proof.

Proposition 2.2. *The square root* $\sqrt[H]{A}$ *(having a nonnegative spectrum) of an* H*-selfadjoint matrix* A *with nonnegative spectrum is a continuous function of* A*. More precisely: Given such a matrix* A*, for every* $\varepsilon > 0$ *there is* $\delta > 0$ *(depending also on* A *and* H*) such that for every* G*-selfadjoint matrix* B *having nonnegative spectrum and satisfying* $\|H-G\| + \|B-A\| < \delta$ *the inequality* $\|\sqrt[G]{B} - \sqrt[H]{A}\| < \varepsilon$ *holds. (*G *is assumed to be Hermitian, and necessarily positive definite if* δ *is chosen sufficiently small.)*

If $H = I$ and G is taken to be equal to H, the result of Proposition 2.2 is a particular case of a general theorem concerning continuity of functions of normal matrices (see Theorem 6.2.37 in [HJ]). The theorem asserts that given a set D in the complex plane and a continuous function $f(\lambda)$ on D, the matrix function $f(X)$ is continuous on the set of normal matrices X with all eigenvalues in D. Taking $D = [0,\infty)$ and $f(\lambda) = \sqrt{\lambda}$ gives Proposition 2.2 when $H = G = I$. Again, Proposition 2.2 can be obtained from this particular case without difficulty. Actually, more can be said: the square root $\sqrt[H]{A}$ is a *locally Lipschitz* function of the pair (A,H) if A is nonsingular, and is a *locally Hölder* function with exponent $\frac{1}{2}$ if A is singular. Namely, let A be as in Proposition 2.2, and let $\alpha = 1$ if A is nonsingular, $\alpha = \frac{1}{2}$ if A is singular. Then there exist positive constants δ and K (depending on A and H) such that for every G-selfadjoint matrix B having nonnegative spectrum and satisfying $\|H-G\| + \|B-A\| < \delta$ the inequality

$$\|\sqrt[G]{B} - \sqrt[H]{A}\| \leq K\left(\|H-G\| + \|B-A\|\right)^\alpha \tag{2.2}$$

holds. The local Lipschitz property of the square root is evident, because, by the functional calculus, the square root is an analytic function of a matrix A with positive spectrum, whereas the local Hölder property follows from the proof of Theorem 6.2.37 in [HJ].

As a byproduct of the above considerations, we obtain:

Proposition 2.3. *The square roots* $\sqrt[H]{X^H X}$ *and* $\sqrt[H]{X X^H}$ *are continuous functions of the pair* (X,H)*, where* $H \in F^{n\times n}$ *is positive definite and* $X \in F^{n\times n}$*. Moreover,* $\sqrt[H]{X^H X}$ *and* $\sqrt[H]{X X^H}$ *are locally Lipschitz continuous on* *the set of pairs (X, H) as above, where X is nonsingular, and are locally Hölder continuous with exponent $\frac{1}{2}$ on the set of pairs (X, H) as above with singular X.*

Passing to polar decompositions, it is a standard result that an H-polar decomposition (1.1) always exists. Moreover, in this case one can take A having nonnegative spectrum; an H-polar decomposition of the form (1.1) with this property of A will be called a *nonnegative* H-polar decomposition. By Proposition 2.1 the factor A in a nonnegative H-polar decomposition is unique and coincides with $\sqrt[H]{X^H X}$. The factor U in a nonnegative H-polar decomposition is unique if and only if X is nonsingular. By Proposition 2.3 we have:

PROPOSITION 2.4. *The factors U and A of the nonnegative H-polar decomposition (1.1) are locally Lipschitz continuous functions on the set of pairs (X, H), where $H \in F^{n\times n}$ is positive definite and $X \in F^{n\times n}$ is nonsingular.*

In connection with Proposition 2.4 we mention the following formula that gives perturbation bounds on the unitary factor of the nonnegative polar decomposition for invertible matrices, assuming $H = I$ (see [Li]; also [Bh], Theorem VII.5.1). Let X and Y be invertible matrices with the nonnegative polar decompositions $X = UA$ and $Y = VB$. Then

$$\|U - V\| \le \frac{2}{\|X^{-1}\|^{-1} + \|Y^{-1}\|^{-1}} \|X - Y\|. \tag{2.3}$$

As shown in [Li], the formula (2.3) is the best possible in the sense that the bound can be achieved. Using the identity $A - B = V^{-1}(X - Y) - U^{-1}(U - V)V^{-1}X$ and the unitarity of U and V, we immediately get the corresponding bounds for the positive semidefinite factor:

$$\|A - B\| \le \left(1 + \frac{2\|X\|}{\|X^{-1}\|^{-1} + \|Y^{-1}\|^{-1}}\right) \|X - Y\|. \tag{2.4}$$

3. STABILITY

In this section we state and prove the main result of this paper concerning stability of polar decompositions (1.1), which are not necessarily nonnegative.

The polar decomposition (1.1) is called *stable* if for every $\varepsilon > 0$ there is $\delta > 0$ such that every pair of matrices (Y, G), where $Y \in F^{n\times n}$ and $G \in F^{n\times n}$ is Hermitian, admits a G-polar decomposition $Y = VB$ with $\|U - V\| + \|A - B\| < \varepsilon$, as soon as $\|Y - X\| + \|H - G\| < \delta$. Restricting this definition to perturbations of Y only, in other words, assuming $G = H$, we obtain the definition of *H-stability* of the polar decomposition (1.1). The polar decomposition (1.1) is called *Lipschitz stable* (resp. *H-Lipschitz stable*) if there exist positive constants δ and K such that every $Y \in F^{n\times n}$ admits a G- (resp. H-) polar decomposition $Y = VB$ with $\|U - V\| + \|A - B\| \le$

$K(\|X-Y\|+\|G-H\|)$ (resp. $\|U-V\|+\|A-B\| \le K\|X-Y\|$) as soon as $\|X-Y\|+\|G-H\| \le \delta$ (resp. $\|X-Y\| \le \delta$). Clearly, stability implies H-stability, Lipschitz stability implies stability and H-Lipschitz stability, and H-Lipschitz stability implies H-stability. The following result shows, in particular, that these a priori distinct notions of stability are in fact equivalent.

THEOREM 3.1. (a) $F = \mathbf{C}$. *There exist H-stable H-polar decompositions of X if and only if X is nonsingular. In this case, the following statements are equivalent for an H-polar decomposition (1.1):*

(i) *(1.1) is H-stable.*
(ii) *(1.1) is Lipschitz stable.*
(iii) *The H-selfadjoint matrices A and $-A$ have no common eigenvalues.*

(b) $F = \mathbf{R}$. *There exist H-stable H-polar decompositions of X if and only if $\dim \operatorname{Ker} X \le 1$. In this case, the following statements are equivalent for an H-polar decomposition (1.1):*

(iv) *(1.1) is H-stable.*
(v) *(1.1) is Lipschitz stable.*
(vi) *The H-selfadjoint matrices A and $-A$ have no common nonzero eigenvalues.*

In particular, the nonnegative H-polar decomposition is Lipschitz stable if and only if X is nonsingular (in the complex case) or $\dim \operatorname{Ker} X \le 1$ (in the real case).

In contrast with (2.3) and (2.4), Theorem 3.1 provides qualitative results. On the other hand, the quantitative formulas (2.3) and (2.4) apply to a more restrictive class of polar decompositions, namely, the nonnegative polar decompositions of nonsingular matrices.

PROOF. In the first part of the proof we reduce the problem in three steps to the case where $H = I$ and X is selfadjoint, and to the polar decomposition with $U = I$, and we show that we have to consider perturbations of X only.

First of all we note that without loss of generality we may restrict ourselves to the case $H = I$. Indeed, write $H = S_1^* S_1$ for some invertible S_1, and observe that $X = UA$ is an H-polar decomposition if and only if $S_1 X S_1^{-1} = S_1 U S_1^{-1} \cdot S_1 A S_1^{-1}$ is an I-polar decomposition.

Secondly, observe that for any $\alpha > 0$ the I-polar decomposition $X = UA$ of X is stable if and only if the I-polar decomposition $\alpha X = U(\alpha A)$ of αX is stable. Thus we may assume in the sequel that $\|X\| = 1$. Now we show that we can restrict our attention to perturbations of X only. Let $X = UA$ be an I-polar decomposition of X. Consider a perturbation Y of X and a perturbation $G = G^*$ of $H = I$. Then we can write $G = S^2$, with positive definite S close to I. In fact, by (2.2) (with $H = G = A = I$, $B = S^2$) we have $\|S - I\| \leq C_1 \|G - I\|$ for some constant $C_1 > 0$, uniformly in G. Moreover, we may assume that $\|Y\| \leq 2$ as $\|X\| = 1$ and we consider sufficiently small perturbations. Put $Y_1 = SYS^{-1}$. Then

$$\begin{aligned} \|Y_1 - X\| &\leq \|Y_1 - Y\| + \|Y - X\| \\ &\leq \|S^{-1}\| \cdot \|(S - I)Y - Y(S - I)\| + \|Y - X\| \\ &\leq C_2 \|G - I\| + \|Y - X\|, \end{aligned}$$

where C_2 is a positive constant, uniformly in G. Moreover, if $Y = VB$ is a G-polar decomposition of Y then with $V_1 = SVS^{-1}$ and $B_1 = S^{-1}(GB)S^{-1} = SBS^{-1}$ we have that $Y_1 = V_1 B_1$ is an I-polar decomposition of Y_1, and vice versa. Indeed, V being G-unitary implies that V_1 is unitary and B being G-selfadjoint implies that GB and hence B_1 are selfadjoint. Therefore, we can restrict our attention to perturbations of X only, keeping $H = I$ fixed. Henceforth, in this proof we shall drop the I in I-stable and in I-polar decomposition. This also shows that the notions of stability and H-stability coincide, as well as the notions of Lipschitz stability and Lipschitz H-stability.

Thirdly, we show that we can restrict the attention to the case where X is selfadjoint and the unitary factor in the polar decomposition is the identity matrix. Indeed, suppose $X = UA$ is a stable polar decomposition. Then for $\tilde{X} = U^* X$ the polar decomposition $\tilde{X} = I \cdot A$ is also stable. Indeed, let $\|\tilde{Y} - \tilde{X}\| < \delta$, then for $Y = U\tilde{Y}$ we have $\|Y - X\| = \|\tilde{Y} - \tilde{X}\| < \delta$, so there exist a unitary V and a selfadjoint B with $Y = VB$ and $\|U - V\| + \|A - B\| < \varepsilon$. Then $\tilde{Y} = (U^* V)B$ is a polar decomposition and

$$\|U^* U - U^* V\| + \|A - B\| = \|U - V\| + \|A - B\| < \varepsilon.$$

On the other hand, if $X = UA$ is a polar decomposition which is not stable, then there exists a sequence $X_n \to X$ such that for any polar decomposition $X_n = U_n A_n$ either A_n does not converge to A or U_n does not converge to U. Consider again $\tilde{X} = U^* X$, and $\tilde{X}_n = U^* X_n$. Then for any polar decomposition of $\tilde{X}_n$ either the unitary factor does not converge to I, or the selfadjoint factor does not converge to A. Hence the polar decomposition $\tilde{X} = I \cdot A$ is not stable. Analogously one proves that (1.1) is Lipschitz stable if and only if $\tilde{X} = I \cdot A$ is Lipschitz stable. We conclude that for the remainder of the proof we may assume that X is selfadjoint and that the unitary factor in the polar decomposition under consideration is the identity matrix.

Now suppose that X is singular and selfadjoint. Decompose the underlying space F^n as $\operatorname{Ker} X \oplus \operatorname{Im} X$. Put $\dim \operatorname{Ker} X = k$. Suppose first that $k > 1$ or $F = \mathbf{C}$. With respect to this decomposition we can write X as $\begin{pmatrix} 0 & 0 \\ 0 & X_{22} \end{pmatrix}$ with X_{22} invertible and selfadjoint. Let X_{11} be an arbitrary invertible $k \times k$ matrix such that in any polar decomposition $X_{11} = V_{11} A_{11}$ the unitary matrix

V_{11} is bounded away from I_k. This is possible as $k > 1$ or $F = \mathbf{C}$. Put $Y = \begin{pmatrix} \varepsilon X_{11} & 0 \\ 0 & X_{22} \end{pmatrix}$, where $\varepsilon \in \mathbf{R}$ is close to zero. As Y is invertible any polar decomposition that needs to be close to the polar decomposition $X = I \cdot X$ must be of the form $Y = \begin{pmatrix} V_{11} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \varepsilon A_{11} & 0 \\ 0 & X_{22} \end{pmatrix}$, with A_{11} taken from a polar decomposition $X_{11} = V_{11} A_{11}$. However, as $V_{11} \neq I_k$, we cannot have that $V_{11} - I_k \to 0$ as ε goes to zero. So no polar decomposition is stable in this case.

Next, suppose that $F = \mathbf{R}$, X is singular and selfadjoint with one-dimensional kernel, and that X and $-X$ don't have a common nonzero eigenvalue. Decompose the space $\mathbf{R}^n$ as in the previous paragraph, and write X as above. By considering the spectral decomposition of X we may as well assume that X is diagonal: $X = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, where $\lambda_1 = 0$, and $\lambda_2, \ldots, \lambda_n$ are not necessarily distinct non-zero numbers such that $\lambda_i^2 = \lambda_j^2 \Rightarrow \lambda_i = \lambda_j$, $(i, j = 2, \ldots, n)$. We also assume that the λ_j are arranged in the non-decreasing order of their absolute values. We continue to assume that the unitary factor in the polar decomposition of X is the identity matrix. Let Y be an arbitrary small perturbation of X. Let B_1 be the positive semidefinite square root of Y^*Y, and write $B_1 = V_1 \operatorname{diag}(\theta_1, \theta_2, \ldots, \theta_n) V_1^*$, with $|\theta_1|$, $|\theta_i - |\lambda_i||$, $(i > 1)$ small and V_1 unitary. Then V_1 can be taken to be close to the identity matrix and actually, we can arrange it so that $\|V_1 - I\|$ is of the same order of magnitude as $\|X - Y\|$ (cf. [Bh], Theorem VII.3.2). Also, $|\theta_i - |\lambda_i||$, $(i > 1)$ are of the same order of magnitude as $\|X - Y\|$; an easy way to verify this, assuming for simplicity that all the λ_j are distinct, is to notice that θ_i (resp. $|\lambda_i|$) is the distance from Y (resp. X) to the set of all real matrices having rank at most $n - i$, the distance being measured in the operator norm. Let $\sigma_1 = \pm\theta_1$, and $\sigma_i = \theta_i \operatorname{sign} \lambda_i$ for $i > 1$, where the sign $\pm$ is chosen so that

$$\sigma_1 \cdot \ldots \cdot \sigma_n \cdot \det Y \geq 0, \tag{3.1}$$

and put $B = V_1 \operatorname{diag}(\sigma_1, \ldots, \sigma_n) V_1^*$. Then B is close to X, and as a matter of fact there is a constant $C_3 > 0$ depending on X only (provided $\|X - Y\|$ is small enough) such that $\|X - B\| \leq C_3 \|X - Y\|$. Note that B is a selfadjoint square root of Y^*Y. Now we determine a matrix $U \in \mathbf{R}^{n \times n}$ by using the equality $Y = UB$. This definition is correct, because $Bx = 0$ implies $Yx = 0$. Moreover, a standard argument shows that

$$\langle Uy_1, Uy_2 \rangle = \langle y_1, y_2 \rangle \quad \text{for every } y_1, y_2 \in \operatorname{Range} B. \tag{3.2}$$

Thus, if B is invertible, then U is uniquely defined by the equality $Y = UB$, and U is necessarily unitary; in addition, (3.1) guarantees that $\det U = 1$. If B is not invertible, then (for Y sufficiently close to X) $\dim \operatorname{Ker} B = 1$, the equality $Y = UB$ determines U uniquely on the range of B, and (3.2) holds. It is easy to see that U can be extended to a unitary matrix, by requiring (in addition to $Y = UB$) that $Uf = \pm g$, where f (resp., g) is a normalized vector in the orthogonal complement to $\operatorname{Range} B$ (resp., $\operatorname{Range} Y$). We adjust the sign $\pm$ so that $\det U = 1$. To summarize: There exists a unitary matrix $U \in \mathbf{R}^{n\times n}$ with $\det U = 1$ such that $Y = UB$.

We now show, under the hypotheses made in the preceding paragraph, that the matrix U is actually close to I. Denote by $e_1, \dots, e_n$ the standard unit coordinate vectors in $\mathbf{R}^n$. We have:

$$YV_1 e_j = UBV_1 e_j = UV_1 \operatorname{diag}(\sigma_1, \dots, \sigma_n) e_j = \sigma_j U V_1 e_j, \tag{3.3}$$

and therefore for $j \geq 2$:

$$\begin{aligned}
\|Ue_j - e_j\| &= \theta_j^{-1} \|\sigma_j U e_j - \sigma_j e_j\| \\
&\leq \theta_j^{-1} \{\|\sigma_j U e_j - \lambda_j e_j\| + |\lambda_j - \sigma_j|\} \\
&\leq \theta_j^{-1} \{\|\sigma_j (U e_j - U V_1 e_j)\| + \|\sigma_j U V_1 e_j - \lambda_j e_j\| + |\lambda_j - \sigma_j|\}
\end{aligned}$$

(by (3.3))

$$\begin{aligned}
&= \theta_j^{-1} \{|\sigma_j|\, \|(V_1 - I)e_j\| + \|YV_1 e_j - Xe_j\| + |\lambda_j - \sigma_j|\} \\
&\leq \theta_j^{-1} \{|\sigma_j|\, \|V_1 - I\| + \|YV_1 e_j - Ye_j\| + \|Ye_j - Xe_j\| + |\lambda_j - \sigma_j|\} \\
&\leq \theta_j^{-1} \{|\sigma_j|\, \|V_1 - I\| + \|Y\| \cdot \|V_1 - I\| + \|Y - X\| + |\lambda_j - \sigma_j|\} \\
&= \theta_j^{-1} \{|\sigma_j|\, \|V_1 - I\| + \|Y\| \cdot \|V_1 - I\| + \|Y - X\| + |\theta_j - |\lambda_j||\} \\
&\leq C_4 \|X - Y\|.
\end{aligned}$$

Here, C_4 is a positive constant depending on X only, provided $\|X - Y\|$ is small enough. Thus, denoting by $\hat{U}$ the $n \times (n-1)$ matrix obtained from U by removing its first column, we see that $\left\| \hat{U} - \begin{pmatrix} 0 \\ I_{n-1} \end{pmatrix} \right\|$ is of the same order of magnitude as $\|X - Y\|$. Now, the first column of U is a normalized vector x satisfying $\hat{U}^* x = 0$. The condition $\det U = 1$ guarantees that x is close to e_1, with $\|x - e_1\| \leq C_5 \|X - Y\|$, where, again, C_5 is a positive constant depending on X only, provided $\|X - Y\|$ is small enough.

This proves that X with one-dimensional kernel in the real case has Lipschitz stable polar decompositions; moreover, we have proved that (vi) implies (v). If X is nonsingular and A is positive definite, the Lipschitz stability of (1.1) is immediate from (2.3) and (2.4). We have verified that there exist stable polar decompositions of X if and only if X is nonsingular (in the complex case), and if and only if the dimension of the kernel of X is at most one (in the real case).

Next, we prove (iii) $\Rightarrow$ (ii). To this end, observe that if X is nonsingular and (iii) holds, then in fact $A = f(X^*X)$, where $f(z)$ is an analytic function defined in a neighborhood of the spectrum of X^*X. Now letting $B = f(Y^GY)$ for Y and $G = G^*$ sufficiently close to X and I, respectively, the Lipschitz stability of (1.1) is easily verified.

It remains to show that under the conditions on $\operatorname{Ker} X$ as stated in the theorem the polar decomposition $X = UA$ is not I-stable if A and $-A$ have a common non-zero eigenvalue. Again, we assume $U = I$, i.e., $A = X$.

Suppose that $X = X^*$ has an eigenvalue $\lambda \neq 0$ and $-\lambda$ is an eigenvalue as well. Let x_1 and x_2 be normalized eigenvectors of X with respect to λ and $-\lambda$, respectively. Then x_1 and x_2 are orthogonal and we may as well assume that X is given by $X = \operatorname{diag}(\lambda, -\lambda, X_{22})$. Again we may assume that the unitary matrix in the polar decomposition under consideration is the identity matrix. Now consider perturbations of X of the form $\tilde{X} = \tilde{X}_{11} \oplus X_{22}$, where $\tilde{X}_{11} = \begin{pmatrix} \lambda & \varepsilon \\ -\varepsilon & -\lambda \end{pmatrix}$, with $\varepsilon \neq 0$ such that $(\lambda+\varepsilon)^2$ and $(\lambda-\varepsilon)^2$ are not eigenvalues of X_{22}^2. Then the matrices in any polar decomposition of $\tilde{X}$ are block diagonal with a 2×2 block in the upper left hand corner. We shall focus our attention on this 2×2 block. Compute $\tilde{X}_{11}^*\tilde{X}_{11} = \begin{pmatrix} \lambda^2+\varepsilon^2 & 2\varepsilon\lambda \\ 2\varepsilon\lambda & \lambda^2+\varepsilon^2 \end{pmatrix}$. This matrix has $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ as eigenvectors, corresponding to the eigenvalues $(\lambda+\varepsilon)^2$ and $(\lambda-\varepsilon)^2$, respectively. If we wish to find a polar decomposition $\tilde{X}_{11} = \tilde{U}_{11}\tilde{A}_{11}$ where $\tilde{A}_{11}$ is close to $X_{11} = \operatorname{diag}(\lambda, -\lambda)$, then this turns out to be impossible, because the eigenvectors of $\tilde{A}_{11}$ are $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, independently of ε. By considering $\tilde{U}_{11} = \tilde{X}_{11}\tilde{A}_{11}^{-1}$ it is also easily checked that the unitary factor in this case does not converge to the identity either, as $\varepsilon \to 0$. So the polar decomposition $X = I\cdot X$ is not stable. As a matter of fact, neither the unitary nor the selfadjoint factor is stable. ∎

In [R], Chapter 4, some parts of Theorem 3.1 were proved in a totally different way. To be precise, it was shown there that, for $H = I$ and $F = \mathbf{C}$, there exist stable polar decompositions of X if and only if X is invertible, and in that case (i) and (iii) are equivalent.

The perturbation part of Theorem 3.1 can also be presented in a stronger, function theoretic form:

THEOREM 3.2. *Assume that the H-polar decomposition (1.1) is not stable. Then there exists $Y \in F^{n\times n}$ such that for every $\varepsilon \neq 0$ sufficiently close to zero and every H-polar decomposition $X + \varepsilon Y = U(\varepsilon)A(\varepsilon)$ the unitary matrices $U(\varepsilon)$ do not converge to U, as ε tends to zero. Moreover, Y can be* *chosen so that* $\operatorname{rank} Y = 2$ *if* A *and* $-A$ *have a common non-zero eigenvalue, or* $\operatorname{rank} Y = \dim \operatorname{Ker} X$ *if* X *is singular.*

Conversely, let $\mathcal{D}$ *be the open set of invertible* $n \times n$ *matrices if* $F = \mathbf{C}$*, or the open set of* $n \times n$ *matrices having rank at least* $n-1$ *if* $F = \mathbf{R}$*, and denote by* $\mathcal{P}$ *the set of positive definite Hermitian matrices in* $F^{n\times n}$*. Then the* H*-polar decompositions* $X = UA$*, where* A *and* $-A$ *have no common non-zero eigenvalues, are real analytic functions of the pair* $(X, H) \in \mathcal{D} \times \mathcal{P}$*. More precisely, let* $X_0 = U_0 A_0$ *be an* H_0*-polar decomposition, where* $(X_0, H_0) \in \mathcal{D} \times \mathcal{P}$*, and* A_0 *and* $-A_0$ *have no common non-zero eigenvalues. Then for every* $(X, H) \in \mathcal{D} \times \mathcal{P}$ *in a neighborhood of* (X_0, H_0) *there exists an* H*-polar decomposition* $X = UA$ *in which* $U = U(X, H)$ *and* $A = A(X, H)$ *are real analytic functions of the real and imaginary parts of the entries of* X *and of the diagonal entries and the real and imaginary parts of the strictly upper triangular entries of* H*, and* $U(X_0, H_0) = U_0$*,* $A(X_0, H_0) = A_0$*. (If* $F = \mathbf{R}$*, the imaginary parts are absent, of course.)*

The proof of Theorem 3.2 is obtained by inspection of the proof of Theorem 3.1.

4. Quadratic Matrix Equations and Invariant Subspaces

The H-polar decompositions (1.1) are closely related to certain matrix quadratic equations. This relationship was first observed and exploited in [R]. Here we use this connection to ascertain stability properties of unitary solutions of such equations.

Indeed, let $X = UA$ be an H-polar decomposition of $X \in F^{n\times n}$. Then, as one readily verifies,

$$UX^H U - X = 0. \tag{4.1}$$

So U is an H- unitary solution of the matrix quadratic equation (4.1). Conversely, if U is an H- unitary solution of this equation, and we put $A = U^{-1}X$, then $A^H = X^H U = U^{-1}X = A$, so $X = UA$ is an H-polar decomposition of X. We say that a unitary solution U of (4.1) is H-*stable* if for every $\varepsilon > 0$ there is $\delta > 0$ such that every equation $VY^H V - Y = 0$ with $Y \in F^{n\times n}$ admits an H-unitary solution V with $\|U - V\| < \varepsilon$, as soon as $\|Y - X\| < \delta$. An H-unitary solution U of (4.1) is called *Lipschitz stable* if there exist positive constants δ and K such that every equation $VY^G V - Y = 0$ admits a G-unitary solution V with $\|U - V\| + \|A - B\| \le K\,(\|X - Y\| + \|G - H\|)$ as soon as $\|X - Y\| + \|G - H\| \le \delta$ and G is Hermitian (necessarily positive definite if δ is sufficiently small). We have the following corollary:

Corollary 4.1. *The equation* (4.1) *has* H*-stable* H*-unitary solutions if and only if* $\dim \operatorname{Ker} X = 0$ *(in case* $F = \mathbf{C}$*) or* $\dim \operatorname{Ker} X \le 1$ *(in case* $F = \mathbf{R}$*). In this case an* H*-unitary solution* U *of* (4.1) *is* H*-stable if and* *only if it is Lipschitz stable if and only if* $U^{-1}X$ *and* $-U^{-1}X$ *have no common non-zero eigenvalues.*

Proof. By the remarks before the statement of the corollary, stability of H-unitary factors in H-polar decompositions translates directly into stability of H-unitary solutions of (4.1) under perturbations of X only. As we have seen in the proof of Theorem 3.1, there exists a H-unitary factor that is stable (in any of the senses introduced in Section 3) if and only if $\dim \operatorname{Ker} X = 0$ (in case $F = \mathbf{C}$) or $\dim \operatorname{Ker} X \leq 1$ (in case $F = \mathbf{R}$). Now apply Theorem 3.1. $\square$

Corollary 4.1 can also be formulated in the function theoretic form, analogously to Theorem 3.2. We leave this formulation to the interested reader.

The connections between matrix quadratic equations and certain invariant subspaces are well-known (see, for example, [BGK], Chapter 17 of [GLR], and Chapters 7 and 8 of [LR]). Applying these ideas to the equation (4.1), we consider

$$\hat{X} = \begin{pmatrix} 0 & X^H \\ X & 0 \end{pmatrix} \in F^{2n \times 2n}, \qquad \hat{H} = \begin{pmatrix} H & 0 \\ 0 & -H \end{pmatrix} \in F^{2n \times 2n},$$

where $X, H \in F^{n \times n}$ and H is positive definite Hermitian. It is easy to see that a subspace $\mathcal{M} \subset F^{2n}$ is *hypermaximal* $\hat{H}$*-neutral*, i. e., $\dim \mathcal{M} = n$ and $\langle \hat{H}x, y \rangle = 0$ for all $x, y \in \mathcal{M}$, if and only if it is a graph subspace $\mathcal{M} = \operatorname{Im} \begin{pmatrix} I \\ U \end{pmatrix}$ with H-unitary U. Furthermore, such subspace $\mathcal{M}$ is $\hat{X}$-invariant if and only if U satisfies the equation (4.1). We can therefore apply Theorems 3.1 and 3.2 to stability problems of $\hat{X}$-invariant hypermaximal $\hat{H}$-neutral subspaces. For example:

Corollary 4.2. *Let* $\mathcal{D}$ *and* $\mathcal{P}$ *be defined as in Theorem* 3.2*. Assume that* $(X_0, H_0) \in \mathcal{D} \times \mathcal{P}$*, and let* $\mathcal{M}_0$ *be an* $\hat{X}_0$*-invariant hypermaximal* $\hat{H}_0$*-neutral subspace such that the restriction* $\hat{X}_0|\mathcal{M}_0$ *has no pairs of non-zero eigenvalues* $\pm\lambda$*. Then for every* $(X, H) \in \mathcal{D} \times \mathcal{P}$ *in a neighborhood of* (X_0, H_0) *there exists an* $\hat{X}$*-invariant hypermaximal* $\hat{H}$*-neutral subspace* $\mathcal{M} = \mathcal{M}(X, H)$ *which is a real analytic function of* (X, H)*, and* $\mathcal{M}(X_0, H_0) = \mathcal{M}_0$*.*

References

[BGK] H. Bart, I. Gohberg, and M. A. Kaashoek. *Minimal Factorization of Matrix and Operator Functions*, OT1, Birkhäuser, Basel, 1979.

[Bh] R. Bhatia. *Matrix Analysis,* Springer, New York, 1997.

[GLR] I. Gohberg, P. Lancaster, and L. Rodman. *Invariant Subspaces of Matrices with Applications,* John Wiley and Sons, New York, 1986.

[HJ] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[LR] P. Lancaster and L. Rodman. *Algebraic Riccati Equations*, Oxford University Press, Oxford, 1995.

[Li] R.-C. Li. New perturbation bounds for the unitary polar factor, *SIAM J. Matrix Analysis and Appl.*, 16 (1995), 327 - 332.

[R] A. C. M. Ran. *Semidefinite Invariant Subspaces. Stability and Applications*, Ph. D. Thesis, Vrije Universiteit, Amsterdam, 1984.

Dipartimento di Matematica
Università di Cagliari
Via Ospedale 72
09124 Cagliari, Italy
E-mail: `cornelis@krein.unica.it`

Faculteit Wiskunde en Informatica
Vrije Universiteit Amsterdam
De Boelelaan 1081
1081 HV Amsterdam, The Netherlands
E-mail: `ran@cs.vu.nl`

Department of Mathematics
The College of William and Mary
Williamsburg, VA 23187-8795, USA
E-mail: `lxrodm@math.wm.edu`

Received: 03.07.98
Revised: 10.12.98