

# STRUCTURED CONDITION NUMBERS AND BACKWARD ERRORS IN SCALAR PRODUCT SPACES\*

FRANÇOISE TISSEUR<sup>†</sup> AND STEF GRAILLAT<sup>‡</sup>

**Abstract.** We investigate the effect of structure-preserving perturbations on the solution to a linear system, matrix inversion, and distance to singularity. Particular attention is paid to linear and nonlinear structures that form Lie algebras, Jordan algebras and automorphism groups of a scalar product. These include complex symmetric, pseudo-symmetric, persymmetric, skew-symmetric, Hamiltonian, unitary, complex orthogonal and symplectic matrices. We show that under mild assumptions on the scalar product, there is little or no difference between structured and unstructured condition numbers and distance to singularity for matrices in Lie and Jordan algebras. Hence, for these classes of matrices, the usual unstructured perturbation analysis is sufficient. We show this is not true in general for structures in automorphism groups. Bounds and computable expressions for the structured condition numbers for a linear system and matrix inversion are derived for these nonlinear structures.

Structured backward errors for the approximate solution of linear systems are also considered. Conditions are given for the structured backward error to be finite. We prove that for Lie and Jordan algebras, whenever the structured backward error is finite, it is within a small factor of the unstructured one. The same conclusion holds for orthogonal and unitary structures but cannot easily be extended to other matrix groups. For some nonlinear structures such as complex orthogonal, we show that the structured backward error can be rewritten in terms of a challenging Procrustes problem.

This work extends and unifies earlier analyses.

**Key words.** structured matrices, normwise structured perturbations, structured linear systems, condition number, backward error, distance to singularity, Lie algebra, Jordan algebra, automorphism group, scalar product, bilinear form, sesquilinear form, orthosymmetric.

**AMS subject classifications.** 15A12, 65F35, 65F15.

**1. Motivations.** Condition numbers and backward errors play an important role in numerical linear algebra. Condition numbers measure the sensitivity of the solution of a problem to perturbation in the data whereas backward errors reveal the stability of a numerical method. Also, when combined with a backward error estimate, condition numbers provide an approximate upper bound on the error in a computed solution.

There is a growing interest in structured perturbation analysis (see for example [1], [7], [16] and the literature cited therein) due to the substantial development of algorithms for structured problems. For these problems it is often more appropriate to define condition numbers that measure the sensitivity to structured perturbations.

Structured perturbations are not always as easy to handle as unstructured ones and the backward error analysis of an algorithm is generally more difficult when

---

\*Numerical Analysis Report 473, Manchester Centre for Computational Mathematics, September 2005.

<sup>†</sup>School of Mathematics, The University of Manchester, Sackville Street, Manchester, M60 1QD, UK ([ftisseur@ma.man.ac.uk](mailto:ftisseur@ma.man.ac.uk), <http://www.ma.man.ac.uk/~ftisseur/>). This work was supported by Engineering and Physical Sciences Research Council grant GR/S31693.

<sup>‡</sup>Laboratoire LP2A, Université de Perpignan Via Domitia, 52, avenue Paul Alduy, F-66860 Perpignan Cedex, France ([graillat@univ-perp.fr](mailto:graillat@univ-perp.fr), <http://gala.univ-perp.fr/~graillat/>).

structured perturbations are concerned. Our aim in this paper is to identify classes of matrices for which the ratio between the structured and the unstructured condition number of a problem and the ratio between the structured and unstructured backward error are close to one. For example, it is well-known that for symmetric linear systems and for normwise distances it makes no difference to the backward error or condition number whether matrix perturbations are restricted to be symmetric or not [1], [6], [17]. Rump [16] recently showed that the same holds for skew-symmetric and persymmetric structures. We unify and extend these results to a large class of structured matrices.

The structured matrices we consider belong to Lie and Jordan algebras or group automorphisms associated with unitary and orthosymmetric scalar products. These include for example symmetric, complex symmetric and complex skew-symmetric matrices, pseudo-symmetric, persymmetric and perskew-symmetric matrices, Hamiltonian and skew-Hamiltonian matrices, Hermitian, skew-Hermitian and  $J$ -Hermitian matrices, but also nonlinear structures such as orthogonal, complex orthogonal, symplectic, unitary and conjugate symplectic structures.

The paper is organized as follows. In Section 2, we set up the notation and recall some necessary background material. In Section 3, we consider the condition number for the matrix inverse and the condition number for linear systems as well as the distance to singularity. We show that for matrices belonging to Lie or Jordan algebras associated with a unitary and orthosymmetric scalar product and for perturbations measured in the Frobenius norm or 2-norm, the structured and unstructured condition numbers are equal or within a small factor of each other and that the distance to singularity is unchanged when the perturbations are forced to be structured. Hence, for these classes of matrices the usual unstructured perturbation analysis is sufficient. We show that this is false in general for nonlinear structures belonging to automorphism groups. We give bounds and derive computable expressions for the matrix inversion and linear system structured condition numbers. Structured backward errors for linear systems are considered in Section 4. Unlike in the unstructured case, the structured backward error may not exist and we say in this case that it is infinite. We give conditions on  $b$  and the approximate solution  $\hat{x}$  to  $Ax = b$  for the structured backward error to be finite. We show that for Lie and Jordan algebras of orthosymmetric and unitary scalar products when the structured backward error is finite it is within a small factor of the unstructured one. Unfortunately, there is no general technique to derive explicit expressions for the structured backward error for structures in automorphism groups. For orthogonal and unitary matrices, we show that the Frobenius norm structured backward error when it exists differs from the unstructured one by at most a factor  $\sqrt{2}$ . For some other particular matrix groups such as complex orthogonal matrices, the structured backward error can be rewritten in terms of Procrustes problems for which an explicit expression is not available.

**2. Preliminaries.** We begin with the basic definitions and properties of scalar products and the structured classes of matrices associated with them. Then we give auxiliary results for these structured matrices.

TABLE 2.1

A sampling of structured matrices associated with scalar products  $\langle \cdot, \cdot \rangle_M$ , where  $M$  is the matrix defining the scalar product.

Space	$M$	Automorphism Group $\mathbb{G} = \{G : G^\star = G^{-1}\}$	Jordan Algebra $\mathbb{J} = \{S : S^\star = S\}$	Lie Algebra $\mathbb{L} = \{K : K^\star = -K\}$
<b>Symmetric bilinear forms</b>				
$\mathbb{R}^n$	$I$	Real orthogonals	Symmetrics	Skew-symmetrics
$\mathbb{C}^n$	$I$	Complex orthogonals	Complex symmetrics	Cplx skew-symmetrics
$\mathbb{R}^n$	$\Sigma_{p,q}$	Pseudo-orthogonals	Pseudo symmetrics	Pseudo skew-symmetrics
$\mathbb{C}^n$	$\Sigma_{p,q}$	Cplx pseudo-orthogonals	Cplx pseudo-symm.	Cplx pseudo-skew-symm.
$\mathbb{R}^n$	$R$	Real perplectics	Persymmetrics	Perskew-symmetrics
<b>Skew-symmetric bilinear forms</b>				
$\mathbb{R}^{2n}$	$J$	Real symplectics	Skew-Hamiltonians	Hamiltonians
$\mathbb{C}^{2n}$	$J$	Complex symplectics	Cplx $J$ -skew-symm.	Complex $J$ -symmetrics
<b>Hermitian sesquilinear forms</b>				
$\mathbb{C}^n$	$I$	Unitaries	Hermitian	Skew-Hermitian
$\mathbb{C}^n$	$\Sigma_{p,q}$	Pseudo-unitaries	Pseudo Hermitian	Pseudo skew-Hermitian
<b>Skew-Hermitian sesquilinear forms</b>				
$\mathbb{C}^{2n}$	$J$	Conjugate symplectics	$J$ -skew-Hermitian	$J$ -Hermitian

Here,  $R = \begin{bmatrix} & & & 1 \\ & \cdot & & \\ & & \cdot & \\ 1 & & & \end{bmatrix}$  and  $\Sigma_{p,q} = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix} \in \mathbb{R}^{n \times n}$  are symmetric and  $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$  is skew-symmetric.

**2.1. Structured matrices in scalar product spaces.** The term *scalar product* will be used to refer to any nondegenerate bilinear or sesquilinear form  $\langle \cdot, \cdot \rangle$  on  $\mathbb{K}^n$ . Here  $\mathbb{K}$  denotes the field  $\mathbb{R}$  or  $\mathbb{C}$ . It is well known that any real or complex bilinear form  $\langle \cdot, \cdot \rangle$  has a unique matrix representation given by  $\langle \cdot, \cdot \rangle = x^T M y$ , while a sesquilinear form can be represented by  $\langle \cdot, \cdot \rangle = x^* M y$ , where the matrix  $M$  is nonsingular. We will denote  $\langle \cdot, \cdot \rangle$  by  $\langle \cdot, \cdot \rangle_M$  as needed.

A bilinear form is symmetric if  $\langle x, y \rangle = \langle y, x \rangle$ , and skew-symmetric if  $\langle x, y \rangle = -\langle y, x \rangle$ . Hence for a symmetric form  $M = M^T$  and for a skew-symmetric form  $M = -M^T$ . A sesquilinear form is Hermitian if  $\langle x, y \rangle = \overline{\langle y, x \rangle}$  and skew-Hermitian if  $\langle x, y \rangle = -\overline{\langle y, x \rangle}$ . The matrices associated with such forms are Hermitian and skew-Hermitian, respectively.

To each scalar product there corresponds a notion of *adjoint*, generalizing the idea of transpose  $^T$  and conjugate transpose  $^*$ , that is, for any matrix  $A \in \mathbb{K}^{n \times n}$  there is a unique adjoint  $A^\star$  with respect to the form defined by  $\langle Ax, y \rangle_M = \langle x, A^\star y \rangle_M$  for all  $x$  and  $y$  in  $\mathbb{K}^n$ . In matrix terms the adjoint is given by

$$(2.1) \quad A^\star = \begin{cases} M^{-1} A^T M & \text{for bilinear forms,} \\ M^{-1} A^* M & \text{for sesquilinear forms.} \end{cases}$$

Associated with the scalar product  $\langle \cdot, \cdot \rangle_M$  are the Lie algebra  $\mathbb{L}$ , the Jordan algebra  $\mathbb{J}$  and the automorphism group  $\mathbb{G}$  defined by

$$\begin{aligned}\mathbb{L} &= \{L \in \mathbb{K}^{n \times n} : \langle Lx, y \rangle_M = -\langle x, Ly \rangle_M\} = \{L \in \mathbb{K}^{n \times n} : L^* = -L\}, \\ \mathbb{J} &= \{S \in \mathbb{K}^{n \times n} : \langle Sx, y \rangle_M = \langle x, Sy \rangle_M\} = \{S \in \mathbb{K}^{n \times n} : S^* = S\}, \\ \mathbb{G} &= \{G \in \mathbb{K}^{n \times n} : \langle Gx, Gy \rangle_M = \langle x, y \rangle_M\} = \{G \in \mathbb{K}^{n \times n} : G^* = G^{-1}\}.\end{aligned}$$

The sets  $\mathbb{L}$  and  $\mathbb{J}$  are linear subspaces. They are not closed under multiplication but are closed under inversion: if  $A \in \mathbb{S}$  is nonsingular, where  $\mathbb{S} = \mathbb{L}$  or  $\mathbb{S} = \mathbb{J}$  then  $A^{-1} \in \mathbb{S}$ . Matrices in  $\mathbb{G}$  form a Lie group under multiplication.

There are two important classes of scalar products termed unitary and orthosymmetric [15]. The scalar product  $\langle \cdot, \cdot \rangle_M$  is *unitary* if  $\alpha M$  is unitary for some  $\alpha > 0$ . A scalar product is *orthosymmetric* if

$$M = \begin{cases} \beta M^T, & \beta = \pm 1, \text{ (bilinear forms),} \\ \beta M^*, & |\beta| = 1, \text{ (sesquilinear forms).} \end{cases}$$

(See [15, Definitions A.4 and A.6] for a list of equivalent properties.) One can show that up to a scalar multiple there are only three distinct types of orthosymmetric scalar products: symmetric and skew-symmetric bilinear, and Hermitian sesquilinear [14]. We will, however, continue to include separately stated results (without separate proofs) for skew-Hermitian forms for convenience, as this is a commonly occurring special case.

In the rest of this paper we concentrate on structured matrices in the Lie algebra, the Jordan algebra or the automorphism group of a scalar product which is both unitary and orthosymmetric. These include, but are not restricted to, the structured matrices listed in Table 2.1, all of which correspond to a unitary and orthosymmetric scalar product with  $\alpha = 1$  and  $\beta = 1$ .

**2.2. Auxiliary results.** The characterizations of various structured condition numbers and backward errors for linear systems relies on the solution to the following problem: *Given a class of structured matrices  $\mathbb{S}$ , for which vectors  $x, b$  does there exist some  $A \in \mathbb{S}$  such that  $Ax = b$ ?* Mackey, Mackey and Tisseur [14] give a solution for this problem when  $\mathbb{S}$  is the Lie algebra, Jordan algebra or automorphism group of an orthosymmetric scalar product. Her and below we write  $\iota = \sqrt{-1}$ .

**THEOREM 2.1** ([14, Thm. 3.2 and Sec. 4.1]). *Let  $\mathbb{S}$  be the Lie algebra  $\mathbb{L}$ , Jordan algebra  $\mathbb{J}$  or automorphism group  $\mathbb{G}$  of an orthosymmetric scalar product  $\langle \cdot, \cdot \rangle_M$  on  $\mathbb{K}^n$ . Then for any given pair of vectors  $x, b \in \mathbb{K}^n$  with  $x \neq 0$ , there exists  $A \in \mathbb{S}$  such that  $Ax = b$  if and only if the conditions given in the following table hold:*

$\mathbb{S}$	Bilinear forms		Sesquilinear forms	
	Symmetric	Skew-symmetric	Hermitian	Skew-Hermitian
$\mathbb{J}$	always	$b^T Mx = 0$	$b^* Mx \in \mathbb{R}$	$b^* Mx \in \iota\mathbb{R}$
$\mathbb{L}$	$b^T Mx = 0$	always	$b^* Mx \in \iota\mathbb{R}$	$b^* Mx \in \mathbb{R}$
$\mathbb{G}$	$x^T Mx = b^T Mb$		$x^* Mx = b^* Mb$	

The next lemma shows that when  $\mathbb{S}$  is the Lie algebra or Jordan algebra of an orthosymmetric scalar product  $\langle \cdot, \cdot \rangle_M$  on  $\mathbb{K}^n$ , left multiplication by  $M$  maps  $\mathbb{S}$  to the sets  $\text{Skew}(\mathbb{K})$  and  $\text{Sym}(\mathbb{K})$  for bilinear forms and, to a scalar multiple of  $\text{Herm}(\mathbb{C})$  for sesquilinear forms, where

$$(2.2) \quad \begin{aligned} \text{Sym}(\mathbb{K}) &= \{A \in \mathbb{K}^{n \times n} : A^T = A\}, \\ \text{Skew}(\mathbb{K}) &= \{A \in \mathbb{K}^{n \times n} : A^T = -A\}, \\ \text{Herm}(\mathbb{C}) &= \{A \in \mathbb{C}^{n \times n} : A^* = A\} \end{aligned}$$

are the sets of symmetric and skew-symmetric matrices on  $\mathbb{K}^{n \times n}$  and Hermitian matrices, respectively. This is a key result for our unified treatment of structured condition numbers and backward errors.

LEMMA 2.2. *Let  $\mathbb{S}$  be the Lie or Jordan algebra of an orthosymmetric scalar product  $\langle \cdot, \cdot \rangle_M$ . Suppose  $A \in \mathbb{S}$ , so that  $A^* = \delta A$  where  $\delta = \pm 1$ .*

- *For bilinear forms on  $\mathbb{K}^n$  ( $\mathbb{K} = \mathbb{R}, \mathbb{C}$ ) write, by orthosymmetry,  $M = \beta M^T$  with  $\beta = \pm 1$ . Then*

$$(2.3) \quad M \cdot \mathbb{S} = \begin{cases} \text{Sym}(\mathbb{K}) & \text{if } \delta = \beta, \\ \text{Skew}(\mathbb{K}) & \text{if } \delta \neq \beta. \end{cases}$$

- *For sesquilinear forms on  $\mathbb{C}^n$  write, by orthosymmetry,  $M = \beta M^*$  with  $|\beta| = 1$ . Then*

$$(2.4) \quad M \cdot \mathbb{S} = \begin{cases} \beta^{1/2} \text{Herm}(\mathbb{C}) & \text{if } \delta = +1, \\ \iota \beta^{1/2} \text{Herm}(\mathbb{C}) & \text{if } \delta = -1. \end{cases}$$

**Proof.** For bilinear forms .

$$A \in \mathbb{S} \iff A^* = M^{-1} A^T M = \delta A \iff MA = \delta A^T M.$$

Invoking the orthosymmetry property we have  $(MA)^T = \delta(A^T M)^T = \beta \delta(MA)$  which proves (2.3).

For sesquilinear forms, an argument similar to the one for bilinear forms shows that for any  $A \in \mathbb{S}$ ,  $\beta^{1/2} MA$  is either Hermitian or skew-Hermitian. But the set of all skew-Hermitian matrices is  $\iota \text{Herm}(\mathbb{C})$  so that (2.4) holds.  $\square$

In practice, when  $A \in \mathbb{J}$  or  $\mathbb{L}$ , Lemma 2.2 suggests to rewrite  $Ax = b$  as  $\tilde{A}x = Mb$ , where  $\tilde{A} := MA$  belongs to one of the sets defined in (2.2). Then we can use a well understood structure preserving algorithm to numerically solve the modified system with  $\tilde{A}$ . For instance, if  $A$  is pseudosymmetric then  $M = \Sigma_{p,q} = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$  and  $\tilde{A} = \Sigma_{p,q} A$  is symmetric. Then we can use a block LDL<sup>T</sup> factorization [9, Ch.11] to solve the modified linear system  $\tilde{A}x = \Sigma_{p,q} b$ . However, since left multiplication of  $A$  by  $M$  changes its spectrum, mapping  $A$  to  $\tilde{A}$  may not be appropriate when solving large sparse systems of equations with an iterative method.

The next result follows immediately from Table 2.1 and Theorem 2.1.

COROLLARY 2.3. *Let  $\mathbb{S}$  be one of  $\text{Sym}(\mathbb{K})$ ,  $\text{Skew}(\mathbb{K})$  and  $\text{Herm}(\mathbb{C})$ . Then for any given pair of vectors  $x, b \in \mathbb{K}^n$  with  $x \neq 0$  there exists  $A \in \mathbb{S}$  such that  $Ax = b$  if and only if the following conditions hold:*

$\mathbb{S}$	Condition
$\text{Sym}(\mathbb{K})$	none
$\text{Skew}(\mathbb{K})$	$b^T x = 0$
$\text{Herm}(\mathbb{C})$	$b^* x \in \mathbb{R}$

For given  $x, b$  of unit 2-norm we now construct a matrix  $A$  belonging to one of the sets in (2.2) such that  $Ax = b$  and  $\|A\|_\nu \approx 1$ ,  $\nu = 2, F$ .

LEMMA 2.4. *Let  $\mathbb{S}$  be one of  $\text{Sym}(\mathbb{K})$ ,  $\text{Skew}(\mathbb{K})$  and  $\text{Herm}(\mathbb{C})$  and let  $x, b \in \mathbb{K}^n$  of unit 2-norm be such that the relevant condition in Corollary 2.3 is satisfied.*

- (i) *There exists  $A \in \mathbb{S}$  such that  $Ax = b$  and  $\|A\|_F = \sqrt{2 - \langle x, b \rangle^2} \leq \sqrt{2}$ .*
- (ii) *If  $\mathbb{S} \neq \text{Sym}(\mathbb{C})$ , there exists  $A \in \mathbb{S}$  such that  $Ax = b$  and  $\|A\|_2 = 1$ .*
- (iii) *If  $\mathbb{S} = \text{Sym}(\mathbb{C})$ , there exists  $A \in \mathbb{S}$  such that  $Ax = \bar{x}$  and  $\|A\|_2 = 1$ .*

**Proof.** (i) If  $\mathbb{S} = \text{Sym}(\mathbb{K})$  then  $A = bx^* + \bar{x}b^T - (b^T x)\bar{x}x^* \in \text{Sym}(\mathbb{K})$  is such that  $Ax = b$ . Straightforward calculation shows that  $\text{trace}(A^*A) = 2 - |b^T x|^2 \leq 2$ .

When  $\mathbb{S} = \text{Skew}(\mathbb{K})$  then by Corollary 2.3 for there to exist  $A \in \mathbb{S}$  mapping  $x$  to  $b$  we must have  $x^T b = 0$ . An easy calculation shows that  $A = bx^* - \bar{x}b^T \in \text{Skew}(\mathbb{K})$  is such that  $Ax = b$  and  $\|A\|_F = \sqrt{2}$ .

If  $\mathbb{S} = \text{Herm}(\mathbb{C})$ , then by Corollary 2.3 we must have  $b^* x \in \mathbb{R}$ . Then  $A = bx^* + xb^* - (b^* x)xx^* \in \text{Herm}(\mathbb{C})$  and satisfies  $Ax = b$ . Also,  $\text{trace}(A^*A) = 2 - |b^* x|^2 \leq 2$ .

(ii) If  $\mathbb{S} = \text{Sym}(\mathbb{R})$ , the Householder reflector  $A$  mapping the real vector  $x$  to the real vector  $b$  is real symmetric and of unit 2-norm.

From Corollary 2.3 there exists  $A \in \text{Skew}(\mathbb{K})$  mapping  $x$  to  $b$  if and only if  $b^T x = 0$ , or equivalently,  $b^* \bar{x} = 0$ . Hence there is a QR factorization  $[\bar{x}, b] = Q[e_1, -e_2]$ , where  $Q$  is unitary and  $e_k$  denotes the  $k$ th column of the identity matrix. Then  $A = Q(e_1 e_2^T - e_2 e_1^T)Q^T$  satisfies  $A^T = -A$ ,  $Ax = b$  and  $\|A\|_2 = 1$ . When  $x, b \in \mathbb{R}^n$ ,  $Q$  is orthogonal and therefore  $A$  is real skew-symmetric.

Suppose that  $\mathbb{S} = \text{Herm}(\mathbb{C})$ . For  $x \neq b$  there exists a unique unitary Householder reflector  $A$  such that  $Ax = b$  [13, Thm. 8.6]. Furthermore  $A = I + \gamma uu^*$ , where  $u = b - x$  and  $\gamma = 1/(u^* x)$ . From Corollary 2.3,  $x^* b \in \mathbb{R}$  so that  $u^* x = b^* x - 1 \in \mathbb{R} \setminus \{0\}$ . Hence  $A \in \text{Herm}(\mathbb{C})$  and  $\|A\|_2 = 1$  since  $A$  is unitary.

(iii) Let  $H$  be a unitary Householder reflector mapping  $\bar{x}$  to  $\|x\|_2 e_1 \in \mathbb{R}^n$  and let  $U = H^*$  so that  $U^* \bar{x} = \|x\|_2 e_1 = \overline{U^T x} = U^T x$  since  $\|x\|_2 e_1$  is real. Let  $A = U e_1 e_1^T U^T$ . Then  $A^T = A$  is complex symmetric and  $Ax = U e_1 e_1^T (U^T x) = U(\|x\|_2 e_1) = UU^* \bar{x} = \bar{x}$ . Also,  $\|A\|_2 = \|e_1 e_1^T\|_2 = 1$ .  $\square$

A slightly weaker bound for  $\|A\|_F$  with  $A$  real symmetric is given in [1, Thm 3] for  $\mathbb{K} = \mathbb{R}$  using a constructing proof based on a QR factorization of  $[x, b]$ . The result for the 2-norm and  $A \in \text{Skew}(\mathbb{R})$  is proved in [16, Lem. 5.1].

**2.3. Nearest structured matrix.** Suppose  $A$  has lost its structure, because of errors in its construction for example. An interesting optimization problem is to find the nearest structured matrix to  $A$ ,

$$(2.5) \quad d_{\mathbb{S}}(A) := \min\{\|A - S\| : S \in \mathbb{S}\}.$$

This problem is easy to solve when  $\mathbb{S}$  is the Jordan algebra  $\mathbb{J}$  or Lie algebra  $\mathbb{L}$  of an orthosymmetric and unitary scalar product  $\langle \cdot, \cdot \rangle_{\mathbb{M}}$  and  $\|\cdot\|$  is any unitarily invariant

norm. Indeed, if  $\langle \cdot, \cdot \rangle_{\mathbb{M}}$  is orthosymmetric then any  $A \in \mathbb{K}^{n \times n}$  can be expressed in the form

$$A = \frac{A + A^{\star}}{2} + \frac{A - A^{\star}}{2} =: A_{\mathbb{J}} + A_{\mathbb{L}},$$

where  $A_{\mathbb{J}} \in \mathbb{J}$  and  $A_{\mathbb{L}} \in \mathbb{L}$ . For any  $S \in \mathbb{J}$ , so that,  $S = S^{\star}$  we have

$$\|A - A_{\mathbb{J}}\| = \|A_{\mathbb{L}}\| = \frac{1}{2}\|A - A^{\star}\| = \frac{1}{2}\|(A - S) + (S^{\star} - A^{\star})\|.$$

Since  $\langle \cdot, \cdot \rangle_{\mathbb{M}}$  is unitary and the norm is unitarily invariant,

$$\|A - A_{\mathbb{J}}\| \leq \frac{1}{2}\|A - S\| + \frac{1}{2}\|(S - A)^{\star}\| = \|A - S\| \quad \forall S \in \mathbb{J}.$$

We find similarly that for Lie algebras  $\mathbb{L}$ ,

$$\|A - A_{\mathbb{L}}\| \leq \|A - S\| \quad \forall S \in \mathbb{L}.$$

Hence,  $S = A_{\mathbb{S}}$  with  $\mathbb{S} = \mathbb{J}$  or  $\mathbb{L}$  is the nearest structured matrix in  $\mathbb{S}$  to  $A$  and

$$d_{\mathbb{J}}(A) = \frac{1}{2}\|A - A^{\star}\|, \quad d_{\mathbb{L}}(A) = \frac{1}{2}\|A + A^{\star}\|.$$

Fan and Hoffman solved this problem for the class of symmetric matrices [4]. More recently it was solved by Cardoso, Kenney and Leite [2, Thm. 5.3] for Lie and Jordan algebras of bilinear scalar products  $\langle \cdot, \cdot \rangle_{\mathbb{M}}$  for which  $MM^T = I$  and  $M^2 = \pm I$ ; these scalar products are a subset of the unitary and orthosymmetric scalar products.

When  $\mathbb{S}$  is the set of real orthogonal matrices or unitary matrices, the solution to the nearness problem (2.5) is well-known for the Frobenius norm and is given by

$$d_{\mathbb{S}}(A) = \|A - U\|_F,$$

where  $U$  is the orthogonal or unitary factor of the polar decomposition of  $A$  [5, p. 149], [8]. To our knowledge, (2.5) is an open problem for matrix groups other than the orthogonal and unitary group. A different formulation appears also in the solution to some nonlinear structured backward errors (see Section 4.2).

**3. Structured condition numbers.** We consider normwise relative structured condition numbers that measure the sensitivity of linear systems and of matrix inversion, and investigate the structured distance to singularity.

**3.1. Linear systems.** Let  $\mathbb{S}$  be a class of structured matrices, not necessarily  $\mathbb{J}$ ,  $\mathbb{L}$  or  $\mathbb{G}$ . We define the structured normwise condition number for the linear system  $Ax = b$  with  $x \neq 0$  by

$$(3.1) \quad \text{cond}(A, x; \mathbb{S}) = \limsup_{\epsilon \rightarrow 0} \left\{ \frac{\|\Delta x\|}{\epsilon \|x\|} : (A + \Delta A)(x + \Delta x) = b + \Delta b, \right. \\ \left. A + \Delta A \in \mathbb{S}, \|\Delta A\| \leq \epsilon \|A\|, \|\Delta b\| \leq \epsilon \|b\| \right\},$$

where  $\|\cdot\|$  is an arbitrary matrix norm. Let  $\text{cond}(A, x) \equiv \text{cond}(A, x; \mathbb{K}^{n \times n})$  denote the unstructured condition number, where  $n$  is the dimension of  $A$ . Clearly,

$$\text{cond}(A, x; \mathbb{S}) \leq \text{cond}(A, x).$$

If this inequality is not always close to being attained then  $\text{cond}(A, x)$  may severely overestimate the worst case effect of structured perturbations.

Let us define for nonsingular  $A \in \mathbb{S}$  and nonzero  $x \in \mathbb{K}^n$ ,

$$(3.2) \quad \phi(A, x; \mathbb{S}) = \limsup_{\epsilon \rightarrow 0} \left\{ \frac{\|A^{-1} \Delta A x\|}{\epsilon \|A\|} : \|\Delta A\| \leq \epsilon \|A\|, A + \Delta A \in \mathbb{S} \right\}.$$

We write  $\phi(A, x) \equiv \phi(A, x; \mathbb{K}^{n \times n})$  when  $A + \Delta A$  is unstructured and  $\phi_2(A, x)$  or  $\phi_F(A, x)$  to specify that the 2-norm or Frobenius norm is used in (3.2). Similarly,  $\text{cond}_2$  and  $\text{cond}_F$  mean that  $\|\cdot\| = \|\cdot\|_2$  or  $\|\cdot\| = \|\cdot\|_F$  in (3.1).

The next lemma is useful when comparing the structured and unstructured condition numbers for linear systems. The result is due to Rump [16] and holds for any class of structured matrices. The proof in [16, Thm. 3.2 and Thm. 3.3] for  $\mathbb{K} = \mathbb{R}$  and the 2-norm extend trivially to  $\mathbb{K} = \mathbb{C}$  and  $\|\cdot\| = \|\cdot\|_F$ .

LEMMA 3.1. *Let  $A \in \mathbb{K}^{n \times n}$  be nonsingular and  $x \in \mathbb{K}^n$  be nonzero. Then for  $\nu = 2, F$ ,*

$$\frac{\phi_\nu(A, x; \mathbb{S})}{\|x\|_2} \left( \|A\|_\nu + \frac{\|b\|_2}{\|x\|_2} \right) \leq \text{cond}_\nu(A, x; \mathbb{S}) \leq \|A^{-1}\|_2 \left( \|A\|_\nu + \frac{\|b\|_2}{\|x\|_2} \right),$$

and

$$\text{cond}_\nu(A, x, \mathbb{S}) = c_\nu \left( \phi_\nu(A, x; \mathbb{S}) \frac{\|A\|_\nu}{\|x\|_2} + \frac{\|A^{-1}\|_2 \|b\|_2}{\|x\|_2} \right), \quad \frac{1}{\sqrt{2}} \leq c_\nu \leq 1.$$

The difficulty in obtaining an explicit expression for  $\phi(A, x; \mathbb{S})$  in (3.2) depends on the nature of  $\mathbb{S}$  and the matrix norm  $\|\cdot\|$ . For example, for the Frobenius norm or the 2-norm and for unstructured perturbations (i.e.,  $\mathbb{S} = \mathbb{K}^{n \times n}$ ), the supremum in (3.2) is attained by  $\Delta A = \epsilon \|A\|_\nu y x^* / \|x\|_2$ , where  $y$  is such that  $\|y\|_2 = 1$  and  $\|A^{-1} y\|_2 = \|A^{-1}\|_2$ . This implies the well known formula

$$(3.3) \quad \phi_2(A, x) = \phi_F(A, x) = \|A^{-1}\|_2 \|x\|_2,$$

so that from Lemma 3.1,

$$(3.4) \quad \text{cond}_\nu(A, x) = \|A^{-1}\|_2 \|A\|_\nu + \frac{\|A^{-1}\|_2 \|b\|_2}{\|x\|_2}, \quad \nu = 2, F.$$

The rest of this section is devoted to the study of  $\phi(A, x; \mathbb{S})$  for the structured matrices in  $\mathbb{L}$ ,  $\mathbb{J}$  or  $\mathbb{G}$  as defined in Section 2.1.

**3.1.1. Lie and Jordan algebras.** D.J. Higham [6] proves that for real symmetric structures,

$$\text{cond}_\nu(A, x; \text{Sym}(\mathbb{R})) = \text{cond}_\nu(A, x), \quad \nu = 2, F$$

and Rump [16] shows that equality also holds in the 2-norm for persymmetric and skew-symmetric structures. These are examples of Lie and Jordan algebras (see Table 2.1). We extend these results to all Lie and Jordan algebras of orthosymmetric and unitary scalar products. Unlike the proofs in [16], our unifying proof does need to consider each algebra individually.

**THEOREM 3.2.** *Let  $\mathbb{S}$  be the Lie algebra or Jordan algebra of a unitary and orthosymmetric scalar product  $\langle \cdot, \cdot \rangle_M$  on  $\mathbb{K}^n$ . For nonsingular  $A \in \mathbb{S}$  and nonzero  $x \in \mathbb{K}^n$ , we have*

$$\phi_\nu(A, x; \mathbb{S}) = c_\nu \phi_\nu(A, x), \quad \frac{1}{\sqrt{2}} \leq c_\nu \leq 1, \quad \nu = 2, F.$$

In particular,  $c_2 = 1$  when  $M \cdot \mathbb{S} \neq \text{Sym}(\mathbb{C})$ .

**Proof.** Clearly  $\phi_\nu(A, x; \mathbb{S}) \leq \phi_\nu(A, x)$ . Since Lie and Jordan algebras are linear subspaces of  $\mathbb{K}^{n \times n}$ ,  $\phi_\nu(A, x; \mathbb{S})$  can be rewritten as

$$\phi_\nu(A, x; \mathbb{S}) = \sup \{ \|A^{-1} \Delta A x\|_2 : \|\Delta A\|_\nu \leq 1, \Delta A \in \mathbb{S} \}.$$

Unitarity of  $\langle \cdot, \cdot \rangle_M$  means that  $\alpha M$  is unitary for some  $\alpha > 0$  and we have  $\phi_\nu(A, x; \mathbb{S}) = \phi_\nu(\alpha M A, x; M \cdot \mathbb{S})$  so that from (3.3) we just have to prove that  $\tilde{c}_\nu \phi_\nu(\alpha M A, x; M \cdot \mathbb{S}) \geq \|A^{-1}\|_2 \|x\|_2$  with  $\tilde{c}_F = \sqrt{2}$  and,  $\tilde{c}_2 = 1$  if  $M \cdot \mathbb{S} \neq \text{Sym}(\mathbb{C})$  and  $\tilde{c}_2 = \sqrt{2}$  otherwise. For that we show that for an appropriate  $u \in \mathbb{K}^n$  of unit 2-norm such that  $\|A^{-1}\|_2 = \|(\alpha M A)^{-1}\|_2 = \|(\alpha M A)^{-1} u\|_2$  there exists  $E \in M \cdot \mathbb{S}$  such that  $Ex = \|x\|_2 u$  and  $\|E\|_\nu \leq \tilde{c}_\nu$ . Thus,

$$\|A^{-1}\|_2 \|x\|_2 = \|(\alpha M A)^{-1} u\|_2 \|x\|_2 = \|(\alpha M A)^{-1} Ex\|_2 \leq \tilde{c}_\nu \phi_\nu(\alpha M A, x; M \cdot \mathbb{S}).$$

Without loss of generality, we assume now that  $\|x\|_2 = 1$ . Since  $\langle \cdot, \cdot \rangle_M$  is orthosymmetric, Lemma 2.2 tells us that  $M \cdot \mathbb{S}$  is one of  $\text{Sym}(\mathbb{K})$ ,  $\text{Skew}(\mathbb{K})$  and  $\text{Herm}(\mathbb{C})$  (strictly speaking this should be  $\beta \text{Herm}(\mathbb{C})$  with  $|\beta| = 1$  but since we are taking norms,  $\beta$  can be dropped). We now examine each of these three cases.

(i)  $M \cdot \mathbb{S} = \text{Sym}(\mathbb{K})$ . Lemma 2.4 says there exists  $E \in \text{Sym}(\mathbb{K})$  such that  $Ex = u$  and  $\|E\|_\nu \leq \tilde{c}_\nu$  with  $\tilde{c}_\nu = \sqrt{2}$  and, in particular,  $\tilde{c}_2 = 1$  if  $\mathbb{K} = \mathbb{R}$ .

(ii)  $M \cdot \mathbb{S} = \text{Skew}(\mathbb{K})$ .  $A \in \mathbb{S}$  nonsingular implies that  $(\alpha M A)^{-1} \in \text{Skew}(\mathbb{K})$ . But the singular values of nonsingular real or complex skew-symmetric matrices have even multiplicity [15, Thm. 8.4]. Hence there exist orthogonal vectors  $u_1$  and  $u_2$  in  $\mathbb{K}^n$  such that  $\|u_1\|_2 = \|u_2\|_2 = 1$  and  $\|(\alpha M A)^{-1} u_1\|_2 = \|(\alpha M A)^{-1} u_2\|_2 = \|(\alpha M A)^{-1}\|_2 = \|A^{-1}\|_2$ . Let  $u \in \text{span}\{u_1, u_2\}$  such that  $u^T x = 0$  and  $\|u\|_2 = 1$ . Then necessarily,  $\|(\alpha M A)^{-1} u\|_2 = \|(\alpha M A)^{-1}\|_2$ . From Lemma 2.4, there exists  $E \in \text{Skew}(\mathbb{K})$  such that  $Ex = u$  and  $\|E\|_\nu = \tilde{c}_\nu$  with  $\tilde{c}_2 = 1$  and  $\tilde{c}_F = \sqrt{2}$  since  $u^T x = 0$ . Note that  $E$  depends on the choice of norm.

(iii)  $M \cdot \mathbb{S} = \text{Herm}(\mathbb{C})$ . We choose  $u$  so that the extra constraint  $u^*x \in \mathbb{R}$  is satisfied. Hence from Lemma 2.4, there exists  $E \in \text{Herm}(\mathbb{C})$  such that  $Ex = u$  and  $\|E\|_\nu \leq \tilde{c}_\nu$  with  $\tilde{c}_2 = 1$  and  $\tilde{c}_F = \sqrt{2}$ .  $\square$

The condition  $M \cdot \mathbb{S} \neq \text{Sym}(\mathbb{C})$  in Theorem 3.2 can also be read as:  $\mathbb{S}$  is not the Jordan algebra of a symmetric bilinear form on  $\mathbb{C}^n$  nor the Lie algebra of a skew-symmetric bilinear form on  $\mathbb{C}^n$ . Examples of structured matrices for which  $M \cdot \mathbb{S} \neq \text{Sym}(\mathbb{C})$  include complex symmetric, complex pseudo-symmetric and complex  $J$ -symmetric matrices (see table 2.1).

Lemma 3.1 and Theorem 3.2 together yield the following result.

**THEOREM 3.3.** *Let  $\mathbb{S}$  be the Lie or Jordan algebra of a unitary and orthosymmetric scalar product  $\langle \cdot, \cdot \rangle_M$  on  $\mathbb{K}^n$ . Let  $A \in \mathbb{S}$  be nonsingular and  $x \in \mathbb{K}^n$  with  $x \neq 0$  be given. Then*

$$\text{cond}_\nu(A, x; \mathbb{S}) = c_\nu \text{cond}_\nu(A, x), \quad \frac{1}{\sqrt{2}} \leq c_\nu \leq 1, \quad \nu = 2, F.$$

In particular,  $c_2 = 1$  when  $M \cdot \mathbb{S} \neq \text{Sym}(\mathbb{C})$ .

We conclude from Theorem 3.3 that for many Lie and Jordan algebras, the constraint  $A + \Delta A \in \mathbb{S}$  has little or no effect on the condition number. This is certainly true for all the examples in Table 2.1.

**3.1.2. Automorphism groups.** When  $\mathbb{S}$  is a smooth manifold the task of computing the supremum (3.2) simplifies to a linearly constrained optimization problem. This was already observed in [11] for a different supremum.

**LEMMA 3.4.** *Let  $\mathbb{S}$  be a smooth real or complex manifold. Then for  $A \in \mathbb{S}$  nonsingular we have*

$$(3.5) \quad \phi(A, x; \mathbb{S}) = \sup \{ \|A^{-1}Ex\| : \|E\| = 1, E \in T_A\mathbb{S} \},$$

where  $T_A\mathbb{S}$  is the tangent space at  $A$ .

**Proof.** Let  $E \in T_A\mathbb{S}$  with  $\|E\| = 1$ . Then there is a smooth curve  $g_E : (-\epsilon, \epsilon) \rightarrow \mathbb{K}^{n \times n}$  satisfying  $g_E(0) = 0$ ,  $g'_E(0) = E$  and  $A + g_E(t) \in \mathbb{S}$  for all  $t$ . We have

$$\lim_{t \rightarrow 0} \frac{g_E(t)}{\|g_E(t)\|} = \lim_{t \rightarrow 0} \frac{Et + O(|t|^2)}{\|Et + O(|t|^2)\|} = E.$$

Hence

$$\lim_{t \rightarrow 0} \frac{\|A^{-1}g_E(t)x\|}{\|g_E(t)\|} = \|A^{-1}Ex\|.$$

This implies  $\phi(A, x; \mathbb{S}) \geq \sup \{ \|A^{-1}Ex\| : \|E\| = 1, E \in T_A\mathbb{S} \}$ . Equality holds since the curves  $g_E$  form a covering of an open neighborhood of  $A \in \mathbb{S}$ .  $\square$

An automorphism group  $\mathbb{G}$  forms a smooth manifold. The Fréchet derivative of the function

$$F(A) = \begin{cases} A^T M A - M & (\text{bilinear forms}) \\ A^* M A - M & (\text{sesquilinear forms}) \end{cases}$$

at  $A \in \mathbb{K}^{n \times n}$  can be represented as the linear function

$$J_A(X) = \begin{cases} A^T M X + X^T M A & \text{(bilinear forms)} \\ A^* M X + X^* M A & \text{(sesquilinear forms)}. \end{cases}$$

The tangent space  $T_A \mathbb{G}$  at  $A \in \mathbb{G}$  is given by the kernel of  $J_A$ ,

$$(3.6) \quad T_A \mathbb{G} = \{X : J_A(X) = 0\} = \{AH : H^* = -H\} = A \cdot \mathbb{L},$$

where  $\mathbb{L}$  is the Lie algebra of  $\langle \cdot, \cdot \rangle_M$ . For  $\mathbb{S} = \mathbb{G}$ , the supremum (3.5) can then be rewritten as

$$(3.7) \quad \phi(A, x; \mathbb{G}) = \sup \{\|Hx\| : \|AH\| = 1, H \in \mathbb{L}\}.$$

In a similar way to [7] and [11], an explicit expression for  $\phi(A, x; \mathbb{G})$  can be obtained for the Frobenius norm. Let us rewrite

$$Hx = (x^T \otimes I_n) \text{vec}(H),$$

where  $\otimes$  denotes the Kronecker product and  $\text{vec}$  is the operator that stacks the columns of a matrix into one long vector. We refer to Lancaster and Tismenetsky [12, Chap. 12] for properties of the  $\text{vec}$  operator and the Kronecker product. Since  $\mathbb{L}$  is a linear vector space of dimension  $m \leq n^2$ , there is an  $n^2 \times m$  matrix  $B$  such that for every  $H \in \mathbb{L}$  there exists a uniquely defined parameter vector  $q$  with

$$(3.8) \quad \text{vec}(H) = Bq.$$

Let  $(I \otimes A)B = QR$  be a reduced QR factorization of  $(I \otimes A)B$ , i.e.,  $Q \in \mathbb{K}^{n^2 \times m}$ ,  $R \in \mathbb{K}^{m \times m}$ , and let  $p = Rq$ . Then,

$$\text{vec}(AH) = (I \otimes A) \text{vec}(H) = (I \otimes A)Bq = Qp$$

so that  $\|AH\|_F = \|\text{vec}(AH)\|_2 = \|p\|_2$  and

$$Hx = (x^T \otimes I_n) \text{vec}(H) = (x^T \otimes I_n)(I \otimes A)^{-1} \text{vec}(AH) = (x^T \otimes A^{-1})Qp.$$

Hence (3.7) becomes

$$\phi_F(A, x; \mathbb{G}) = \sup \{\|(x^T \otimes A^{-1})Qp\|_2 : \|p\|_2 = 1, p \in \mathbb{K}^m\} = \|(x^T \otimes A^{-1})Q\|_2.$$

Using Lemma 3.1 we then have, up to a small scalar multiple, a directly computable expression for the structured condition number.

**THEOREM 3.5.** *Let  $\mathbb{G}$  be the automorphism group of any scalar product and let  $\mathbb{L}$  be the associated Lie algebra. Then*

$$\text{cond}_F(A, x, \mathbb{G}) = c_F \left( \|(x^T \otimes A^{-1})Q\|_2 \frac{\|A\|_F}{\|x\|_2} + \frac{\|A^{-1}\|_2 \|b\|_2}{\|x\|_2} \right), \quad \frac{1}{\sqrt{2}} \leq c_F \leq 1,$$

where  $(I \otimes A)B = QR$  is a reduced QR factorization of  $(I \otimes A)B$  and  $B$  is a pattern matrix for  $\mathbb{L}$  in the sense of (3.8).

The expression for  $\text{cond}_F(A, x; \mathbb{G})$  in Theorem 3.5 has two disadvantages. First, it is expensive to compute, since it requires the QR factorization of an  $n^2 \times m$  matrix. Note that since  $A \in \mathbb{G}$ ,  $A^{-1} = A^\star$  with  $A^\star$  given in (2.1) so that we do not necessarily need to invert  $A$ . Second, it is difficult to compare with the unstructured condition number  $\text{cond}_F(A, x)$  in (3.4). However, when  $\langle \cdot, \cdot \rangle_M$  is unitary and orthosymmetric, we can bound the ratio  $\phi_\nu(A, x; \mathbb{G})/\phi_\nu(A, x)$  below.

**LEMMA 3.6.** *Let  $\mathbb{G}$  be the automorphism group of a unitary and orthosymmetric scalar product on  $\mathbb{K}^n$ . Let  $A \in \mathbb{G}$  and  $x \in \mathbb{K}^n$  with  $x \neq 0$  be given. Then*

$$\frac{\phi_\nu(A, x)}{\sqrt{2}\|A\|_2\|A^{-1}\|_2} \leq \phi_\nu(A, x; \mathbb{G}) \leq \phi_\nu(A, x), \quad \nu = 2, F,$$

**Proof.** Since  $\alpha M$  is unitary for some  $\alpha > 0$ , (3.7) can be rewritten as

$$\phi_\nu(A, x; \mathbb{G}) = \sup \left\{ \|\tilde{H}x\|_2 : \|A(\alpha M)^{-1}\tilde{H}\|_\nu = 1, \tilde{H} \in \tilde{\mathbb{L}} \right\},$$

where  $\tilde{H} = \alpha MH$  and  $\tilde{\mathbb{L}} = M \cdot \mathbb{S}$ . Since  $\langle \cdot, \cdot \rangle_M$  is orthosymmetric, Lemma 2.2 implies that  $\tilde{\mathbb{L}}$  is one of  $\text{Sym}(\mathbb{K})$ ,  $\text{Skew}(\mathbb{K})$  or  $\text{Herm}(\mathbb{C})$ . Let  $u \in \mathbb{K}^n$  be of unit 2-norm and such that  $x, u$  satisfy the conditions of Corollary 2.3. Lemma 2.4 tells us there exists  $S \in \tilde{\mathbb{L}}$  such that  $S(x/\|x\|_2) = u$  with  $\|S\|_2 \leq \|S\|_F \leq \sqrt{2}$ . Let  $\tilde{H} = \xi S \in \tilde{\mathbb{L}}$  with  $\xi > 0$  such that  $\|A(\alpha M)^{-1}\tilde{H}\|_\nu = 1$ ,  $\nu = 2, F$ . This implies  $1 = \xi\|A(\alpha M)^{-1}S\|_\nu \leq \xi\|A\|_2\|S\|_\nu \leq \sqrt{2}\xi\|A\|_2$  so that

$$\phi_\nu(A, x; \mathbb{G}) \geq \|\tilde{H}x\|_2 = \xi\|x\|_2 \geq \frac{\|x\|_2}{\sqrt{2}\|A\|_2} = \frac{\phi_\nu(A, x)}{\sqrt{2}\|A\|_2\|A^{-1}\|_2}. \quad \square$$

Lemma 3.6, together with Lemma 3.1 and the explicit expression for  $\text{cond}_\nu(A, x)$  in (3.4), yields bounds for  $\text{cond}(A, x; \mathbb{G})$ .

**THEOREM 3.7.** *Let  $\mathbb{G}$  be the automorphism group of a unitary and orthosymmetric scalar product on  $\mathbb{K}^n$ . Let  $A \in \mathbb{G}$  and  $x \in \mathbb{K}^n$  with  $x \neq 0$  be given. Then*

$$\frac{\text{cond}_\nu(A, x)}{\sqrt{2}\|A\|_2\|A^{-1}\|_2} \leq \text{cond}_\nu(A, x; \mathbb{G}) \leq \text{cond}_\nu(A, x), \quad \nu = 2, F.$$

Theorem 3.7 shows that when  $A$  is well-conditioned (i.e.,  $\|A\|_2\|A^{-1}\|_2 \approx 1$ ), the structured and unstructured condition numbers for  $Ax = b$  are equal or nearly equal. This is the case for the real orthogonal group and the unitary group. For ill-conditioned  $A$ , the bounds may not be sharp, as illustrated by the following example. Suppose that  $M = J$  and that  $\langle \cdot, \cdot \rangle_J$  is a real bilinear form. Then  $\mathbb{G}$  is the set of real symplectic matrices (see Table 2.1). Consider the symplectic matrix

$$(3.9) \quad A = \begin{bmatrix} D & D \\ 0 & D^{-1} \end{bmatrix}, \quad D = \text{diag}(10^{-6}, 10^2, 2)$$

and define the ratio

$$\rho = \text{cond}_F(A, x; \mathbb{G}) / \text{cond}_F(A, x)$$

between the structured and unstructured condition numbers. For this particular example, Theorem 3.7 provides the non-sharp bounds  $7 \times 10^{-13} \leq \rho \leq 1$ . We use the computable expression in Theorem 3.5 with  $c_F = 1$  to approximate  $\text{cond}_F(A, x; \mathbb{G})$ . For  $x = [1, 0, 0.1, 0, 0.1, -1]^T$  we obtain  $\rho \approx 8 \times 10^{-5}$  showing that  $\text{cond}_F(A, x; \mathbb{G}) \ll \text{cond}_F(A, x)$  may happen.

**3.2. Matrix inversion.** Let  $\mathbb{S}$  be a class of structured matrices. The structured condition number for the matrix inverse can be defined by

$$(3.10) \quad \kappa(A; \mathbb{S}) = \limsup_{\epsilon \rightarrow 0} \left\{ \frac{\|(A + \Delta A)^{-1} - A^{-1}\|}{\epsilon \|A^{-1}\|} : A + \Delta A \in \mathbb{S}, \|\Delta A\| \leq \epsilon \|A\| \right\}.$$

When  $\Delta A$  is unstructured we write  $\kappa(A) \equiv \kappa(A; \mathbb{K}^{n \times n})$ . It is well-known that for the 2- and Frobenius norms  $\kappa(A)$  has the characterization

$$(3.11) \quad \kappa_2(A) = \|A\|_2 \|A^{-1}\|_2, \quad \kappa_F(A) = \frac{\|A\|_F \|A^{-1}\|_2^2}{\|A^{-1}\|_F}.$$

See [9, Thm. 6.4] for  $\kappa_2$  and [6] for  $\kappa_F$ . Again, when  $\mathbb{S}$  is a smooth manifold (3.10) simplifies to a linearly constrained optimization problem.

**LEMMA 3.8.** *Let  $\mathbb{S}$  be a smooth real or complex manifold. Then for  $A \in \mathbb{S}$  nonsingular we have*

$$(3.12) \quad \kappa(A; \mathbb{S}) = \frac{\|A\|}{\|A^{-1}\|} \sup\{\|A^{-1}EA^{-1}\| : \|E\| = 1, E \in T_A\mathbb{S}\},$$

where  $T_A\mathbb{S}$  is the tangent space at  $A$ .

**Proof.** The proof is similar to that of Lemma 3.4 and makes use of the expansion

$$(A + \Delta A)^{-1} - A^{-1} = -A^{-1}\Delta AA^{-1} + O(\|\Delta A\|^2),$$

in the definition of  $\kappa(A; \mathbb{S})$  in (3.10).  $\square$

When  $\mathbb{S}$  is the Lie algebra or Jordan algebra of a scalar product, the tangent space at  $\mathbb{S}$  is  $\mathbb{S}$  itself:  $T_A\mathbb{S} = \mathbb{S}$ . If we restrict  $\langle \cdot, \cdot \rangle_M$  to be orthosymmetric and unitary then there is equality between the structured and unstructured condition numbers for both the 2-norm and Frobenius norm.

**THEOREM 3.9.** *Let  $\mathbb{S}$  be the Lie algebra or Jordan algebra of a unitary and orthosymmetric scalar product  $\langle \cdot, \cdot \rangle_M$  on  $\mathbb{K}^n$ . For nonsingular  $A \in \mathbb{S}$  we have*

$$\kappa_\nu(A; \mathbb{S}) = \kappa_\nu(A), \quad \nu = 2, F.$$

**Proof.** Using the inequality  $\|ABC\|_\nu \leq \|A\|_2 \|B\|_\nu \|C\|_2$  in (3.12) gives

$$\kappa_\nu(A; \mathbb{S}) \leq \|A\|_\nu \|A^{-1}\|_2^2 / \|A^{-1}\|_\nu = \kappa_\nu(A).$$

Since  $\alpha M$  is unitary for some  $\alpha > 0$ , we have  $\kappa_\nu(A; \mathbb{S}) = \kappa_\nu(\tilde{A}; M \cdot \mathbb{S})$  with  $\tilde{A} = \alpha M A$ . Now assume that there exists  $E \in M \cdot \mathbb{S}$  such that

$$(3.13) \quad \|E\|_\nu = 1, \quad \|\tilde{A}^{-1} E \tilde{A}^{-1}\|_\nu = \|A^{-1}\|_2^2.$$

Then (3.12) implies

$$\kappa(\tilde{A}, M \cdot \mathbb{S}) \geq \frac{\|\tilde{A}\|_\nu}{\|\tilde{A}^{-1}\|_\nu} \|\tilde{A}^{-1} E \tilde{A}^{-1}\|_\nu = \frac{\|A\|_\nu}{\|A^{-1}\|_\nu} \|A^{-1}\|_2^2 = \kappa_\nu(A).$$

Hence to complete the proof we just need to construct  $E \in M \cdot \mathbb{S}$  satisfying (3.13). Since the scalar product is orthosymmetric, we have from Lemma 2.2 that  $M \cdot \mathbb{S}$  is one of  $\text{Sym}(\mathbb{K})$ ,  $\text{Skew}(\mathbb{K})$  and  $\text{Herm}(\mathbb{C})$ .

(i)  $M \cdot \mathbb{S} = \text{Sym}(\mathbb{K})$ . Let  $\tilde{A}^{-1} = U \Sigma U^T$  be the Takagi factorization of  $\tilde{A}^{-1} \in \text{Sym}(\mathbb{K})$ , where  $U$  is unitary and  $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$  with  $\|A^{-1}\|_2 = \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$  [10, Cor. 4.4.4]. When  $\mathbb{K} = \mathbb{R}$ ,  $U$  is orthogonal and  $\tilde{A}^{-1} = U \Sigma U^T$  is the singular value decomposition of  $\tilde{A}^{-1}$ . Take  $E = \bar{U} e_1 e_1^T U^*$ .

(ii)  $M \cdot \mathbb{S} = \text{Skew}(\mathbb{K})$ . We consider the skew-symmetric analog of the Takagi factorization (Problems 25 and 26 in [10, Sec. 4.4]). Since  $\tilde{A}^{-1} \in \text{Skew}(\mathbb{K})$ , there exists a unitary matrix  $U$  such that  $\tilde{A}^{-1} = U D U^T$  where  $D = D_1 \oplus \dots \oplus D_{n/2}$ , with  $D_j = \begin{bmatrix} 0 & +z_j \\ -z_j & 0 \end{bmatrix}$ ,  $0 \neq z_j \in \mathbb{C}$ ,  $j = 1:n/2$ . When  $\mathbb{K} = \mathbb{R}$ ,  $U$  is orthogonal. Assume that the  $z_j$  are ordered such that  $\|A^{-1}\|_2 = |z_1| \geq |z_2| \geq \dots \geq |z_k|$ . Take  $E = c_\nu \bar{U} (e_1 e_2^T - e_2 e_1^T) U^*$ ,  $\nu = 2$ ,  $F$  with  $c_2 = 1$  and  $c_F = 1/\sqrt{2}$ .

(iii)  $M \cdot \mathbb{S} = \text{Herm}(\mathbb{C})$ . Take  $E = U e_1 e_1^* U^*$  where  $U$  is the unitary factor in the singular value decomposition of  $\tilde{A}^{-1}$ .  $\square$

When  $\mathbb{S}$  is the automorphism group  $\mathbb{G}$  of a scalar product there is a directly computable expression for  $\kappa_F(A; \mathbb{G})$ . Its derivation is similar to that of  $\text{cond}(A, x; \mathbb{G})$  described in section 3.1.2.

**THEOREM 3.10.** *Let  $\mathbb{G}$  be the automorphism group of any scalar product and let  $\mathbb{L}$  be the associated Lie algebra. Then*

$$\kappa_F(A; \mathbb{G}) = \frac{\|A\|_F}{\|A^{-1}\|_F} \|(A^T \otimes A)^{-1} Q\|_2.$$

where  $(I \otimes A)B = QR$  is a reduced QR factorization of  $(I \otimes A)B$  and  $B$  is a pattern matrix for  $\mathbb{L}$  in the sense of (3.8).

It is well-known that  $\kappa_2(A) \geq 1$ . This is also true for  $\kappa_2(A; \mathbb{G})$  when  $\mathbb{G}$  is the automorphism group of a unitary and orthosymmetric scalar product. Indeed let  $u, v \in \mathbb{K}^n$  be such that  $A^{-1}u = \|A^{-1}\|_2 v$  with  $\|u\|_2 = \|v\|_2 = 1$ . Then, from (3.12), (3.6) and (3.7),

$$\begin{aligned} \kappa_2(A; \mathbb{G}) &= \frac{\|A\|_2}{\|A^{-1}\|_2} \sup\{\|HA^{-1}\|_2 : \|AH\|_2 = 1, H \in \mathbb{L}\} \\ &\geq \frac{\|A\|_2}{\|A^{-1}\|_2} \|A^{-1}\|_2 \sup\{\|Hv\|_2 : \|AH\|_2 = 1, H \in \mathbb{L}\} \\ &= \|A\|_2 \phi_2(A, v; \mathbb{G}). \end{aligned}$$

But from Lemma 3.6,  $\phi_2(A, v; \mathbb{G}) \geq 1/\|A\|_2$  so that

$$1 \leq \kappa_2(A; \mathbb{G}) \leq \kappa_2(A).$$

Clearly, if  $A$  is well conditioned then  $\kappa_2(A; \mathbb{G}) \approx \kappa_2(A)$ . This is certainly true for the orthogonal and unitary groups. For the symplectic matrix  $A$  defined in (3.9), we find that  $\kappa_{\mathbb{F}}(A; \mathbb{G}) \approx \kappa_2(A; \mathbb{G}) = 1.0001$  but that  $\kappa_2(A) = 1 \times 10^{12}$ , showing that  $\kappa_2(A; \mathbb{G}) \ll \kappa_2(A)$  may happen.

**3.3. Nearness to singularity.** The *structured distance to singularity* of a matrix is defined by

$$(3.14) \quad \delta(A; \mathbb{S}) = \min \left\{ \epsilon : \|\Delta A\| \leq \epsilon \|A\|, A + \Delta A \text{ singular}, A + \Delta A \in \mathbb{S} \right\}.$$

D. J. Higham [6] showed that for symmetric perturbations,

$$\delta_\nu(A; \text{Sym}(\mathbb{R})) = \delta_\nu(A), \quad \nu = 2, F,$$

so that the constraint  $\Delta A = \Delta A^T$  has no effect on the distance. This is a special case of the following more general result.

**THEOREM 3.11.** *Let  $\mathbb{S}$  be the Lie or Jordan algebra of a unitary and orthosymmetric scalar product. Let nonsingular  $A \in \mathbb{S}$  be given. Then*

$$c_\nu \delta_\nu(A; \mathbb{S}) = \delta_\nu(A) = \frac{1}{\|A\|_\nu \|A^{-1}\|_2}, \quad \nu = 2, F,$$

with  $c_2 = 1$ ,  $c_F = 1$  if  $M \cdot \mathbb{S} \neq \text{Skew}(\mathbb{K})$  and  $1 \leq c_F \leq \sqrt{2}$ .

**Proof.** It is well-known that the relative distance to singularity is the reciprocal of the condition number  $\kappa(A)$  or a scalar multiple of it [9, p.111]. Clearly,  $\delta_\nu(A; \mathbb{S}) \geq \delta_\nu(A)$ .

Since the scalar product  $\langle \cdot, \cdot \rangle_M$  is unitary,  $\alpha M$  is unitary for some  $\alpha > 0$  and we have  $\delta_\nu(A; \mathbb{S}) = \delta_\nu(\tilde{A}, \tilde{\mathbb{S}})$ , where  $\tilde{A} = \alpha M A$  and  $\tilde{\mathbb{S}} = M \cdot \mathbb{S}$ . Since the scalar product is orthosymmetric we have from Lemma 2.2 that  $\tilde{\mathbb{S}}$  is one of  $\text{Sym}(\mathbb{K})$ ,  $\text{Skew}(\mathbb{K})$  and  $\beta \text{Herm}(\mathbb{C})$  with  $|\beta| = 1$ . Without loss of generality, assume that  $\|\tilde{A}\|_\nu = 1$ . Since  $(\tilde{A} + \Delta \tilde{A})^{-1} = (I + \tilde{A}^{-1} \Delta \tilde{A})^{-1} \tilde{A}^{-1}$  we just need to find  $\Delta \tilde{A} \in M \cdot \mathbb{S}$  such that  $I + \tilde{A}^{-1} \Delta \tilde{A}$  is singular and  $\|\Delta \tilde{A}\|_\nu \leq 1/\|A^{-1}\|_2$ . This is achieved by taking  $\Delta \tilde{A} = -E/\|A^{-1}\|_2$ , where  $E$  is the perturbation used in the proof of Theorem 3.9.  $\square$

Matrices in automorphism groups are nonsingular so  $\delta(A; \mathbb{G}) = \infty$ . This ends our treatment of various structured condition numbers.

**4. Normwise structured backward errors.** When solving a linear system  $Ax = b$  with  $A \in \mathbb{S} \subset \mathbb{K}^{n \times n}$ , one is interested in whether a computed solution  $\hat{x}$  solves a nearby structured system: for example whether the particular structured backward error

$$(4.1) \quad \mu(\hat{x}; \mathbb{S}) = \min \{ \|\Delta A\| : (A + \Delta A)\hat{x} = b, A + \Delta A \in \mathbb{S} \}$$

is relatively small. Note that in (4.1), only the coefficient matrix  $A$  is perturbed. Let  $\Delta A_{\text{opt}}$  be the optimal solution defined by  $\|\Delta A_{\text{opt}}\| = \mu(\hat{x}; \mathbb{S})$  and let  $r = b - A\hat{x}$  be the residual vector. It is well known that for unstructured perturbations,

$$\mu_\nu(\hat{x}) \equiv \mu_\nu(\hat{x}; \mathbb{K}^{n \times n}) = \|r\|_2 / \|\hat{x}\|_2, \quad \nu = 2, F$$

and this is achieved by  $\Delta A_{\text{opt}} = r\hat{x}^*/(\hat{x}^*\hat{x})$ . Unlike the unstructured case, there may not be any solution  $\Delta A$  to  $(A + \Delta A)\hat{x} = b$  with  $A + \Delta A \in \mathbb{S}$ . In this case we write that  $\mu(\hat{x}; \mathbb{S}) = \infty$ . If one uses a structure preserving algorithm to solve the linear system then  $A + \Delta A$  is structured and the backward error is guaranteed to be finite. This may not be the case if the system is solved using a linear solver that does not take advantage of the structure even if it is backward stable.

When both  $A$  and  $b$  are perturbed we define the corresponding structured backward error by

$$(4.2) \quad \eta(\hat{x}; \mathbb{S}) = \min\{\epsilon : (A + \Delta A)\hat{x} = b + \Delta b, \quad A + \Delta A \in \mathbb{S}, \\ \|\Delta A\| \leq \epsilon\|A\|, \quad \|\Delta b\| \leq \epsilon\|b\|\}.$$

Again, for unstructured perturbations we have the well-known explicit expression [9, Thm. 7.1],

$$(4.3) \quad \eta_\nu(\hat{x}) \equiv \eta_\nu(\hat{x}; \mathbb{K}^{n \times n}) = \frac{\|r\|_2}{\|A\|_\nu \|\hat{x}\|_2 + \|b\|_2}, \quad \nu = 2, F.$$

The structured backward error  $\eta(\hat{x}; \mathbb{S})$  is harder to analyze than  $\mu(\hat{x}; \mathbb{S})$ . The constraint  $A + \Delta A \in \mathbb{S}$  generally implies an extra constraint on the perturbation  $\Delta b$ . For example, if  $\mathbb{S} = \text{Skew}(\mathbb{R})$  then for there to exist  $\Delta A$  skew-symmetric such that  $(A + \Delta A)\hat{x} = b + \Delta b$  one need  $\hat{x}^T(b + \Delta b) = 0$  (see Corollary 2.3). On the other hand one may have  $\eta(\hat{x}; \mathbb{S})$  finite but  $\mu(\hat{x}; \mathbb{S}) = \infty$ . The next theorem shows that when the particular backward error  $\mu(\hat{x}; \mathbb{S})$  exists, its study provides useful information on  $\eta(\hat{x}; \mathbb{S})$ .

**THEOREM 4.1.** *Let  $A \in \mathbb{S}$ , where  $\mathbb{S}$  is a class of structured matrices and let  $\hat{x} \neq 0$  be an approximate solution to  $Ax = b$ . If  $\mu_\nu(\hat{x}; \mathbb{S})$  is finite with  $\mu_\nu(\hat{x}; \mathbb{S}) \leq c_\nu \mu_\nu(\hat{x})$  then for  $\eta_\nu(\hat{x}) < 1$ ,*

$$(4.4) \quad \eta_\nu(\hat{x}) \leq \eta_\nu(\hat{x}; \mathbb{S}) \leq \frac{2c_\nu \eta_\nu(\hat{x})}{1 - \eta_\nu(\hat{x})}, \quad \nu = 2, F.$$

**Proof.** The proof is modelled from the solution to [9, Exercise 7.7] concerning the unstructured case. Clearly, from the definitions of  $\eta_\nu(\hat{x}; \mathbb{S})$  and  $\mu_\nu(\hat{x}; \mathbb{S})$ ,

$$(4.5) \quad \eta_\nu(\hat{x}) \leq \eta_\nu(\hat{x}; \mathbb{S}) \leq \frac{\mu_\nu(\hat{x}; \mathbb{S})}{\|A\|_\nu} \leq c_\nu \frac{\mu_\nu(\hat{x})}{\|A\|_\nu} = c_\nu \frac{\|r\|_2}{\|A\|_\nu \|\hat{x}\|_2}.$$

Let  $\epsilon = \eta_\nu(\hat{x})$ . Then from (4.3) we have that  $\|r\|_2 \leq \epsilon(\|A\|_\nu \|\hat{x}\|_2 + \|b\|_2)$  with  $\|b\|_2 = \|(A + \Delta A)\hat{x} - \Delta b\|_2 \leq (1 + \epsilon)\|A\|_\nu \|\hat{x}\|_2 + \epsilon\|b\|_2$  yielding  $\|b\|_2 \leq (1 - \epsilon)^{-1}(1 + \epsilon)\|A\|_\nu \|\hat{x}\|_2$ . Thus

$$\|r\|_2 \leq \frac{2\epsilon}{1 - \epsilon} \|A\|_\nu \|\hat{x}\|_2$$

and (4.5) yields the required bounds.  $\square$

Theorem 4.1 shows that when  $\eta_\nu(\hat{x})$  and  $c_\nu$  are small, the structured relative backward error is small. In the rest of this section, we concentrate on the study of  $\mu(\hat{x}; \mathbb{S})$  and particular structures  $\mathbb{S}$ .

**4.1. Lie and Jordan algebras.** For symmetric or Hermitian structures, the optimal solution to the minimization problem (4.1) is well-known for the Frobenius norm [1], [3]. Mackey, Mackey, and Tisseur extend this result to structures that are in the Lie or Jordan algebra ( $\mathbb{S} = \mathbb{L}$  or  $\mathbb{J}$ ) of a unitary and orthosymmetric scalar product  $\langle \cdot, \cdot \rangle_M$  on  $\mathbb{K}^n$  [14]. They show that when  $\hat{x}$  and  $b$  satisfy the conditions of Theorem 2.1 then

$$\mu_F(\hat{x}; \mathbb{S}) = \sqrt{2 \frac{\|r\|_2^2}{\|\hat{x}\|_2^2} - \alpha^2 \frac{|\langle r, \hat{x} \rangle_M|^2}{\|\hat{x}\|_2^4}},$$

where  $\alpha > 0$  is such that  $\alpha M$  is unitary. The optimal solution to (4.1) is given by

$$\Delta A_{\text{opt}} = \frac{r\hat{x}^*}{\hat{x}^*\hat{x}} + \epsilon \left( \frac{r\hat{x}^*}{\hat{x}^*\hat{x}} \right)^* \left( I - \frac{\hat{x}\hat{x}^*}{\hat{x}^*\hat{x}} \right), \quad \epsilon = \begin{cases} 1 & \text{if } \mathbb{S} = \mathbb{J}, \\ -1 & \text{if } \mathbb{S} = \mathbb{L}. \end{cases}$$

If  $\hat{x}, b$  do not satisfy the conditions of Theorem 2.1 then the backward error is infinite. Since  $\alpha |\langle \hat{x}, r \rangle_M| \leq \|\hat{x}\|_2 \|\alpha M r\|_2 = \|\hat{x}\|_2 \|r\|_2$ , and  $\mu_2(\hat{x}; \mathbb{S}) \leq \mu_F(\hat{x}; \mathbb{S})$  we obtain the following bounds on the structured backward error  $\mu_\nu(\hat{x}; \mathbb{S})$ .

**THEOREM 4.2.** *Let  $\mathbb{S}$  be the Lie algebra or Jordan algebra of an orthosymmetric and unitary scalar product. Let  $A \in \mathbb{S}$  be nonsingular and  $\hat{x} \in \mathbb{K}^n$  be an approximate nonzero solution to  $Ax = b$ . If  $\hat{x}$  and  $b$  satisfy the conditions in Theorem 2.1 then, for  $\nu = 2, F$ ,*

$$(4.6) \quad \mu_\nu(\hat{x}; \mathbb{S}) = c_\nu \mu_\nu(\hat{x}), \quad 1 \leq c_\nu \leq \sqrt{2},$$

otherwise  $\mu_\nu(\hat{x}; \mathbb{S}) = \infty$ .

Theorem 4.2 shows that forcing the backward error matrix to have the structure of  $\mathbb{S}$  has either little effect or a drastic effect on its norm. Not all structures require conditions on  $\hat{x}$  and  $b$  for  $\mu(\hat{x}; \mathbb{S})$  to be finite. For example, for symmetric, complex symmetric, pseudo-symmetric, persymmetric and Hamiltonian structures the structured backward error always exists and satisfies (4.6).

**4.2. Automorphism groups.** Structured backward errors are difficult to analyze for nonlinear structures. Some progress is made in this section for specific automorphism groups. For that we need a characterization of the set of all matrices in the automorphism group  $\mathbb{G}$  of an orthosymmetric scalar product mapping  $x$  to  $b$ .

**THEOREM 4.3.** [14, Thm. 4.7] *Let  $\mathbb{G}$  be the automorphism group of any orthosymmetric scalar product on  $\mathbb{K}^n$ . Then for any given nonzero  $x, b \in \mathbb{K}^n$  such that  $\langle x, x \rangle_M = \langle b, b \rangle_M = 1$ , we have*

$$\{A \in \mathbb{G} : Ax = b\} = \{S_{b,x} + Q(I - P_x) : Q \in \mathbb{G}, QP_x = P_bQ\},$$

where

$$S_{b,x} = \begin{cases} bx^T M & (\text{bilinear}), \\ bx^* M & (\text{sesquilinear}), \end{cases} \quad P_y = \begin{cases} yy^T M & (\text{bilinear}), \\ yy^* M & (\text{sesquilinear}). \end{cases}$$

Now assume that  $\mathbb{G}$  is the group of unitary matrices ( $M = I$ ). Let  $B \in \mathbb{G}$  be a particular solution to  $Ax = b$ , where  $x, b$  are given vectors of unit norms. Let  $X = [x, \tilde{X}]$  be unitary. Substituting  $P_x = xx^* = X \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} X^*$  into the constraint  $QP_x = P_b Q$  with  $P_b = bb^*$  and replacing  $b$  by  $Bx$  yields

$$QX \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} X^* = BX \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} X^* B^* Q$$

or, equivalently,

$$X^*(B^*Q)X \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} X^*(B^*Q)X$$

so that  $\tilde{X}^*(B^*Q)x = 0$  and  $x^*(B^*Q)\tilde{X} = 0$ . Hence

$$X^*B^*QX = \begin{bmatrix} e^{i\theta} & 0 \\ 0 & G \end{bmatrix} \iff Q = BX \begin{bmatrix} e^{i\theta} & 0 \\ 0 & G \end{bmatrix} X^*$$

for some  $(n-1) \times (n-1)$  unitary  $G$  and  $\theta \in \mathbb{R}$ . This latter result combined with Theorem 4.3 yields an explicit characterization of all the unitary matrices mapping  $x$  to  $b$ .

**COROLLARY 4.4.** *Let  $\mathbb{G}$  be the set of  $n \times n$  unitary matrices. Then for any nonzero  $x, b \in \mathbb{C}^n$  such that  $\|x\|_2 = \|b\|_2 = 1$ ,*

$$\{A \in \mathbb{G} : Ax = b\} = \left\{ bx^* + B\tilde{X}G\tilde{X}^* : G^*G = I_{n-1} \right\},$$

where  $B \in \mathbb{G}$  is a particular solution to  $Ax = b$  and  $[x, \tilde{X}]$  is unitary.

The following lemmas are also needed.

**LEMMA 4.5.** *Let  $A \in \mathbb{C}^{m \times m}$ ,  $X_1 \in \mathbb{C}^{m \times k}$  with  $m \geq k$  and  $X = [X_1, X_2]$  be unitary and let  $B \in \mathbb{C}^{k \times k}$ . Then*

$$\|X_1 B - AX_1\|_F^2 = \|B - X_1^* A X_1\|_F^2 + \|X_2^* A X_1\|_F^2.$$

**Proof.** The proof is immediate using  $X_1 X_1^* + X_2 X_2^* = I$  and

$$X_1 B - AX_1 = X \begin{bmatrix} B - X_1^* A X_1 \\ -X_2^* A X_1 \end{bmatrix}. \quad \square$$

**LEMMA 4.6.** [18, Lem. 2.4] *Let  $A \in \mathbb{C}^{m \times m}$  be unitary,  $X_1 \in \mathbb{C}^{m \times k}$  with  $2k \leq m$ ,  $X = [X_1, X_2]$  be unitary, and let  $H_1$  and  $H_2$  be the Hermitian polar factors of  $X_1^* A X_1$  and  $X_2^* A X_2$  respectively. Then, for any unitarily invariant norm,*

$$\|I - H_1\| = \|I - H_2\| \quad \text{and} \quad \|X_1^* A X_2\| = \|X_2^* A X_1\|.$$

Note that the results of Corollary 4.4 and Lemmas 4.5–4.6 hold for real orthogonal structures with  $\mathbb{C}$  and  $*$  replaced by  $\mathbb{R}$  and  $T$ , respectively. The next theorem shows that for orthogonal and unitary matrices, whenever the structured backward error is finite, it is a small factor of the unstructured one. This generalizes Sun's result on backward errors for the unitary eigenvalue problem, for which  $b$  has the special form  $b = e^{i\theta}x$ .

We have the following result.

**THEOREM 4.7.** *Let  $\mathbb{G} \subset \mathbb{K}^{n \times n}$  be the set of orthogonal matrices ( $\mathbb{K} = \mathbb{R}$ ) or unitary matrices ( $\mathbb{K} = \mathbb{C}$ ). Then for any nonzero approximate solution  $\hat{x} \in \mathbb{K}^n$  to  $Ax = b$ ,*

$$\mu_F(\hat{x}; \mathbb{G}) = \begin{cases} c_F \mu_F(\hat{x}) & \text{if } \|\hat{x}\|_2 = \|b\|_2, \text{ where } 1 \leq c_F \leq \sqrt{2}, \\ \infty & \text{otherwise.} \end{cases}$$

**Proof.** Let  $\Delta A$  be such that  $(A + \Delta A)\hat{x} = b$  and  $A + \Delta A \in \mathbb{G}$ . Without loss of generality, we can assume that  $\|b\|_2 = 1$ . If  $\|\hat{x}\|_2 \neq \|b\|_2$ , there is no solution to (4.1) and  $\mu_F(\hat{x}; \mathbb{G}) = \infty$ . If  $\|\hat{x}\|_2 = \|b\|_2$ , Corollary 4.4 implies that

$$A + \Delta A = b\hat{x}^* + B\tilde{X}G\tilde{X}^*,$$

where  $G, B$  and  $[\hat{x}, \tilde{X}]$  are all unitary,  $B$  being a particular solution to  $(A + \Delta A)\hat{x} = b$  (e.g.  $B$  is an Householder reflector mapping  $\hat{x}$  to  $b$ ). Then

$$\Delta A = [(b - A\hat{x}) (B\tilde{X}G - A\tilde{X})]X^*$$

so that  $\|\Delta A\|_F^2 = \|r\|_2^2 + \|B\tilde{X}G - A\tilde{X}\|_F^2$  and

$$(4.7) \quad \mu_F(\hat{x}; \mathbb{G})^2 = \|r\|_2^2 + \min_{G^*G=I} \|B\tilde{X}G - A\tilde{X}\|_F^2.$$

The minimization problem in (4.7) is a well-known Procrustes problem whose solution is given by [5, p. 149]

$$\min_{G^*G=I} \|B\tilde{X}G - A\tilde{X}\|_F = \|B\tilde{X}U - A\tilde{X}\|_F,$$

where  $U$  is the unitary factor of the polar decomposition of  $\tilde{X}^*B^*A\tilde{X}$ . Finally, by applying Lemma 4.5, then Lemma 4.6 and finally Lemma 4.5 again, we have

$$\begin{aligned} \|B\tilde{X}U - A\tilde{X}\|_F^2 &= \|\hat{x}^*B^*A\tilde{X}\|_F^2 + \|U - \tilde{X}^*B^*A\tilde{X}\|_F^2 \\ &= \|\tilde{X}^*B^*A\hat{x}\|_F^2 + \|e^{i\theta} - \hat{x}^*B^*A\hat{x}\|_F^2 \\ &= \|\hat{x}e^{i\theta} - B^*A\hat{x}\|_F^2, \end{aligned}$$

where  $\theta = \arg(\hat{x}^*B^*A\hat{x})$ . But

$$\|\hat{x}e^{i\theta} - B^*A\hat{x}\|_F = \min_{|\rho|=1} \|\hat{x}\rho - B^*A\hat{x}\|_F \leq \|\hat{x} - B^*A\hat{x}\|_F = \|b - A\hat{x}\|_2 = \|r\|_2. \quad \square$$

The unitary and orthogonal groups are not the only groups  $\mathbb{G}$  for which  $\mu_F(\hat{x}; \mathbb{G})$  can be rewritten in term of a Procrustes problem. Another example is the complex orthogonal group for which we can show using the characterization Theorem 4.3 that when  $\hat{x}, b \in \mathbb{C}^n$  are such that  $\hat{x}^T \hat{x} = b^T b = 1$ ,

$$\mu_F(\hat{x}; \mathbb{G})^2 = \|r\|_2^2 + \min \left\{ \|B\tilde{X}Q - A\tilde{X}\|_F^2 : Q^T Q = I, Q \in \mathbb{C}^{(n-1) \times (n-1)} \right\},$$

where  $B \in \mathbb{G}$  is a particular solution to  $B\hat{x} = b$  and  $[\hat{x}, \tilde{X}] =: X$  is complex orthogonal. We refer to [14, sec. 4.2.2] on how to construct a particular solution  $B$ ;  $X$  can be taken as the complex orthogonal  $\mathbb{G}$ -reflector mapping  $e_1$  to  $x$  [13]. Unlike the unitary and orthogonal case, the solution to this Procrustes problem is an open problem.

The important task of finding computable expressions for the structured backward error of nonlinearly structured linear systems is still largely open and remains to be addressed.

#### REFERENCES

- [1] J. R. BUNCH, J. W. DEMMEL, AND C. F. VAN LOAN, *The strong stability of algorithms for solving symmetric linear systems*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 494–499.
- [2] J. R. CARDOSO, C. S. KENNEY, AND F. S. LEITE, *Computing the square root and logarithm of a real  $P$ -orthogonal matrix*, Appl. Numer. Math., 46 (2003), pp. 173–196.
- [3] J. E. DENNIS, JR. AND J. J. MORÉ, *Quasi-Newton methods, motivation and theory*, SIAM Rev., 19 (1977), pp. 46–89.
- [4] K. FAN AND A. J. HOFFMAN, *Some metric inequalities in the space of matrices*, Proc. Amer. Math. Soc., 6 (1955), pp. 111–116.
- [5] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, USA, third ed., 1996.
- [6] D. J. HIGHAM, *Condition numbers and their condition numbers*, Linear Algebra Appl., 214 (1995), pp. 193–213.
- [7] D. J. HIGHAM AND N. J. HIGHAM, *Backward error and condition of structured linear systems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 162–175.
- [8] N. J. HIGHAM, *Matrix nearness problems and applications*, in Applications of Matrix Theory, M. J. C. Gover and S. Barnett, eds., Oxford University Press, 1989, pp. 1–27.
- [9] ———, *Accuracy and Stability of Numerical Algorithms*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, second ed., 2002.
- [10] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, 1985.
- [11] M. KAROW, D. KRESSNER, AND F. TISSEUR, *Structured eigenvalue condition numbers*, Numerical Analysis Report No. 467, Manchester Centre for Computational Mathematics, Manchester, England, Apr. 2005.
- [12] P. LANCASTER AND M. TISMENETSKY, *The Theory of Matrices*, Academic Press, London, second ed., 1985.
- [13] D. S. MACKEY, N. MACKEY, AND F. TISSEUR,  *$\mathbb{G}$ -reflectors: Analogues of Householder transformations in scalar product spaces*, Linear Algebra Appl., 385 (2004), pp. 187–213.
- [14] ———, *Structured mapping problems for automorphism groups, Lie algebras and Jordan algebras associated with scalar products*, numerical analysis report, Manchester Centre for Computational Mathematics, Manchester, England, 2004. In preparation.
- [15] ———, *Structured factorizations in scalar product spaces*, To appear in SIAM J. Matrix Anal. Appl., (2005).
- [16] S. M. RUMP, *Structured perturbations part I: Normwise distances*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 1–30.
- [17] J. SUN, *On optimal backward perturbation bounds for linear systems and linear least squares problems*, Report UMINF 96-15, Department of Computing Science, University of Umeå, Sweden, Dec. 1996.

- [18] ———, *Backward errors for the unitary eigenproblem*, Tech. Report UMINF-97.25, Department of Computing Science, University of Umeå, Umeå, Sweden, 1997.