Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

STRUCTURED CONDITIONING OF MATRIX FUNCTIONS*

PHILIP I. DAVIES†

Abstract. The existing theory of conditioning for matrix functions $f(X)\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ does not cater for structure in the matrix X. An extension of this theory is presented in which when X has structure, all perturbations of X are required to have the same structure. Two classes of structured matrices are considered, those comprising the Jordan algebra $\mathbb{J}$ and the Lie algebra $\mathbb{L}$ associated with a nondegenerate bilinear or sesquilinear form on $\mathbb{R}^n$ or $\mathbb{C}^n$. Examples of such classes are the symmetric, skew-symmetric, Hamiltonian and skew-Hamiltonian matrices. Structured condition numbers are defined for these two classes. Under certain conditions on the underlying scalar product, explicit representations are given for the structured condition numbers. Comparisons between the unstructured and structured condition numbers are then made. When the underlying scalar product is a sesquilinear form, it is shown that there is no difference between the values of the two condition numbers for (i) all functions of $X \in \mathbb{J}$, and (ii) odd and even functions of $X \in \mathbb{L}$. When the underlying scalar product is a bilinear form then equality is not guaranteed in all these cases. Where equality is not guaranteed, bounds are obtained for the ratio of the unstructured and structured condition numbers.

Key words. Matrix functions, Fréchet derivative, Condition numbers, Bilinear forms, Sesquilinear forms, Structured Matrices, Jordan algebra, Symmetric matrices, Lie algebra, Skew-symmetric matrices.

AMS subject classifications. 15A57, 15A63, 65F30, 65F35.

1. Introduction.

A theory of conditioning for matrix functions was developed by Kenney and Laub [6]. Condition numbers of $f(X)$ are obtained in terms of the norm of the Fréchet derivative of the function at X. In this work a function $f(X)$ of a matrix $X \in \mathbb{C}^{n\times n}$ has the usual meaning, which can be defined in terms of a Cauchy integral formula, a Hermite interpolating polynomial, or the Jordan canonical form. It is assumed throughout that f is defined on the spectrum of X. A large body of theory on matrix functions exists, with a comprehensive treatment available in [5].

In this work we extend the ideas of Kenney and Laub to structured matrices. That is, when X has structure then all perturbations of X are required to have the same structure. Enforcing structure on the perturbations enables the theory to respect the underlying physical problem. The structured matrices considered in this work arise in the context of nondegenerate bilinear or sesquilinear forms on $\mathbb{R}^n$ or $\mathbb{C}^n$. This allows a wide variety of structured matrices to be considered. Examples of such classes of structured matrices are the symmetric, skew-symmetric, Hamiltonian and skew-Hamiltonian matrices.

In section 2 we review the original theory of conditioning of matrix functions [6] and discuss the unstructured condition number, $K(f, X)$. In section 3 we briefly

*Received by the editors 19 January, 2004. Accepted for publication 11 May, 2004. Handling Editor: Ludwig Elsner. Numerical Analysis Report No. 441, Manchester Centre for Computational Mathematics, Manchester, England, Jan 2004.

†Department of Mathematics, University of Manchester, Manchester, M13 9PL, England (`pdavies@ma.man.ac.uk, http://www.maths.man.ac.uk/~ieuan`). This work was supported by Engineering and Physical Sciences Research Council grant GR/R22612.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

review definitions and some relevant properties of bilinear forms, $\langle x, y\rangle = x^T My$ and sesquilinear forms $\langle x, y\rangle = x^* My$. We introduce some important classes of structured matrices associated with a bilinear or sesquilinear form. For two of these classes, the Jordan and Lie algebras, denoted by $\mathbb{J}$ and $\mathbb{L}$ respectively, we define structured condition numbers, $K_{\mathbb{J}}(f, X)$ and $K_{\mathbb{L}}(f, X)$. Then, assuming that $M \in \mathbb{R}^{n\times n}$ satisfies $M = \pm M^T$ and $M^T M = I$, we give an explicit representation for these condition numbers. If the underlying scalar product is a bilinear form, then the structured condition numbers are equal to the 2-norm of a matrix. We then present an algorithm, based on the power method, for approximating these structured condition numbers. Numerical examples are given to show that after a few cycles of our algorithm, reliable estimates can be obtained.

In section 4, using the explicit representation of $K_{\mathbb{J}}(f, X)$, we compare the unstructured and the structured condition number when $X \in \mathbb{J}$. We first consider the case where $\mathbb{J}$ is the Jordan algebra associated with a bilinear form. Experimental and theoretical evidence is used to show that $K_{\mathbb{J}}(f, X)$ and $K(f, X)$ are often equal to each other. However, examples can be found where the unstructured condition number is larger than the structured condition number. A bound for $K(f, X)/K_{\mathbb{J}}(f, X)$ is then given which is linear in n. We then investigate the class of real symmetric matrices, which is the Jordan algebra associated with the bilinear form $\langle x, y\rangle = x^T y$. We show that $K(f, X) = K_{\mathbb{J}}(f, X)$ for all symmetric X. We also show that when $\mathbb{J}$ is the Jordan algebra associated with a sesquilinear form, $K(f, X) = K_{\mathbb{J}}(f, X)$ for all $X \in \mathbb{J}$.

In section 5, using the explicit representation of $K_{\mathbb{L}}(f, X)$, we compare the unstructured and the structured condition number when $X \in \mathbb{L}$. We first consider the case where $\mathbb{L}$ is the Lie algebra associated with a bilinear form. For general f, we have been unable to show anything about the relationship between the two condition numbers. However, progress can be made if we restrict f to odd or even functions. For odd f, we show that the ratio between the unstructured and the structured condition number is unbounded. However, for even f, we are able to show that $K(f, X) = K_{\mathbb{L}}(f, X)$ for all $X \in \mathbb{L}$. We then investigate the class of real skew-symmetric matrices, which is the Lie algebra associated with the bilinear form $\langle x, y\rangle = x^T y$. For the exponential, cosine and sine functions we show that $K(f, X) = K_{\mathbb{L}}(f, X)$ for all skew-symmetric X. We then consider the case where $\mathbb{L}$ is the Lie algebra associated with a sesquilinear form. We show that $K(f, X) = K_{\mathbb{L}}(f, X)$ for odd and even functions f at all $X \in \mathbb{L}$.

Finally, in section 6 we give our conclusions and some suggestions for future work. Table 6.1 gives a summary of the main results of this paper.

2. Conditioning of Matrix Functions. Using the theory developed by Kenney and Laub [6] we start by considering the effect of general perturbations of X at f. The function f is Fréchet differentiable at X if and only if there exists a bounded linear operator $L(\,\cdot\,, X)\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ such that for all $Z \in \mathbb{C}^{n\times n}$ where $\|Z\| = 1$,

$$\lim_{\delta\to 0} \left\| \frac{f(X + \delta Z) - f(X)}{\delta} - L(Z, X) \right\| = 0. \tag{2.1}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

The operator L is known as the Fréchet derivative of f at X. The unstructured condition number of the matrix function is then defined using the Fréchet derivative:

$$K(f,X) = \|L(\,\cdot\,,X)\|_F = \max_{Z\neq 0} \frac{\|L(Z,X)\|_F}{\|Z\|_F}. \tag{2.2}$$

Just as in [6], the Frobenius norm has been used because of its nice properties with respect to the Kronecker matrix product. The condition number defined in (2.2) relates the absolute errors of f at X. An alternative condition number is described in [1] that relates the relative errors of f at X:

$$k(f,X) = \frac{\|L(\,\cdot\,,X)\|_F \|X\|_F}{\|f(X)\|_F}.$$

These two condition numbers are closely related:

$$k(f,X) = K(f,X)\frac{\|X\|_F}{\|f(X)\|_F}. \tag{2.3}$$

The focus of this paper shall be on the "absolute" condition number, $K(f,X)$. Results for the "relative" condition number, $k(f,X)$, can then be obtained using (2.3). We shall assume throughout that $f(X)$ can be expressed as a power series,

$$f(X) = \sum_{m=0}^{\infty} \alpha_m X^m, \tag{2.4}$$

where $\alpha_m \in \mathbb{R}$ and the equivalent scalar power series $f(x) = \sum_{m=0}^{\infty} \alpha_m x^m$ is absolutely convergent for all $|x| < r$ where $\|X\|_2 < r$. This assumption encompasses a wide range of functions, including functions such as the exponential, trigonometric and hyperbolic functions, whose Taylor series have an infinite radius of convergence. We shall also assume that $\delta > 0$, $\|Z\|_2 \leq 1$ and $\|X\|_2 + \delta < r$, so that $f(X+\delta Z)$ is well defined in terms of the power series in (2.4). Then

$$f(X+\delta Z) = f(X) + \delta \sum_{m=1}^{\infty} \alpha_m \sum_{k=0}^{m-1} X^k Z X^{m-1-k} + O(\delta^2). \tag{2.5}$$

Using (2.5) together with (2.1), an explicit representation can be given for the Fréchet derivative:

$$L(Z,X) = \sum_{m=1}^{\infty} \alpha_m \sum_{k=0}^{m-1} X^k Z X^{m-1-k}.$$

Applying the vec operator (which forms a vector by stacking the columns of a matrix) to $L(Z,X)$ and using the relation $\mathrm{vec}(AXB) = (B^T \otimes A)\mathrm{vec}(X)$, we obtain

$$\mathrm{vec}(L(Z,X)) = D(X)\mathrm{vec}(Z),$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

where $D(X)$ is the Kronecker form of the Fréchet derivative:

$$D(X) = \sum_{m=1}^{\infty} \alpha_m \sum_{k=0}^{m-1} (X^T)^{m-1-k} \otimes X^k. \tag{2.6}$$

As $\|X\|_F = \|\mathrm{vec}(X)\|_2$ for all X, we have

$$K(f, X) = \max_{Z \neq 0} \frac{\|D(X)\mathrm{vec}(Z)\|_2}{\|\mathrm{vec}(Z)\|_2} = \|D(X)\|_2. \tag{2.7}$$

We shall now consider the effect on the condition number when structure is imposed on X and the perturbed matrix $X + \delta Z$.

3. Structured Matrices and Condition Numbers. In [8], Mackey, Mackey and Tisseur define classes of structured matrices that arise in the context of nondegenerate bilinear and sesquilinear forms. We shall briefly review the definitions and properties of such forms.

Consider a map $(x, y) \to \langle x, y\rangle$ from $\mathbb{K}^n \times \mathbb{K}^n$ to $\mathbb{K}$, where $\mathbb{K}$ denotes the field $\mathbb{R}$ or $\mathbb{C}$. If the map is linear in both arguments x and y, that is,

$$\langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle = \alpha_1\langle x_1, y\rangle + \alpha_2\langle x_2, y\rangle,$$
$$\langle x, \beta_1 y_1 + \beta_2 y_2\rangle = \beta_1\langle x, y_1\rangle + \beta_2\langle x, y_2\rangle,$$

then this map is called a bilinear form. If $\mathbb{K} = \mathbb{C}$ and the map $\langle x, y\rangle$ is conjugate linear in the first argument and linear in the second, that is,

$$\langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle = \alpha_1^*\langle x_1, y\rangle + \alpha_2^*\langle x_2, y\rangle,$$
$$\langle x, \beta_1 y_1 + \beta_2 y_2\rangle = \beta_1\langle x, y_1\rangle + \beta_2\langle x, y_2\rangle,$$

then this map is called a sesquilinear form.

For each bilinear form on $\mathbb{K}^n$, there exists a unique $M \in \mathbb{K}^{n\times n}$ such that $\langle x, y\rangle = x^T My, \forall x, y \in \mathbb{K}^n$. Similarly, for each sesquilinear form on $\mathbb{C}^n$, there exists a unique $M \in \mathbb{C}^{n\times n}$ such that $\langle x, y\rangle = x^*My, \forall x, y \in \mathbb{C}^n$. A bilinear or sesquilinear form is nondegenerate if

$$\langle x, y\rangle = 0,\ \forall y \Rightarrow x = 0,$$
$$\langle x, y\rangle = 0,\ \forall x \Rightarrow y = 0.$$

It can be shown that a bilinear or sesquilinear form is nondegenerate if and only if M is nonsingular. We shall use the term *scalar product* to refer to a nondegenerate bilinear or sesquilinear form on $\mathbb{K}^n$.

A bilinear form is said to be symmetric if $\langle x, y\rangle = \langle y, x\rangle$ and skew-symmetric if $\langle x, y\rangle = -\langle y, x\rangle$ for all $x, y \in \mathbb{K}^n$. It can easily be shown that the matrix M associated with these forms are symmetric and skew-symmetric respectively. Similarly, a sesquilinear form is Hermitian if $\langle x, y\rangle = \langle y, x\rangle^*$ or skew-Hermitian if $\langle x, y\rangle = -\langle y, x\rangle^*$ for all $x, y \in \mathbb{C}^n$. The matrices associated with these forms are Hermitian and skew-Hermitian respectively. Symmetric, skew-symmetric, Hermitian and skew-Hermitian forms will be the main focus of this paper.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

TABLE 3.1
Sampling of structured matrices where $J = \begin{bmatrix} & I_n \\ -I_n & \end{bmatrix}$, $\Sigma_{p,q} = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix}$ *and* $R = \mathrm{antidiag}(1,\dots,1)$.

Space	**Bilinear Form** $\langle x,y\rangle$	**Automorphism Gp.** $\{G : G^TMG = M\}$	**Jordan Algebra** $\{S : S^TM = MS\}$	**Lie Algebra** $\{K : K^TM = -MK\}$
$\mathbb{R}^n$	x^Ty symmetric	Real orthogonals $O(n,\mathbb{R})$	Real Symmetrics	Real Skew-symmetrics
$\mathbb{C}^n$	x^Ty symmetric	Complex orthogonals $O(n,\mathbb{C})$	Complex symmetrics	Complex skew-symmetrics
$\mathbb{R}^{2n}$	x^TJy skew-symm.	Real symplectics $Sp(2n,\mathbb{R})$	Skew-Hamiltonians	Hamiltonians
$\mathbb{C}^{2n}$	x^TJy skew-symm.	Complex symplectics $Sp(2n,\mathbb{C})$	J-skew-symmetric	J-symmetric
$\mathbb{R}^n$	x^TRy symmetric	Real perplectics $\mathcal{P}(n)$	Persymmetrics	Perskew-symmetrics
$\mathbb{R}^n$	$x^T\Sigma_{p,q}y$ symmetric	Pseudo-orthogonals $O(p,q)$	Pseudo symmetrics	Pseudo skew-symmetrics

Space	**Sesquilinear Form** $\langle x,y\rangle$	**Automorphism Gp.** $\{G : G^*MG = M\}$	**Jordan Algebra** $\{S : S^*M = MS\}$	**Lie Algebra** $\{K : K^*M = -MK\}$
$\mathbb{C}^n$	x^*y Hermitian	Unitaries $U(n)$	Hermitian	Skew-Hermitian
$\mathbb{C}^n$	$x^*\Sigma_{p,q}y$ Hermitian	Pseudo-unitaries $U(p,q)$	Pseudo Hermitian	Pseudo skew-Hermitian
$\mathbb{C}^{2n}$	x^*Jy skew-Herm.	Conjugate symplectics $Sp^*(2n,\mathbb{C})$	J-skew-Hermitian	J-Hermitian

Three important classes of structured matrices are associated with each scalar product.

1. The matrices G which preserve the value of the scalar product, that is,
$$\langle Gx, Gy\rangle = \langle x,y\rangle \quad \forall x,y \in \mathbb{K}^n.$$
The set $\mathbb{G}$, known as the automorphism group of the scalar product, is thus defined as
$$\mathbb{G} \stackrel{\text{def}}{=} \{G \in \mathbb{K}^{n\times n} : G^TMG = M\} \quad \text{for a bilinear form,}$$
$$\mathbb{G} \stackrel{\text{def}}{=} \{G \in \mathbb{C}^{n\times n} : G^*MG = M\} \quad \text{for a sesquilinear form.}$$
2. The matrices S that are *self-adjoint* with respect to the scalar product, that is,
$$\langle Sx, y\rangle = \langle x, Sy\rangle \quad \forall x,y \in \mathbb{K}^n.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

The set $\mathbb{J}$, known as the Jordan algebra related to the scalar product, is thus defined as

$$\mathbb{J} \stackrel{\text{def}}{=} \{S \in \mathbb{K}^{n\times n}\colon S^T M = MS\} \quad \text{for a bilinear form,}$$
$$\mathbb{J} \stackrel{\text{def}}{=} \{S \in \mathbb{C}^{n\times n}\colon S^* M = MS\} \quad \text{for a sesquilinear form.}$$

3. The matrices K that are *skew-adjoint* with respect to the scalar product, that is,

$$\langle Kx, y\rangle = -\langle x, Ky\rangle \quad \forall x, y \in \mathbb{K}^n.$$

The set $\mathbb{L}$, known as the Lie algebra related to the scalar product, is thus defined as

$$\mathbb{L} \stackrel{\text{def}}{=} \{K \in \mathbb{K}^{n\times n}\colon K^T M = -MK\} \quad \text{for a bilinear form,}$$
$$\mathbb{L} \stackrel{\text{def}}{=} \{K \in \mathbb{C}^{n\times n}\colon K^* M = -MK\} \quad \text{for a sesquilinear form.}$$

While $\mathbb{G}$ is a multiplicative group, it is not a linear subspace. However, $\mathbb{J}$ and $\mathbb{L}$ do form linear subspaces. This means that if X and the perturbed matrix $X + \delta Z$ belong to $\mathbb{J}$ (or $\mathbb{L}$), then the perturbation matrix Z must also belong to $\mathbb{J}$ (or $\mathbb{L}$). Because of this linear property, the rest of this paper focuses only on matrices in $\mathbb{J}$ and $\mathbb{L}$. Table 3.1 shows some examples of well-known structured matrices associated with a scalar product.

The structured condition numbers are defined in a similar manner to the unstructured condition number given in (2.2), except the perturbation matrix Z is restricted to either $\mathbb{J}$ or $\mathbb{L}$. Therefore, we define

$$K_{\mathbb{J}}(f, X) = \max_{Z\neq 0, Z\in\mathbb{J}} \frac{\|L(Z, X)\|_F}{\|Z\|_F},$$
$$K_{\mathbb{L}}(f, X) = \max_{Z\neq 0, Z\in\mathbb{L}} \frac{\|L(Z, X)\|_F}{\|Z\|_F},$$

where $L(Z, X)$ is defined in (2.1). Notice that no structure has been assumed on the matrix X. Imposing a similar structure on X will be considered in sections 4 and 5. We can also define "relative" structured condition numbers

$$k_{\mathbb{J}}(f, X) = K_{\mathbb{J}}(f, X)\frac{\|X\|_F}{\|f(X)\|_F},$$
$$k_{\mathbb{L}}(f, X) = K_{\mathbb{L}}(f, X)\frac{\|X\|_F}{\|f(X)\|_F}.$$

Notice that

$$\frac{k(f, X)}{k_{\mathbb{J}}(f, X)} = \frac{K(f, X)}{K_{\mathbb{J}}(f, X)} \quad \text{and} \quad \frac{k(f, X)}{k_{\mathbb{L}}(f, X)} = \frac{K(f, X)}{K_{\mathbb{L}}(f, X)}.$$

In sections 4 and 5 we shall consider the ratios $\frac{K(f,X)}{K_{\mathbb{J}}(f,X)}$ and $\frac{K(f,X)}{K_{\mathbb{L}}(f,X)}$. Identical results can then be shown to hold for the ratios of the "relative" condition numbers.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

We shall now consider the structured condition numbers when the underlying scalar product is a bilinear form in section 3.1, and when the scalar product is a sesquilinear form in section 3.2.

3.1. Bilinear forms. In this section we shall assume that $\mathbb{J}$ and $\mathbb{L}$ denote a Jordan algebra and a Lie algebra associated with a nondegenerate bilinear form on $\mathbb{K}^n$. For $S \in \mathbb{J}$, it can be shown that $\mathrm{vec}(S)$ satisfies

$$\big((M^T \otimes I)P - I \otimes M\big)\,\mathrm{vec}(S) = 0, \tag{3.1}$$

where P is the vec-permutation matrix that satisfies $P\mathrm{vec}(X) = \mathrm{vec}(X^T)$ for all $X \in \mathbb{C}^{n\times n}$. From (3.1) we see that $\mathrm{vec}(S)$ is contained in the null space of $(M^T \otimes I)P - I \otimes M$. Therefore the structured condition number of f at X can be expressed as

$$K_{\mathbb{J}}(f,X) = \max_{Z\neq 0, Z\in\mathbb{J}} \frac{\|D(X)\mathrm{vec}(Z)\|_2}{\|\mathrm{vec}(Z)\|_2} = \|D(X)B\|_2, \tag{3.2}$$

where the columns of B form an orthonormal basis for the null space of $(M^T \otimes I)P - I \otimes M$. For $K \in \mathbb{L}$, $\mathrm{vec}(K)$ satisfies

$$\big((M^T \otimes I)P + I \otimes M\big)\,\mathrm{vec}(K) = 0,$$

and the structured condition number of f at X can be given as

$$K_{\mathbb{L}}(f,X) = \max_{Z\neq 0, Z\in\mathbb{L}} \frac{\|D(X)\mathrm{vec}(Z)\|_2}{\|\mathrm{vec}(Z)\|_2} = \|D(X)B\|_2,$$

where the columns of B form an orthonormal basis for the null space of $(M^T \otimes I)P + I \otimes M$. If M has certain properties, then more can be said about the null space of $(M^T \otimes I)P \pm I \otimes M$.

LEMMA 3.1. *Let $M \in \mathbb{R}^{n\times n}$ be nonsingular and $M = \delta M^T$ where $\delta = \pm 1$. Define*

$$S_{\mathbb{J}}(M) = (M^T \otimes I)P - I \otimes M, \tag{3.3}$$

$$S_{\mathbb{L}}(M) = (M^T \otimes I)P + I \otimes M, \tag{3.4}$$

where P is the vec-permutation matrix. Then

$$\mathrm{rank}(S_{\mathbb{J}}(M)) = n(n-\delta)/2,$$
$$\mathrm{rank}(S_{\mathbb{L}}(M)) = n(n+\delta)/2.$$

Furthermore, if $M^TM = I$, then the nonzero singular values of $S_{\mathbb{J}}(M)$ and $S_{\mathbb{L}}(M)$ are equal to 2.

Proof. Using the fact that $P(A\otimes B) = (B \otimes A)P$ for all $A, B \in \mathbb{C}^{n\times n}$ [2] we can write

$$S_{\mathbb{J}}(M) = (\delta P - I)(I \otimes M).$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

The vec-permutation matrix P is symmetric and has eigenvalues 1 and -1 with multiplicities $\frac{1}{2}n(n+1)$ and $\frac{1}{2}n(n-1)$, respectively [2]. Since the matrix $I \otimes M$ has full rank when M is nonsingular,

$$\operatorname{rank}(S_{\mathbb{J}}(M)) = \operatorname{rank}(\delta P - I) = n(n-\delta)/2.$$

A similar argument shows that $\operatorname{rank}(S_{\mathbb{L}}(M)) = \operatorname{rank}(\delta P + I) = \frac{1}{2}n(n+\delta)$.

When $M^T M = I$ the matrix $I \otimes M$ is orthogonal. Therefore, the matrix $\delta P - I$ and $S_{\mathbb{J}}(M)$ have the same singular values. A similar argument shows that $\delta P + I$ and $S_{\mathbb{L}}(M)$ have the same singular values. ∎

Lemma 3.2. *Let $M \in \mathbb{R}^{n\times n}$ be nonsingular and $M = \pm M^T$. Then*

$$\operatorname{null}(S_{\mathbb{J}}(M)) = \operatorname{range}(PS_{\mathbb{L}}(M)^T),$$
$$\operatorname{null}(S_{\mathbb{L}}(M)) = \operatorname{range}(PS_{\mathbb{J}}(M)^T),$$

where P is the vec-permutation matrix and $S_{\mathbb{J}}(M), S_{\mathbb{L}}(M)$ are defined in (3.3) *and* (3.4)*.*

Proof. As a consequence of Lemma 3.1

$$\dim(\operatorname{null}(S_{\mathbb{J}}(M))) = \dim(\operatorname{range}(PS_{\mathbb{L}}(M)^T)),$$
$$\dim(\operatorname{null}(S_{\mathbb{L}}(M))) = \dim(\operatorname{range}(PS_{\mathbb{J}}(M)^T)),$$

and hence all we need to show is that $S_{\mathbb{J}}(M)PS_{\mathbb{L}}(M)^T = 0$. Now

$$S_{\mathbb{J}}(M)PS_{\mathbb{L}}(M)^T = (M^T \otimes I)P(M \otimes I) - (I \otimes M)P(I \otimes M^T) + M^T \otimes M^T - M \otimes M.$$

As $M = \pm M^T$ the third and fourth terms cancel. Further rearranging yields

$$S_{\mathbb{J}}(M)PS_{\mathbb{L}}(M)^T = P(M \otimes M^T) - P(M \otimes M^T) = 0. \qquad ∎$$

Theorem 3.3. *Let $X \in \mathbb{C}^{n\times n}$ and the scalar function $f(x) = \sum_{m=0}^{\infty} \alpha_m x^m$ be absolutely convergent for all $|x| < r$ where $\|X\|_2 < r$. Also, let $\mathbb{J}$ and $\mathbb{L}$ denote a Jordan algebra and a Lie algebra associated with a nondegenerate bilinear form $\langle x, y\rangle = x^T M y$, where $M \in \mathbb{R}^{n\times n}$ satisfies $M^T M = I$ and $M = \delta M^T$ where $\delta = \pm 1$. Then*

$$K_{\mathbb{J}}(f, X) = \frac{1}{2}\|D(X)PS_{\mathbb{L}}(M)^T\|_2, \tag{3.5}$$

$$K_{\mathbb{L}}(f, X) = \frac{1}{2}\|D(X)PS_{\mathbb{J}}(M)^T\|_2. \tag{3.6}$$

Proof. First, we consider $Z \in \mathbb{J}$. We have shown that $\operatorname{vec}(Z) \in \operatorname{null}(S_{\mathbb{J}}(M))$ and by Lemma 3.2 that $\operatorname{vec}(Z) \in \operatorname{range}(PS_{\mathbb{L}}(M)^T)$. As $M^T M = I$, Lemma 3.1 shows that there exist orthogonal matrices U and V such that

$$PS_{\mathbb{L}}(M)^T = U \begin{bmatrix} 2I & \\ & 0 \end{bmatrix} V^T.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

Let $r = \text{rank}(PS_{\mathbb{L}}(M)^T) = \frac{1}{2}n(n+\delta)$ and partition

$$U = \begin{array}{c} \begin{array}{cc} \scriptstyle r & \scriptstyle n^2-r \end{array} \\ \begin{pmatrix} U_1 & U_2 \end{pmatrix} \end{array}.$$

Then we can see that $\text{range}(PS_{\mathbb{L}}(M)^T) = \text{range}(U_1)$ and $U_1^T U_1 = I$. Therefore the structured condition number of f at X is

$$K_{\mathbb{J}}(f,X) = \max_{Z \neq 0, Z \in \mathbb{J}.} \frac{\|D(X)\text{vec}(Z)\|_2}{\|\text{vec}(Z)\|_2} = \|D(X)U_1\|_2.$$

Define

$$B = \begin{array}{c} \begin{array}{cc} \scriptstyle r & \scriptstyle n^2-r \end{array} \\ \begin{pmatrix} U_1 & 0 \end{pmatrix} \end{array} = U \begin{bmatrix} I & \\ & 0 \end{bmatrix}.$$

Then

$$\begin{aligned} K_{\mathbb{J}}(f,X) = \|D(X)B\|_2 &= \frac{1}{2}\|D(X)PS_{\mathbb{L}}(M)^T V\|_2, \\ &= \frac{1}{2}\|D(X)PS_{\mathbb{L}}(M)^T\|_2. \end{aligned}$$

A similar argument can be used to give the result for $K_{\mathbb{L}}(f,X)$. □

In Theorem 3.3 we have assumed that $M = \pm M^T$ and $M^T M = I$. This may seem a restrictive condition. However, a wide variety of structured matrices are associated with bilinear forms that satisfy these conditions, including all those in Table 3.1.

By showing that $S_{\mathbb{J}}(M)S_{\mathbb{L}}(M)^T = 0$ when $M = \pm M^T$ and $M^T M = I$, we can show, using an almost identical proof of Lemma 3.2, that

$$\begin{aligned} \text{null}(S_{\mathbb{J}}(M)) &= \text{range}(S_{\mathbb{L}}(M)^T), \\ \text{null}(S_{\mathbb{L}}(M)) &= \text{range}(S_{\mathbb{J}}(M)^T). \end{aligned}$$

Hence, using an almost identical proof of Theorem 3.3, we can show

$$\begin{aligned} K_{\mathbb{J}}(f,X) &= \frac{1}{2}\|D(X)S_{\mathbb{L}}(M)^T\|_2, \\ K_{\mathbb{L}}(f,X) &= \frac{1}{2}\|D(X)S_{\mathbb{J}}(M)^T\|_2. \end{aligned}$$

We shall be using the forms (3.5) and (3.6) given in Theorem 3.3 as these involve less algebraic manipulation later.

3.1.1. Condition estimation. Kenney and Laub [6] presented a method for estimating the condition number $K(f,X)$ by using the power method. The power method can be used to approximate $\|A\|_2$ for a matrix $A \in \mathbb{C}^{m\times n}$. This method starts with a vector $z \in \mathbb{C}^n$ and iterates the following cycle:

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

for $i = 1, 2, \dots$
  $z = z/\|z\|_2$, then compute $w = Az$.
  $w = w/\|w\|_2$, then compute $z = A^T w$.
end
$\|A\|_2 \approx \|z\|_2$

As long as the starting vector is not orthogonal to the singular subspace corresponding to the largest singular value of A, $\|z\|_2$ converges to $\|A\|_2$. The power method can then be used to approximate $\|D(X)\|_2$. As forming $D(X)$ can be prohibitively expensive, it is difficult to compute $D(X)z_0$, where $z_0 = \mathrm{vec}(Z_0)$ for some $Z_0 \in \mathbb{C}^{n\times n}$. Therefore we can use the "finite difference" relation

$$\mathrm{unvec}(D(X)\mathrm{vec}(Z_0)) = \frac{1}{\delta}\left(f(X+\delta Z_0) - f(X)\right) + O(\delta) \tag{3.7}$$

where $\|Z_0\|_F = 1$. An approximation to $D(X)\mathrm{vec}(Z_0)$ can be formed using a sufficiently small δ. Starting with $\|Z_0\|_F = 1$, the two steps of the power method can then be approximated by

$$W = \frac{1}{\delta}\left(f(X+\delta Z_0) - f(X)\right),$$
$$Z_1 = \frac{1}{\delta}\left(f(X^T+\delta W_0) - f(X^T)\right),$$

where $W_0 = W/\|W\|_F$. Then $\|Z_1\|_F \approx \|D(X)\|_2$. More accurate estimates can be obtained by repeating the cycle with $Z_0 = Z_1/\|Z_1\|_F$.

Now we consider how to estimate our structured condition number $K_{\mathbb{J}}(f,X)$ by using the power method to estimate $\frac{1}{2}\|D(X)PS_{\mathbb{L}}(M)^T\|_2$. Starting with Z_0 where $\|Z_0\|_F = 1$, let $y = \mathrm{vec}(Y) = \frac{1}{2}PS_{\mathbb{L}}(M)^T\mathrm{vec}(Z_0)$. Then

$$Y = \frac{1}{2}\left(Z_0M^T + Z_0^TM\right). \tag{3.8}$$

Let $w = \mathrm{vec}(W) = D(X)y$. Then we can approximate W using

$$W \approx \frac{1}{\delta}\left(f(X+\delta Y) - f(X)\right).$$

The next stage is to scale w such that

$$w_0 = \mathrm{vec}(W_0) = \mathrm{vec}(W/\|W\|_F) = w/\|w\|_2.$$

The final step is to compute $z_1 = \mathrm{vec}(Z_1) = \frac{1}{2}(D(X)PS_{\mathbb{L}}(M)^T)^Tw_0$. Rearranging we get $z_1 = \frac{1}{2}S_{\mathbb{L}}(M)Pu$ where $u = \mathrm{vec}(U) = D(X)^Tw_0$. Therefore we can approximate U using

$$U \approx \frac{1}{\delta}\left(f(X^T+\delta W_0) - f(X^T)\right).$$

We can then form z_1 by

$$Z_1 = \frac{1}{2}\left(UM + MU^T\right). \tag{3.9}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

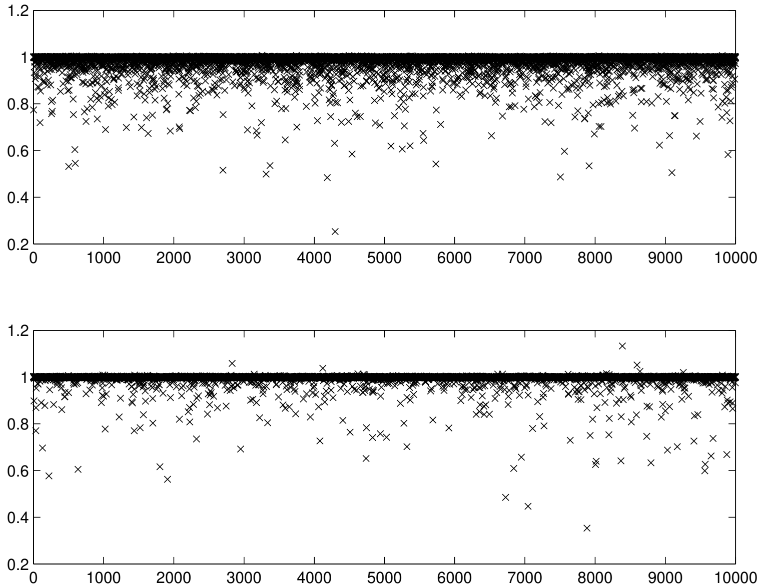

FIG. 3.1. *Accuracy of estimates for structured condition numbers. (top) Measures* $K_{\mathbb{J}}^{\mathrm{est}}(f,X)/K_{\mathbb{J}}(f,X)$ *for the 10000 examples in Experiment* 4.2. *(bottom) Measures* $K_{\mathbb{L}}^{\mathrm{est}}(f,X)/K_{\mathbb{L}}(f,X)$ *for the 10000 examples in Experiment* 5.1.

Then $\|Z_1\|_F \approx \frac{1}{2}\|D(X)PS_{\mathbb{L}}(M)^T\|_2$. More accurate estimates can be obtained by repeating the cycle with $Z_0 = Z_1/\|Z_1\|_F$. To estimate $K_{\mathbb{L}}(f,X)$, we have exactly the same procedure, except with the $+$ sign in (3.8) and (3.9) changed to a $-$ sign.

ALGORITHM 3.4 (Estimation of structured condition numbers). *Given* $X \in \mathbb{K}^{n\times n}$ *and the scalar function* $f(x) = \sum_{m=0}^{\infty} \alpha_m x^m$ *for which the series is absolutely convergent for all* $|x| < r$ *where* $\|X\|_2 < r$*, this algorithm computes an approximation to the condition number* $K_{\mathbb{J}}(f,X)$ *or* $K_{\mathbb{L}}(f,X)$*.*

If we are computing $K_{\mathbb{J}}(f,X)$ then $k=0$. Otherwise $k=1$.
Let $\delta = 100u\|X\|_F$ where u is the unit roundoff.
Compute $f(X)$.
Choose random nonzero $Z \in \mathbb{K}^{n\times n}$.
for $i = 1, 2, \ldots$
  if $\|Z\|_F \neq 0$, $Z = Z/\|Z\|_F$, end
  $Y = \frac{1}{2}(ZM^T + (-1)^k Z^T M)$.
  $W = \frac{1}{\delta}(f(X+\delta Y) - f(X))$.
  if $\|W\|_F \neq 0$, $W = W/\|W\|_F$, end
  $U = \frac{1}{\delta}(f(X^T + \delta W) - f(X)^T)$.
  $Z = \frac{1}{2}(UM + (-1)^k MU^T)$.
end
$K_{\mathbb{J}}^{\mathrm{est}}(f,X)$ or $K_{\mathbb{L}}^{\mathrm{est}}(f,X) = \|Z\|_F$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

In order to test whether Algorithm 3.4 produces reliable estimates to our structured condition numbers we used random polynomials of random matrices in $\mathbb{J}$ or $\mathbb{L}$, where $\mathbb{J}$ and $\mathbb{L}$ are the Jordan and Lie algebras relating to a bilinear form $\langle x, y\rangle = x^T My$, where M is a random symmetric orthogonal matrix. See Experiment 4.2, in section 4, and Experiment 5.1, in section 5, for more details on how this random data was produced. We used three cycles in Algorithm 3.4 to compute our estimates, $K_{\mathbb{J}}^{\rm est}(f, X)$ and $K_{\mathbb{L}}^{\rm est}(f, X)$. Figure 3.1 plots the ratios $K_{\mathbb{J}}^{\rm est}(f, X)/K_{\mathbb{J}}(f, X)$ and $K_{\mathbb{L}}^{\rm est}(f, X)/K_{\mathbb{L}}(f, X)$ for the 10000 examples in each experiment. We see from Figure 3.1 that Algorithm 3.4 can overestimate the true condition number. This can be explained by the fact that approximations are used in some steps of the algorithm, for example (3.7). Also massive cancellation can occur in the computation of (3.7). However, in virtually all examples, the estimates are within a factor of 2 of the correct value. As we are often only interested in the order of magnitude of the condition number, these results are acceptable.

An alternative to Algorithm 3.4 can be obtained by applying the power method to $\frac{1}{2}D(X)S_{\mathbb{L}}(M)^T$ and $\frac{1}{2}D(X)S_{\mathbb{J}}(M)^T$. Consider $K_{\mathbb{J}}(f, X)$ and the first step $y = \mathrm{vec}(Y) = \frac{1}{2}S_{\mathbb{L}}(M)^T\mathrm{vec}(Z_0)$. Then

$$Y = \frac{1}{2}(MZ_0^T + M^T Z_0)$$

would replace the step (3.8). Then

$$W = \frac{1}{\delta}\left(f(X + \delta Y) - f(X)\right),$$
$$U = \frac{1}{\delta}\left(f(X^T + \delta W_0) - f(X^T)\right),$$

where $W_0 = W/\|W\|_F$. We would also have to replace the final step (3.9) with

$$Z_1 = \frac{1}{2}(U^T M + MU).$$

To estimate $K_{\mathbb{L}}(f, X)$ we would again change the $+$ signs to $-$ signs in the equations for Y and Z_1. No appreciable difference can be seen in practice between this alternative algorithm and Algorithm 3.4.

3.2. Sesquilinear forms. In this section we shall assume that $\mathbb{J}$ and $\mathbb{L}$ denote a Jordan algebra and a Lie algebra associated with a nondegenerate sesquilinear form. For $S \in \mathbb{J}$, it can be shown that $\mathrm{vec}(S)$ satisfies

$$(M^T \otimes I)\mathrm{vec}(S^*) - (I \otimes M)\mathrm{vec}(S) = 0.$$

Provided $M \in \mathbb{R}^{n\times n}$ then it can be shown that $\mathrm{vec}(\mathrm{Real}(S)) \in \mathrm{null}(S_{\mathbb{J}}(M))$ and $\mathrm{vec}(\mathrm{Imag}(S)) \in \mathrm{null}(S_{\mathbb{L}}(M))$. Using Theorem 3.2 we can show that

$$\mathrm{vec}(S) = PS_{\mathbb{L}}(M)^T x + iPS_{\mathbb{J}}(M)^T y, \tag{3.10}$$

for some $x, y \in \mathbb{R}^{n^2}$. Similarly, for $K \in \mathbb{L}$, $\mathrm{vec}(K)$ satisfies

$$\mathrm{vec}(K) = PS_{\mathbb{J}}(M)^T x + iPS_{\mathbb{L}}(M)^T y,$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

for some $x, y \in \mathbb{R}^{n^2}$.

THEOREM 3.5. *Let $X \in \mathbb{C}^{n\times n}$ and the scalar function $f(x) = \sum_{m=0}^{\infty} \alpha_m x^m$ be absolutely convergent for all $|x| < r$ where $\|X\|_2 < r$. Also, let $\mathbb{J}$ and $\mathbb{L}$ denote a Jordan algebra and a Lie algebra associated with a nondegenerate sesquilinear form $\langle x, y\rangle = x^* My$, where $M \in \mathbb{R}^{n\times n}$ satisfies $M^T M = I$ and $M = \delta M^T$ where $\delta = \pm 1$. Then*

$$K_{\mathbb{J}}(f, X) = \max_{v\in\mathbb{C}^{n^2}} \frac{\|D(X)(M \otimes I)(v + \delta P\bar{v})\|_2}{\|(v + \delta P\bar{v})\|_2}, \tag{3.11}$$

$$K_{\mathbb{L}}(f, X) = \max_{v\in\mathbb{C}^{n^2}} \frac{\|D(X)(M \otimes I)(v - \delta P\bar{v})\|_2}{\|(v - \delta P\bar{v})\|_2}. \tag{3.12}$$

Proof. First we consider $Z \in \mathbb{J}$. As $M = \delta M^T$ where $\delta = \pm 1$ we can show, using (3.10), that

$$\mathrm{vec}(Z) = (M \otimes I)(v + \delta P\bar{v})$$

for some $v \in \mathbb{C}^{n^2}$. Substituting this into (3.2) and using the fact that $M \otimes I$ is orthogonal when M is orthogonal gives the result in (3.11). Using a similar argument we can show that for $Z \in \mathbb{L}$,

$$\mathrm{vec}(Z) = (M \otimes I)(v - \delta P\bar{v})$$

for some $v \in \mathbb{C}^{n^2}$ and hence obtain the result in (3.12). ◻

4. Jordan Algebra.

In this section we shall compare the unstructured condition number of f, $K(f, X)$, with the structured condition number of f, $K_{\mathbb{J}}(f, X)$, for $X \in \mathbb{J}$. We shall first consider the case where the underlying scalar product is a bilinear form. Then in section 4.2 we shall consider the case where the underlying scalar product is a sesquilinear form.

4.1. Bilinear forms.

We shall first assume that $\mathbb{J}$ denotes the Jordan algebra relating to a nondegenerate bilinear form $\langle x, y\rangle = x^T My$ where $M \in \mathbb{R}^{n\times n}$ satisfies $M = \delta M^T$, $\delta = \pm 1$ and $M^T M = I$. Recall that (3.5) shows that the structured condition number of f is equal to the 2-norm of the matrix $\frac{1}{2}D(X)PS_{\mathbb{L}}(M)^T$. As pre- or post-multiplication by an orthogonal matrix does not affect the singular values, we shall consider the matrix $\frac{1}{2}(I \otimes M)D(X)PS_{\mathbb{L}}(M)^T$. It can be shown that

$$\frac{1}{2}(I \otimes M)D(X)PS_{\mathbb{L}}(M)^T = \frac{1}{2}H(X)(I_{n^2} + \delta P), \tag{4.1}$$

where

$$H(X) = (I \otimes M)D(X)(M \otimes I). \tag{4.2}$$

It is clear that the singular values of $H(X)$ are the same as those of $D(X)$. Therefore, to compare the unstructured condition number, $K(f, X)$, with the structured

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

condition number, $K_{\mathbb{J}}(f, X)$, we shall compare the singular values of $H(X)$ and $\frac{1}{2}H(X)(I_{n^2} + \delta P)$.

When $X \in \mathbb{J}$, the matrix $H(X)$ is highly structured. We can rearrange $H(X)$ to get:

$$
\begin{aligned}
H(X) &= (I \otimes M) \sum_{m=1}^{\infty} \alpha_m \sum_{k=0}^{m-1} \left((X^T)^{m-1-k} \otimes X^k\right)(M \otimes I) \\
&= (M \otimes M) \sum_{m=1}^{\infty} \alpha_m \sum_{k=0}^{m-1} X^{m-1-k} \otimes X^k. \qquad (4.3)
\end{aligned}
$$

From (4.3) it is easy to see that:

- $H(X) = H(X)^T$.
- $H(X)$ commutes with P, that is $PH(X) = H(X)P$.
- $H(X) = ([H(X)]_{ij})$ is a block $n \times n$ matrix where each block satisfies

$$[H(X)]_{ij} = \delta[H(X)]_{ij}^T \in \mathbb{K}^{n \times n}.$$

The following result about matrices that commute with unitary matrices allows us to compare the singular values of $H(X)$ and $\frac{1}{2}H(X)(I_{n^2} + \delta P)$.

LEMMA 4.1. *Let $A, B \in \mathbb{C}^{n \times n}$, where B is a Hermitian unitary matrix, satisfy*

$$AB = \pm BA.$$

Let B have eigenvalues 1 with multiplicity p and -1 with multiplicity $n-p$. Then $\frac{1}{2}A(I+B)$ has p singular values in common with A and $n-p$ zero singular values. Also $\frac{1}{2}A(I-B)$ has the other $n-p$ singular values in common with A plus p zero singular values.

Proof. We can write the singular value decomposition of A as $A = U\Sigma V^*$ where Σ is partitioned in block diagonal form such that $\Sigma = \mathrm{diag}(\sigma_j I)$ where $\sigma_1 > \cdots > \sigma_k$ are the k distinct singular values of A. If $AB = \pm BA$, then $(U\Sigma V^*)B = \pm B(U\Sigma V^*)$ and

$$\Sigma = \pm(U^*BU)\Sigma(V^*BV)^*. \qquad (4.4)$$

As B is unitary, this is just a singular value decomposition of a diagonal matrix. Therefore U^*BU and V^*BV must be block diagonal matrices and conformably partitioned with Σ [5, Theorem 3.1.1′]. Using this, we can see that

$$
\begin{aligned}
\frac{1}{2}A(I+B) &= \frac{1}{2}U\left(\Sigma(I + V^*BV)\right)V^* \\
&= U\mathrm{diag}\left(\frac{\sigma_j}{2}(I + E_j)\right)V^*
\end{aligned}
$$

where $E = V^*BV = \mathrm{diag}(E_j)$. As B is Hermitian, the block diagonal matrix $\mathrm{diag}\left(\frac{\sigma_j}{2}(I+E_j)\right)$ is also Hermitian. Therefore the singular values of $\frac{1}{2}A(I+B)$ are equal to the absolute values of the eigenvalues of $\mathrm{diag}\left(\frac{\sigma_j}{2}(I+E_j)\right)$. We recall that B

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

has eigenvalues ± 1 and hence so do the diagonal blocks E_j. Let E_j have eigenvalues 1 with multiplicity p_j and -1 with multiplicity n_j. Then the diagonal block $\frac{\sigma_j}{2}(I+E_j)$ has σ_j as an eigenvalue with multiplicity p_j and 0 as an eigenvalue with multiplicity n_j. As $\sum_{j=1}^{k} p_j = p$, $\frac{1}{2}A(I+B)$ has p singular values in common with A. A similar argument shows that

$$\frac{1}{2}A(I-B) = U\mathrm{diag}\left(\frac{\sigma_j}{2}(I-E_j)\right)V^*$$

and therefore $\frac{1}{2}A(I-B)$ has the other $n-p$ singular values of A that are "missing" from $\frac{1}{2}A(I+B)$. $\square$

Using Lemma 4.1 we can see that $\frac{1}{2}H(X)(I_{n^2}+\delta P)$ has $\frac{1}{2}n(n+\delta)$ singular values in common with $H(X)$. The natural question arises: which singular values of $H(X)$ does $\frac{1}{2}H(X)(I_{n^2}+\delta P)$ share and do they have the same largest singular value? To gain insight into this question we performed the following experiment 10000 times.

Experiment 4.2. *Using normally distributed random variables with mean 0 and variance 1, generate*

- *Random Householder matrix M such that $My = \|y\|_2 e_1$ where y is a random vector in $\mathbb{R}^3$.*
- *Random polynomial $f(x) = a_6x^6 + \cdots + a_1x + a_0$ where the coefficients a_i are randomly distributed.*
- *Random $X \in \mathbb{J}$. This is formed using random $A \in \mathbb{R}^{3\times 3}$ and forming $X = AM^T + A^TM$.*

Using this data, $H(X)$ is formed, from which the condition numbers $K(f,X)$ and $K_{\mathbb{J}}(f,X)$ are computed. In all 10000 examples we found $K(f,X) = K_{\mathbb{J}}(f,X)$. This seemed to suggest that equality may hold for all f, M where $M = \pm M^T$ and $M^TM = I$, and $X \in \mathbb{J}$. However, examples where $K(f,X) > K_{\mathbb{J}}(f,X)$ can be found. For example, take

- Householder matrix M such that $My = \|y\|_2 e_1$ where

$$y = [\,-0.4442 \quad -0.5578 \quad -0.2641\,]^T.$$

- A polynomial

$$f(x) = -0.2879x^6 + 1.2611x^5 + 2.3149x^4 - 0.2079x^3 + 2.1715x^2 + 0.6125x.$$

- $X = AM^T + A^TM \in \mathbb{J}$ where

$$A = \begin{bmatrix} -2.0820 & -0.1532 & 1.4778 \\ -0.1035 & 0.1206 & -0.7404 \\ 1.0344 & 1.1157 & -0.9895 \end{bmatrix}.$$

In this example, $K(f,X) = 10.5813$ while $K_{\mathbb{J}}(f,X) = 8.7644$. This example was generated using direct search methods in MATLAB to try to maximize the ratio $K(f,X)/K_{\mathbb{J}}(f,X)$. The ratio in this example is just over 1.2073, which suggests that this ratio may remain small. A bound for this ratio, based on the properties of $H(X)$, is given in section 4.1.2.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

Examples where $K(f,X) > K_{\mathbb{J}}(f,X)$ seem very rare and hard to characterize. However, for the class of real symmetric matrices, which is the Jordan algebra relating to the bilinear form $\langle x,y\rangle = x^Ty$, we show in section 4.1.3 that $K(f,X) = K_{\mathbb{J}}(f,X)$ for all symmetric X.

4.1.1. What about the other singular values of $H(X)$? To see what happened to the singular values of $H(X)$ that are not singular values of $\frac{1}{2}H(X)(I_{n^2}+\delta P)$, consider the condition number of f at X, where $X \in \mathbb{J}$, subject to perturbations from the Lie algebra:

$$K_{\mathbb{L}}(f,X) = \frac{1}{2}\|(I\otimes M)D(X)PS_{\mathbb{J}}(M)^T\|_2.$$

It is easily seen that

$$\frac{1}{2}(I\otimes M)D(X)PS_{\mathbb{J}}(M)^T = \frac{1}{2}H(X)(I_{n^2}-\delta P),$$

where $H(X)$ is defined in (4.2). Using Lemma 4.1 we can see that $\frac{1}{2}H(X)(I_{n^2}-\delta P)$ has the other $\frac{1}{2}n(n-\delta)$ singular values of $H(X)$ that are "missing" from $\frac{1}{2}H(X)(I_{n^2}+\delta P)$. This also shows that when $X \in \mathbb{J}$,

$$K(f,X) = \max\{K_{\mathbb{J}}(f,X), K_{\mathbb{L}}(f,X)\}.$$

4.1.2. Bounding $K(f,X)/K_{\mathbb{J}}(f,X)$. In order to bound the ratio of the unstructured and structured condition numbers, where $X \in \mathbb{J}$, we shall consider the set

$$\mathbb{H} = \{H \in \mathbb{K}^{n^2\times n^2} : H = (H_{ij}) \text{ with } H_{ij} = \delta H_{ij}^T \in \mathbb{K}^{n\times n} \text{ and } PH = HP\}, \tag{4.5}$$

where $\delta = \pm 1$ and P is the vec-permutation matrix. All possible $H(X)$, formed from a function f at $X \in \mathbb{J}$, belong to $\mathbb{H}$. Therefore

$$\max_{X\in\mathbb{J}} \frac{K(f,X)}{K_{\mathbb{J}}(f,X)} \le \max_{G\in\mathbb{H}} \frac{2\|G\|_2}{\|G(I_{n^2}+\delta P)\|_2}.$$

The interesting case is where $\|H(X)\|_2 > \frac{1}{2}\|H(X)(I_{n^2}+\delta P)\|_2$. We have shown that when this happens $\|H(X)\|_2 = \frac{1}{2}\|H(X)(I_{n^2}-\delta P)\|_2$. Therefore we can equivalently consider

$$\max_{G\in\mathbb{H}} \frac{\|G(I_{n^2}-\delta P)\|_2}{\|G(I_{n^2}+\delta P)\|_2}.$$

In order to exploit the properties of the matrices in $\mathbb{H}$ it is convenient to introduce a 4-point coordinate system to identify elements of $G \in \mathbb{H}$:

$$(a,b,c,d) = G_{n(a-1)+b,n(c-1)+d}, \quad 1 \le a,b,c,d \le n.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

Therefore (a,b,c,d) refers to the element of G in the bth row of the ath block row and the dth column of the cth block column. We can now interpret the two properties of $\mathbb{H}$:

$$(a,b,c,d) = \delta(a,d,c,b) \quad (\text{regarding } G_{ij} = \delta G_{ij}^T), \tag{4.6}$$

$$(a,b,c,d) = \ (b,a,d,c) \quad (\text{regarding } GP = PG). \tag{4.7}$$

Using an alternate application of (4.6) and (4.7) we can show that

$$\begin{aligned} (a,b,c,d) &= \delta(a,d,c,b) = \\ \delta(d,a,b,c) &= \ (d,c,b,a) = \\ (c,d,a,b) &= \delta(c,b,a,d) = \\ \delta(b,c,d,a) &= \ (b,a,d,c). \end{aligned} \tag{4.8}$$

It can be seen from (4.8) that $(a,b,c,d) = (c,d,a,b)$ and therefore G is symmetric. Also noticeable is the fact that the left hand side of (4.8) are all the cyclic permutations of (a,b,c,d) while the right hand side of (4.8) are all the cyclic permutations of the reverse ordering (d,c,b,a). Other permutations of a,b,c and d give different elements of G:

$$\begin{aligned} (a,b,d,c) &= \delta(a,c,d,b) = \\ \delta(c,a,b,d) &= \ (c,d,b,a) = \\ (d,c,a,b) &= \delta(d,b,a,c) = \\ \delta(b,d,c,a) &= \ (b,a,c,d), \end{aligned} \tag{4.9}$$

and

$$\begin{aligned} (a,c,b,d) &= \delta(a,d,b,c) = \\ \delta(d,a,c,b) &= \ (d,b,c,a) = \\ (b,d,a,c) &= \delta(b,c,a,d) = \\ \delta(c,b,d,a) &= \ (c,a,d,b). \end{aligned} \tag{4.10}$$

All 4! permutations of a,b,c and d are accounted for. When a,b,c and d are all distinct integers then we have three sets of eight elements where all elements in the same set have the same value (up to signs). However, when a,b,c and d are not all distinct integers then all the elements will be repeated the same number of times in the lists (4.8), (4.9) and (4.10). Also, the integers a,b,c and d won't refer to three unique values of G. For example, when $a=b$, a,b,c and d refer to just two unique values of G. This is seen from the fact that list (4.8) is identical to the list (4.9) and the list (4.10) only refers to four unique elements of G.

LEMMA 4.3. *Let $\mathbb{H}$ be as defined in* (4.5). *Then*

$$\max_{G\in\mathbb{H}} \frac{\|G(I_{n^2}-\delta P)\|_2}{\|G(I_{n^2}+\delta P)\|_2} \le \sqrt{\frac{3n(n+\delta)}{2}},$$

where $\delta = \pm 1$ and P is the vec-permutation matrix.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

Proof. We can first show that

$$\max_{G\in\mathbb{H}} \frac{\|G(I_{n^2}-\delta P)\|_2}{\|G(I_{n^2}+\delta P)\|_2} \le \max_{G\in\mathbb{H}} \sqrt{\operatorname{rank}(G(I_{n^2}+\delta P))}\frac{\|G(I_{n^2}-\delta P)\|_F}{\|G(I_{n^2}+\delta P)\|_F}$$
$$\le \sqrt{\frac{n(n+\delta)}{2}} \max_{G\in\mathbb{H}} \frac{\|G(I_{n^2}-\delta P)\|_F}{\|G(I_{n^2}+\delta P)\|_F}.$$

Now we have to bound $\max_{G\in\mathbb{H}} \frac{\|G(I_{n^2}-\delta P)\|_F}{\|G(I_{n^2}+\delta P)\|_F}$. Define the ordered set

$$\mathbb{S} = \{\{a,b,c,d\}\colon 1\le a\le b\le c\le d\le n\}$$

and

$$\mathbb{T}_{\{a,b,c,d\}} = \{\{p,q,r,s\}\colon \text{All distinct permutations of } \{a,b,c,d\}\in\mathbb{S}\}.$$

As $(GP)_{n(a-1)+b,n(c-1)+d} = (a,b,d,c)$,

$$\frac{\|G(I_{n^2}-\delta P)\|_F^2}{\|G(I_{n^2}+\delta P)\|_F^2} = \frac{\sum_{\{a,b,c,d\}\in\mathbb{S}} f(a,b,c,d)}{\sum_{\{a,b,c,d\}\in\mathbb{S}} g(a,b,c,d)}$$

where

$$f(a,b,c,d) = \sum_{\{p,q,r,s\}\in\mathbb{T}_{\{a,b,c,d\}}} ((p,q,r,s)-\delta(p,q,s,r))^2,$$
$$g(a,b,c,d) = \sum_{\{p,q,r,s\}\in\mathbb{T}_{\{a,b,c,d\}}} ((p,q,r,s)+\delta(p,q,s,r))^2.$$

It can easily be shown that

$$\max_{G\in\mathbb{H}} \frac{\|G(I_{n^2}-\delta P)\|_F^2}{\|G(I_{n^2}+\delta P)\|_F^2} \le \max_{\{a,b,c,d\}\in\mathbb{S}} \left(\max \frac{f(a,b,c,d)}{g(a,b,c,d)}\right).$$

Now we have to bound $\max \frac{f(a,b,c,d)}{g(a,b,c,d)}$ for all possible $\{a,b,c,d\}\in\mathbb{S}$.

All a,b,c and d different. Let $x=(a,b,c,d)$, $y=(a,c,d,b)$ and $z=(a,d,b,c)$. Then we can show that for all $\{a,b,c,d\}\in\mathbb{S}$,

$$\frac{f(a,b,c,d)}{g(a,b,c,d)} = \frac{8\left((x-y)^2+(y-z)^2+(z-x)^2\right)}{8\left((x+y)^2+(y+z)^2+(z+x)^2\right)}. \tag{4.11}$$

Rearranging yields

$$\frac{f(a,b,c,d)}{g(a,b,c,d)} = \frac{3(x^2+y^2+z^2)-(x+y+z)^2}{(x^2+y^2+z^2)+(x+y+z)^2}$$

and therefore $\max \frac{f(a,b,c,d)}{g(a,b,c,d)} = 3$ which is obtained at $x+y+z=0$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

Two integers equal, two different. Let $x = (a, a, c, d)$, $y = (a, c, d, a)$, and $z = (a, d, a, c)$. For $\delta = 1$, it is easy to see from (4.8) and (4.9) that $x = y$. Then,

$$f(a, a, c, d) = 8(x - z)^2,$$
$$g(a, a, c, d) = 8(x + z)^2 + 16x^2.$$

Notice that $\frac{f(a,a,c,d)}{g(a,a,c,d)}$ is a special case of (4.11) where $x = y$ and therefore has a maximum of 3 at $2x + z = 0$. For $\delta = -1$,

$$f(a, a, c, d) = 24x^2,$$
$$g(a, a, c, d) = 8x^2.$$

This is also a special case of (4.11) where $x = -y$ and $z = 0$ ($z = (d, a, c, a)$ which is on the diagonal of a block) and $\frac{f(a,a,c,d)}{g(a,a,c,d)}$ is always the maximum 3.

Two integers equal twice. Let $x = (a, a, d, d)$, $y = (a, d, d, a)$, and $z = (a, d, a, d)$. For $\delta = 1$, $x = y$ and

$$f(a, a, d, d) = 4(x - z)^2,$$
$$g(a, a, d, d) = 4(x + z)^2 + 8x^2.$$

Therefore $\max \frac{f(a,a,d,d)}{g(a,a,d,d)} = 3$ at $2x + z = 0$. For $\delta = -1$, $x = -y$, $z = 0$ and

$$f(a, a, d, d) = 12x^2,$$
$$g(a, a, d, d) = 4x^2,$$

and $\frac{f(a,a,d,d)}{g(a,a,d,d)} = 3$.

More than two integers equal. Let $x = (a, a, a, d)$, $y = (a, a, d, a)$, and $z = (a, d, a, a)$. For $\delta = 1$, it can be shown that $x = y = z$. This means, all permutations of (a, a, a, d) are equal. Hence $f(a, a, a, d) = 0$ and $\max \frac{f(a,a,a,d)}{g(a,a,a,d)} = 0$ for all x, y and z. For $\delta = -1$, we have $x = y = z = 0$ and therefore $f(a, a, a, d) = 0$.

Therefore $\max_{\{a,b,c,d\}\in\mathbb{S}}(\max \frac{f(a,b,c,d)}{g(a,b,c,d)}) = 3$ and the result follows immediately from this. ∎

THEOREM 4.4. *Let $X \in \mathbb{J}$, where $\mathbb{J}$ denotes the Jordan algebra relating to a nondegenerate bilinear form $\langle x, y\rangle = x^T My$, where $M \in \mathbb{R}^{n\times n}$ satisfies $M^T M = I$ and $M = \delta M^T$ for $\delta = \pm 1$. Let the scalar function*

$$f(x) = \sum_{m=0}^{\infty} \alpha_m x^m$$

be absolutely convergent for all $|x| < r$ where $\|X\|_2 < r$. Then

$$\frac{K(f, X)}{K_{\mathbb{J}}(f, X)} \le \sqrt{\frac{3n(n + \delta)}{2}}.$$

Proof. This result comes easily from Lemma 4.3. ∎

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

4.1.3. Real symmetric case ($M = I$). It is known that restricting perturbations to symmetric linear systems or eigenvalue problems to be symmetric makes little difference to the backward error or the condition of the problem [3], [4]. The same can be shown for the condition of matrix functions. We shall start with the following lemma, which is essentially the same as a result given in [6, Lemma 2.1], but written as a matrix factorization.

LEMMA 4.5. *Let $X \in \mathbb{R}^{n\times n}$ be diagonalizable and have the eigendecomposition $X = QDQ^{-1}$ where $D = \mathrm{diag}(\lambda_k)$. Then the Kronecker form of the Fréchet derivative of $f(X)$ is also diagonalizable and has the eigendecomposition $D(X) = V\Phi V^{-1}$ where $V = Q^{-T} \otimes Q$, $\Phi = \mathrm{diag}(\phi_k)$ and*

$$\phi_{n(i-1)+j} = \begin{cases} \frac{f(\lambda_i)-f(\lambda_j)}{\lambda_i-\lambda_j} & \lambda_i \neq \lambda_j, \\ f'(\lambda_j) & \lambda_i = \lambda_j. \end{cases}$$

Proof. Considering $\Phi = V^{-1}D(X)V$, we see that

$$\Phi = \sum_{m=1}^{\infty} \alpha_m \sum_{k=0}^{m-1} D^{m-1-k} \otimes D^k = \mathrm{diag}(\phi_k).$$

The kth diagonal element of Φ, where $k = n(i-1) + j$ for some unique $1 \leq i, j \leq n$, is then given by

$$\phi_k = \sum_{m=1}^{\infty} \alpha_m \sum_{k=0}^{m-1} \lambda_i^{m-1-k}\lambda_j^k.$$

If $\lambda_i = \lambda_j$, including the case where $i = j$, then $\sum_{k=0}^{m-1} \lambda_i^{m-1-k}\lambda_j^k = m\lambda_i^{m-1}$ and

$$\phi_k = \sum_{m=1}^{\infty} \alpha_m m\lambda_i^{m-1} = f'(\lambda_i).$$

If $\lambda_i \neq \lambda_j$, then $\sum_{k=0}^{m-1} \lambda_i^{m-1-k}\lambda_j^k = \frac{\lambda_i^m-\lambda_j^m}{\lambda_i-\lambda_j}$ and

$$\phi_k = \sum_{m=1}^{\infty} \alpha_m \frac{\lambda_i^m - \lambda_j^m}{\lambda_i - \lambda_j} = \frac{f(\lambda_i) - f(\lambda_j)}{\lambda_i - \lambda_j}. \qquad \square$$

We now consider a symmetric matrix X and the condition number of f at X, subject to symmetric perturbations.

THEOREM 4.6. *Let $X = X^T \in \mathbb{R}^{n\times n}$ and the scalar function $f(x) = \sum_{m=0}^{\infty} \alpha_m x^m$ be absolutely convergent for all $|x| < r$ where $\|X\|_2 < r$. Then*

$$K(f, X) = K_{\mathbb{J}}(f, X).$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

Proof. We have seen that $K(f,X) = \|D(X)\|_2$ and Theorem 3.3 shows that

$$K_{\mathbb{J}}(f,X) = \max_{Z\neq 0, Z\text{symm.}} \frac{\|L(Z,X)\|_F}{\|Z\|_F} = \frac{1}{2}\|D(X)PS_{\mathbb{L}}(I)^T\|_2.$$

Therefore, we shall compare the singular values of $D(X)$ and $\frac{1}{2}D(X)PS_{\mathbb{L}}(I)^T$ to show our result. As $M = I$ we can see that $D(X) = H(X)$, where $H(X)$ is defined in (4.2), and from (4.1), $\frac{1}{2}D(X)PS_{\mathbb{L}}(I)^T = \frac{1}{2}H(X)(I_{n^2} + \delta P)$ and therefore both matrices are symmetric.

As X is symmetric, we can write its eigendecomposition $X = QDQ^T$ where Q is orthogonal and $D = \mathrm{diag}(\lambda_k)$. ¿From Lemma 4.5 we see that $D(X) = V\Phi V^T$ where $V = Q \otimes Q$ is also orthogonal and $\Phi = \mathrm{diag}(\phi_k)$. It is easy to see that P commutes with V, and therefore

$$\frac{1}{2}D(X)PS_{\mathbb{L}}(I)^T = \frac{1}{2}V(\Phi P + \Phi)V^T.$$

As $\phi_{n(i-1)+j} = \phi_{n(j-1)+i}$, a similarity transformation can be applied to $\frac{1}{2}(\Phi P + \Phi)$ using a permutation matrix to get a block diagonal matrix consisting of

- n 1×1 blocks $[\phi_{n(i-1)+i}]$ for $1 \le i \le n$.
- $\frac{1}{2}n(n-1)$ 2×2 blocks $\frac{1}{2}\left[\begin{smallmatrix}\phi & \phi\\ \phi & \phi\end{smallmatrix}\right]$ where $\phi = \phi_{n(i-1)+j}$ for $1 \le i < j \le n$.

Hence the eigenvalues μ_k of $\frac{1}{2}(\Phi P + \Phi)$ are

$$\mu_{n(i-1)+j} = \begin{cases} \phi_{n(i-1)+j} & i \le j, \\ 0 & i > j. \end{cases}$$

The nonzero parts of the spectra of $D(X)$ and $\frac{1}{2}D(X)PS_{\mathbb{L}}(I)^T$ are equal, if multiplicities are ignored. As both matrices are symmetric, their singular values are equal to the absolute values of their eigenvalues and so $\|D(X)\|_2 = \frac{1}{2}\|D(X)PS_{\mathbb{L}}(I)^T\|_2$. $\square$

Theorem 4.6 shows that the condition number of f at a symmetric matrix X is unaffected if the perturbations are restricted to just symmetric perturbations.

4.2. Sesquilinear forms. We shall now assume that $\mathbb{J}$ denotes the Jordan algebra relating to a nondegenerate sesquilinear form $\langle x,y\rangle = x^*My$ where $M \in \mathbb{R}^{n\times n}$ satisfies $M = \delta M^T$, $\delta = \pm 1$ and $M^TM = I$. From (3.11) we can show that

$$K_{\mathbb{J}}(f,X) = \max_{v\in\mathbb{C}^{n^2}} \frac{\|H(X)(v + \delta P\bar{v})\|_2}{\|v + \delta P\bar{v}\|_2},$$

where $H(X)$ is defined in (4.2). When $X \in \mathbb{J}$, the matrix $H(X)$ is highly structured and it can be shown that

- $H(X)$ is Hermitian.
- $H(X)P = PH(X)^T$.

Using these properties of $H(X)$ we can obtain the following result.

LEMMA 4.7. *Let $H \in \mathbb{R}^{n^2\times n^2}$ satisfy $HP = PH^T$ and $H = \mu H^*$ where $\mu = \pm 1$. Then*

$$\max_{v\in\mathbb{C}^{n^2}} \frac{\|H(v + \alpha P\bar{v})\|_2}{\|v + \alpha P\bar{v}\|_2} = \|H\|_2$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

where $\alpha \in \mathbb{C}$.

Proof. Consider an eigenpair (λ, y) of H. If H is Hermitian, its eigenvalues are real. If H is skew-Hermitian, its eigenvalues are purely imaginary. Therefore

$$\begin{aligned} H(y + \alpha P\bar{y}) &= Hy + \alpha P H^T \bar{y}, \\ &= \lambda y + \alpha P(\mu\bar{\lambda}\bar{y}), \\ &= \lambda(y + \alpha P\bar{y}). \end{aligned}$$

If $y + \alpha P\bar{y} = 0$, then we can always replace y with βy where $\beta \in \mathbb{C}$ and $\mathrm{Imag}(\beta) \neq 0$. Then $y+\alpha P\bar{y}$ is an eigenvector of H corresponding to the eigenvalue λ. As $H = \pm H^*$, its singular values are equal to the absolute values of its eigenvalues. Therefore, using the eigenvectors y_k of H, we have

$$\sigma_k = \frac{\|H(y_k + \alpha P\bar{y_k})\|_2}{\|y_k + \alpha P\bar{y_k}\|_2}$$

for all singular values σ_k of H. The result follows immediately from this. □

Theorem 4.8. *Let $X \in \mathbb{J}$, where $\mathbb{J}$ denotes the Jordan algebra relating to a nondegenerate sesquilinear form $\langle x, y\rangle = x^* My$, where $M \in \mathbb{R}^{n\times n}$ satisfies $M^T M = I$ and $M = \delta M^T$ for $\delta = \pm 1$. Let the scalar function*

$$f(x) = \sum_{m=0}^{\infty} \alpha_m x^m$$

be absolutely convergent for all $|x| < r$ where $\|X\|_2 < r$. Then

$$K(f, X) = K_{\mathbb{J}}(f, X).$$

Proof. This result comes easily from Lemma 4.7 using $\alpha = \delta$. □

We can also show that

$$K_{\mathbb{L}}(f, X) = \max_{v\in\mathbb{C}^{n^2}} \frac{\|H(X)(v - \delta P\bar{v})\|_2}{\|v - \delta P\bar{v}\|_2}.$$

Therefore using Lemma 4.7 with $\alpha = -\delta$ we can also show that $K(f, X) = K_{\mathbb{L}}(f, X)$ under the same conditions on f, X and $\mathbb{J}$ that are used in Theorem 4.8.

5. Lie Algebra. In this section we shall compare the unstructured condition number of f, $K(f, X)$, with the structured condition number of f, $K_{\mathbb{L}}(f, X)$, for $X \in \mathbb{L}$. We shall first consider the case where the underlying scalar product is a bilinear form. Then in section 5.2 we shall consider the case where the underlying scalar product is a sesquilinear form.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

5.1. Bilinear forms. We shall first assume that $\mathbb{L}$ denotes the Lie algebra relating to a nondegenerate bilinear form $\langle x, y\rangle = x^T M y$ where $M \in \mathbb{R}^{n\times n}$ satisfies $M = \delta M^T$, $\delta = \pm 1$ and $M^T M = I$. To compare the unstructured condition number $K(f, X)$, with the structured condition number $K_{\mathbb{L}}(f, X)$, we shall compare the singular values of $D(X)$ and those of $\frac{1}{2}D(X)PS_{\mathbb{J}}(M)^T$. Equivalently, we can compare the singular values of $H(X)$, defined in (4.2), and $\frac{1}{2}H(X)(I_{n^2} - \delta P)$. We have not been able to find a pattern between these singular values. For an arbitrary function $f(X)$ and $X \in \mathbb{L}$, the matrix $H(X)$ is not necessarily symmetric nor does it commute with P. However, if f is restricted to being an odd or even function, then a pattern arises. This is a natural restriction, as f is an odd function if and only if $f(\mathbb{L}) \subseteq \mathbb{L}$, while f is an even function if and only if $f(\mathbb{L}) \subseteq \mathbb{J}$ [7].

First, consider an odd function $f(X) = \sum_{m=0}^{\infty} \alpha_{2m+1} X^{2m+1}$. The Kronecker form of the Fréchet derivative of f is

$$D(X) = \sum_{m=1}^{\infty} \alpha_{2m-1} \sum_{k=0}^{2m-2} (X^T)^{2m-2-k} \otimes X^k.$$

When $X \in \mathbb{L}$, the matrix $H(X)$ is highly structured. We can rearrange $H(X)$ to get:

$$\begin{aligned} H(X) &= (I \otimes M)D(X)(M \otimes I), \\ &= (M \otimes M) \sum_{m=1}^{\infty} \alpha_{2m-1} \sum_{k=0}^{2m-2} (-1)^k X^{2m-2-k} \otimes X^k. \end{aligned}$$

It can be shown that
- $H(X) = H(X)^T$.
- $H(X)$ commutes with P, that is $PH(X) = H(X)P$.

Using Lemma 4.1 we can show that the matrix $\frac{1}{2}H(X)(I_{n^2} - \delta P)$ has $\frac{1}{2}n(n-\delta)$ singular values in common with $H(X)$. The natural question arises: which singular values of $H(X)$ does $\frac{1}{2}H(X)(I_{n^2} - \delta P)$ share and do they have the same largest singular values? To gain insight into this question we performed the following experiment 10000 times.

Experiment 5.1. *Using normally distributed random variables with mean 0 and variance 1, generate*
- *Random Householder matrix M such that $My = \|y\|_2 e_1$ where y is a random vector in $\mathbb{R}^3$.*
- *Random polynomial $f(x) = a_5 x^5 + a_3 x^3 + a_1 x$ where the coefficients a_i are randomly distributed.*
- *Random $X \in \mathbb{L}$. This is formed using random $A \in \mathbb{R}^{3\times 3}$ and forming $X = AM^T - A^T M$.*

Using this data, $H(X)$ is formed, from which the condition numbers $K(f, X)$ and $K_{\mathbb{L}}(f, X)$ are computed. We found a marked difference between the results of Experiment 4.2 and 5.1. Out of 10000 examples, on just 740 occasions did we find $K(f, X) = K_{\mathbb{L}}(f, X)$. Also $K(f, X)/K_{\mathbb{L}}(f, X)$ could grow large. In this experiment we achieved a maximum of $K(f, X)/K_{\mathbb{L}}(f, X) = 349$. In fact, this ratio is unbounded. Let

$$X = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathbb{L},$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

where $\mathbb{L}$ is the Lie algebra associated with the bilinear form $\langle x, y\rangle = x^T y$, that is, the class of skew-symmetric matrices. Let $f(X) = X^3 + 3X$. Then

$$H(X) = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}.$$

Note that $H(X)P$ is the same as $H(X)$ except that the second and third columns have been swapped over. Then, it is easy to see that $\frac{1}{2}H(X)(I_{n^2} - \delta P) = 0$ and therefore $K_{\mathbb{L}}(f, X) = 0$. Hence, $K(f, X)/K_{\mathbb{L}}(f, X)$ is unbounded at X. We can call X a "stationary point" of the function $f(X) = X^3 + 3X$ when X is restricted to the class of skew-symmetric matrices.

Now, consider an even function $f(X) = \sum_{m=0}^{\infty} \alpha_{2m} X^{2m}$. The Kronecker form of the Fréchet derivative of f is

$$D(X) = \sum_{m=1}^{\infty} \alpha_{2m} \sum_{k=0}^{2m-1} (X^T)^{2m-1-k} \otimes X^k.$$

When $X \in \mathbb{L}$, the matrix $H(X)$ is highly structured. We can rearrange $H(X)$ to get:

$$\begin{aligned} H(X) &= (I \otimes M) D(X) (M \otimes I), \\ &= (M \otimes M) \sum_{m=1}^{\infty} \alpha_{2m} \sum_{k=0}^{2m-1} (-1)^{k+1} X^{2m-1-k} \otimes X^k. \end{aligned}$$

It can be shown that

- $H(X) = -H(X)^T$.
- $H(X)$ also satisfies $PH(X) = -H(X)P$.

These conditions are more restrictive on the singular values of $H(X)$ than those for odd functions. Because of this, more can be said about the structured condition number at even f.

THEOREM 5.2. *Let $X \in \mathbb{L}$, where $\mathbb{L}$ denotes the Lie algebra relating to a non-degenerate bilinear form $\langle x, y\rangle = x^T M y$, where $M \in \mathbb{R}^{n\times n}$ satisfies $M = \pm M^T$ and $M^T M = I$. Let the even scalar function*

$$f(x) = \sum_{m=0}^{\infty} \alpha_{2m} x^{2m}$$

be absolutely convergent for all $|x| < r$ where $\|X\|_2 < r$. Then

$$K(f, X) = K_{\mathbb{L}}(f, X).$$

Proof. We have shown that $K(f, X) = \|H(X)\|_2$ and

$$K_{\mathbb{L}}(f, X) = \frac{1}{2}\|H(X)(I_{n^2} - \delta P)\|_2, \quad \delta = \pm 1.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

Therefore, we shall compare the singular values of $H(X)$ and $\frac{1}{2}H(X)(I_{n^2} - \delta P)$ to show our result. Define

$$M(X) = \frac{1}{2}H(X)(I_{n^2} - P),$$
$$N(X) = \frac{1}{2}H(X)(I_{n^2} + P).$$

When $X \in \mathbb{L}$ and f is an even function, it can be shown that $M(X) = -N(X)^T$. Therefore $M(X)$ and $N(X)$ have the same singular values. Using Lemma 4.1 we can show that $M(X)$ has $\frac{1}{2}n(n-1)$ singular values in common with $H(X)$ (plus $\frac{1}{2}n(n+1)$ zero singular values) while $N(X)$ has the other $\frac{1}{2}n(n+1)$ singular values of $H(X)$ (plus $\frac{1}{2}n(n-1)$ zero singular values). As we have shown $M(X)$ and $N(X)$ have the same singular values, then

$$\|H(X)\|_2 = \|M(X)\|_2 = \|N(X)\|_2. \qquad \square$$

Recall that

$$K_{\mathbb{J}}(f, X) = \frac{1}{2}\|H(X)(I_{n^2} + \delta P)\|_2$$

where $\delta = \pm 1$. Therefore, the proof of Theorem 5.2 also shows that for $X \in \mathbb{L}$ and an even function f, the condition number is unaffected if the perturbations are restricted to $\mathbb{J}$. That is $K(f, X) = K_{\mathbb{J}}(f, X)$.

5.1.1. Real skew-symmetric case ($M = I$). We now compare the unstructured and the structured condition numbers of f at X when X is a skew-symmetric matrix. We have seen that $K(f, X) = \|D(X)\|_2$ and Theorem 3.3 shows that

$$K_{\mathbb{L}}(f, X) = \max_{Z \neq 0, Z \text{skew.}} \frac{\|L(Z, X)\|_F}{\|Z\|_F} = \frac{1}{2}\|D(X)PS_{\mathbb{J}}(I)^T\|_2.$$

Therefore, we shall compare the singular values of $D(X)$ and $\frac{1}{2}D(X)PS_{\mathbb{J}}(I)^T$. Using a slightly modified version of Lemma 4.5, where X^T is replaced by X^* in the Kronecker form of the Fréchet derivative, we can show that if X has the eigendecomposition $X = QDQ^*$ where Q is unitary and $D = \operatorname{diag}(\lambda_i)$ then $D(X) = V\Phi V^*$ where

$$\Phi = \sum_{m=1}^{\infty} \alpha_m \sum_{k=0}^{m-1} (D^*)^{m-1-k} \otimes D^k = \operatorname{diag}(\phi_k)$$

and $V = Q \otimes Q$. The diagonal elements of Φ are given by

$$\phi_{n(i-1)+j} = \begin{cases} \frac{f(\lambda_i^*) - f(\lambda_j)}{\lambda_i^* - \lambda_j} & \lambda_i^* \neq \lambda_j, \\ f'(\lambda_j) & \lambda_i^* = \lambda_j. \end{cases} \tag{5.1}$$

As P commutes with V, it is easy to see that

$$\frac{1}{2}D(X)PS_{\mathbb{J}}(I)^T = \frac{1}{2}V(\Phi P - \Phi)V^*.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

By applying a similarity transformation to $\frac{1}{2}(\Phi P - \Phi)$ using a permutation matrix we can get a block diagonal matrix consisting of n 1×1 blocks whose elements are zero and $\frac{1}{2}n(n-1)$ 2×2 blocks

$$\Lambda_{ij} = \frac{1}{2}\begin{bmatrix} -\phi_{n(i-1)+j} & \phi_{n(j-1)+i} \\ \phi_{n(i-1)+j} & -\phi_{n(j-1)+i} \end{bmatrix}$$

for $1 \le i < j \le n$. Using the fact that $f(\lambda_k^*) = f(\lambda_k)^*$, it is easy to see from (5.1) that $\phi_{n(i-1)+j} = \phi^*_{n(j-1)+i}$. Therefore the singular values of Λ_{ij} are $|\phi_{n(i-1)+j}|$ and 0 and the singular values of $\frac{1}{2}(\Phi P - \Phi)$ are

$$\sigma_{n(i-1)+j} = \begin{cases} 0 & i \le j \\ |\phi_{n(i-1)+j}| & i > j. \end{cases} \tag{5.2}$$

As in the symmetric case, the singular values of $D(X)$ exist in pairs, $|\phi_{n(i-1)+j}| = |\phi_{n(j-1)+i}|$. However, ignoring multiplicities, not all the nonzero singular values of $D(X)$ appear as singular values of $\frac{1}{2}D(X)PS_{\mathbb{J}}(I)^T$. The "missing" singular values are

$$\widehat{\sigma}_i = |\phi_{n(i-1)+i}| = \begin{cases} |\frac{f(\lambda_i^*)-f(\lambda_i)}{\lambda_i^*-\lambda_i}| & \lambda_i \ne 0, \\ |f'(0)| & \lambda_i = 0. \end{cases} \tag{5.3}$$

For certain functions these "missing" singular values are never the largest singular values of $D(X)$ and so, for these functions, we have $K(f,X) = K_{\mathbb{L}}(f,X)$ for all skew-symmetric X.

LEMMA 5.3. *Let the scalar function $f(x) = \sum_{m=0}^{\infty} \alpha_m x^m$ be absolutely convergent for all $|x| < r$, and let*

$$\max\left\{|f'(0)|, \left|\frac{f(\mu i) - f(-\mu i)}{2\mu}\right|\right\} \le |f'(\mu i)|. \tag{5.4}$$

for all real μ such that $0 < |\mu| < r$. Then

$$K(f,X) = K_{\mathbb{L}}(f,X)$$

for all skew-symmetric X such that $0 < \|X\|_2 < r$.

Proof. A skew-symmetric matrix has purely imaginary eigenvalues which we denote by $\lambda_k = i\mu_k$. Then using (5.3) we see that

$$\widehat{\sigma}_k = \begin{cases} |\frac{f(-i\mu_k)-f(i\mu_k)}{2\mu_k}| & \lambda_k \ne 0, \\ |f'(0)| & \lambda_k = 0. \end{cases}$$

are the singular values of $D(X)$ that are not singular values of $\frac{1}{2}D(X)PS_{\mathbb{J}}(I)^T$. All we have to show is that, providing (5.4) holds, there exists a singular value of $\frac{1}{2}D(X)PS_{\mathbb{J}}(I)^T$ that is greater than or equal to $\max \widehat{\sigma}_k$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

As X is a real matrix, its eigenvalues exist in complex conjugate pairs. Therefore, for each nonzero eigenvalue λ_i, then $\lambda_i = \lambda_j^*$ for some j. From (5.1), we can see that

$$|\phi_{n(i-1)+j}| = |f'(\lambda_j)| \quad \text{and} \quad |\phi_{n(j-1)+i}| = |f'(\lambda_i)|$$

are singular values of $D(X)$, and from (5.2), we can see that one of them is also a singular value of $\frac{1}{2}D(X)PS_{\mathbb{J}}(I)^T$. As $|f'(\lambda_i)| = |f'(\lambda_j)|$, we can use (5.4) to show there exists a singular value of $\frac{1}{2}D(X)PS_{\mathbb{J}}(I)^T$ greater than or equal to $\max \widehat{\sigma}_k$. $\square$

The condition (5.4) in Lemma 5.3 holds when f is a wide range of functions, including exponential, sine, cosine and cosh. For cosine and cosh and other even functions, this result has already been proved in Theorem 5.2. It can also be seen that the left hand side of (5.4) is zero for even functions. However this condition (5.4) does not hold for sinh. Examples where $K(\sinh, X) > K_{\mathbb{L}}(\sinh, X)$ are easily generated.

5.1.2. Skew-symmetric case, symmetric perturbations. If we consider the condition number of f at X, where X is skew-symmetric, subject to symmetric perturbations, then Theorem 3.3 shows that

$$K_{\mathbb{J}}(f, X) = \max_{Z \neq 0, Z \text{symm.}} \frac{\|L(Z, X)\|_F}{\|Z\|_F} = \frac{1}{2}\|D(X)PS_{\mathbb{L}}(I)^T\|_2.$$

It is easy to see that

$$\frac{1}{2}D(X)PS_{\mathbb{L}}(I)^T = \frac{1}{2}V(\Phi P + \Phi)V^*,$$

and the singular values of $\frac{1}{2}(\Phi P + \Phi)$ are

$$\sigma_{n(i-1)+j} = \begin{cases} |\phi_{n(i-1)+j}| & i \leq j \\ 0 & i > j. \end{cases}$$

As $|\phi_{n(i-1)+j}| = |\phi_{n(j-1)+i}|$ for all $1 \leq i, j \leq n$, all the nonzero singular values of $D(X)$ appear as singular values of $\frac{1}{2}D(X)PS_{\mathbb{L}}(I)^T$, if multiplicities are ignored. Therefore $K(f, X) = K_{\mathbb{J}}(f, X)$ which means the condition number of f at a skew-symmetric X, is unaffected if the perturbations are restricted to just symmetric perturbations.

5.2. Sesquilinear forms. We shall now assume that $\mathbb{L}$ denotes the Lie algebra relating to a nondegenerate sesquilinear form $\langle x, y\rangle = x^* M y$ where $M \in \mathbb{R}^{n\times n}$ satisfies $M = \delta M^T$, $\delta = \pm 1$ and $M^T M = I$. From (3.12) we can show that

$$K_{\mathbb{L}}(f, X) = \max_{v \in \mathbb{C}^{n^2}} \frac{\|H(X)(v - \delta P\bar{v})\|_2}{\|v - \delta P\bar{v}\|_2},$$

where $H(X)$ is defined in (4.2). We shall again restrict our attention to odd or even functions of X, as this gives properties of $H(X)$ which we can work with. For an odd function $f(x) = \sum_{m=0}^{\infty} \alpha_{2m+1}x^{2m+1}$ and $X \in \mathbb{L}$, we can show that

- $H(X)$ is Hermitian.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004
ELA
http://math.technion.ac.il/iic/ela

- $H(X)P = PH(X)^T$.

For an even function $f(x) = \sum_{m=0}^{\infty} \alpha_{2m} x^{2m}$ and $X \in \mathbb{L}$, we can show that

- $H(X)$ is skew-Hermitian.
- $H(X)P = PH(X)^T$.

These properties enable us to prove the following result.

THEOREM 5.4. *Let $X \in \mathbb{L}$, where $\mathbb{L}$ denotes the Lie algebra relating to a nondegenerate sesquilinear form $\langle x, y\rangle = x^* My$ where $M \in \mathbb{R}^{n\times n}$ satisfies $M^T M = I$ and $M = \delta M^T$ for $\delta = \pm 1$. Let the $f(x)$ be either*

- *An odd scalar function $f_{\text{odd}}(x) = \sum_{m=0}^{\infty} \alpha_{2m+1} x^{2m+1}$.*
- *An even scalar function $f_{\text{even}}(x) = \sum_{m=0}^{\infty} \alpha_{2m} x^{2m}$.*

Also, let $f(x)$ be absolutely convergent for all $|x| < r$ where $\|X\|_2 < r$. Then

$$K(f, X) = K_{\mathbb{L}}(f, X).$$

Proof. This result comes from the fact that $H(X)$ satisfies the conditions of Lemma 4.7 for both functions $f_{\text{odd}}(x)$ and $f_{\text{even}}(x)$. □

Using Lemma 4.7 with $\alpha = \delta$ we can also show that $K(f, X) = K_{\mathbb{J}}(f, X)$ under the same conditions on f, X and $\mathbb{L}$ that are used in Theorem 5.4.

6. Concluding Remarks. Kenney and Laub [6] presented a theory of conditioning of matrix functions $f(X)$ based on the Fréchet derivative at X. We have extended this theory by imposing structure on X and its perturbations. Structured conditioned numbers have been defined and, under certain conditions on the underlying scalar products, explicit representations have been given for them in Theorem 3.3 and Theorem 3.5. Comparisons between the unstructured and the structured condition numbers were made and Table 6.1 summarizes the main results of this paper. If the underlying scalar product is a sesquilinear form, we have shown that imposing structure does not affect the condition number of f for

- All functions f of $X \in \mathbb{J}$.
- Odd or even functions f of $X \in \mathbb{L}$.

However, when the underlying scalar product is a bilinear form, equality between the two condition numbers is not guaranteed in these cases. For general f and $X \in \mathbb{J}$, we have provided experimental and theoretical evidence to show that $K_{\mathbb{J}}(f, X)$ and $K(f, X)$ are often equal to each other. However, equality does not always hold and a bound

$$\frac{K(f, X)}{K_{\mathbb{J}}(f, X)} < \sqrt{\frac{3n(n+\delta)}{2}} \tag{6.1}$$

was proved. A few questions merit further investigation:

- Is it possible to characterize when $K(f, X) = K_{\mathbb{J}}(f, X)$?
- We have not been able to construct examples where the left hand side of (6.1) is much larger than 1.2. Are better bounds obtainable? Or, can we generate examples where $\frac{K(f,X)}{K_{\mathbb{J}}(f,X)}$ is as large as the bound suggests is possible?

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

ELA
http://math.technion.ac.il/iic/ela

Bilinear form $\langle x,y\rangle = x^T My$	$X \in \mathbb{J}$ **or** $\mathbb{L}$	**Matrix class**	**Function** f	**Result**
$M = I$	$\mathbb{J}$	Real symmetrics	all	$K(f,X) = K_{\mathbb{J}}(f,X)$
$M^TM=I$, $M=\pm M^T$	$\mathbb{J}$		all	$\frac{K(f,X)}{K_{\mathbb{J}}(f,X)} \le \sqrt{\frac{3n(n+\delta)}{2}}$
$M = I$	$\mathbb{L}$	Real skew-symmetrics	exp, sin, cos and cosh	$K(f,X) = K_{\mathbb{L}}(f,X)$
$M = I$	$\mathbb{L}$	Real skew-symmetrics	all	$K(f,X) = K_{\mathbb{J}}(f,X)$
$M^TM=I$, $M=\pm M^T$	$\mathbb{L}$		odd	$\frac{K(f,X)}{K_{\mathbb{L}}(f,X)}$ is unbounded
$M^TM=I$, $M=\pm M^T$	$\mathbb{L}$		even	$K(f,X) = K_{\mathbb{L}}(f,X)$
$M^TM=I$, $M=\pm M^T$	$\mathbb{L}$		even	$K(f,X) = K_{\mathbb{J}}(f,X)$

Sesquilinear form $\langle x,y\rangle = x^* My$	$X \in \mathbb{J}$ **or** $\mathbb{L}$	**Matrix class**	**Function** f	**Result**
$M^TM=I$, $M=\pm M^T$	$\mathbb{J}$		all	$K(f,X) = K_{\mathbb{J}}(f,X)$
$M^TM=I$, $M=\pm M^T$	$\mathbb{J}$		all	$K(f,X) = K_{\mathbb{L}}(f,X)$
$M^TM=I$, $M=\pm M^T$	$\mathbb{L}$		odd	$K(f,X) = K_{\mathbb{L}}(f,X)$
$M^TM=I$, $M=\pm M^T$	$\mathbb{L}$		odd	$K(f,X) = K_{\mathbb{J}}(f,X)$
$M^TM=I$, $M=\pm M^T$	$\mathbb{L}$		even	$K(f,X) = K_{\mathbb{L}}(f,X)$
$M^TM=I$, $M=\pm M^T$	$\mathbb{L}$		even	$K(f,X) = K_{\mathbb{J}}(f,X)$

TABLE 6.1

Summary of main results comparing unstructured and structured condition numbers. "all" means all functions that can be written as a convergent power series.

For general f and $X \in \mathbb{L}$ less is known. The matrix $H(X)$ has no observably nice properties to work with. A natural restriction is to consider odd and even functions of $X \in \mathbb{L}$. For even f we have shown that $K(f,X) = K_{\mathbb{L}}(f,X)$. For odd f, the ratio $K(f,X)/K_{\mathbb{L}}(f,X)$ is unbounded. More information about f, X and $\mathbb{L}$ is required to form more meaningful bounds on this ratio.

Acknowledgments. I would like to thank Rafikul Alam, Nick Higham, Niloufer Mackey, Steven Mackey and Françoise Tisseur for their help during the production of this work.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 132-161, June 2004

REFERENCES

[1] I. Gohberg and I. Koltracht. Condition numbers for functions of matrices. *Appl. Numer. Math.*, 12(1-3):107–117, 1993.

[2] Harold V. Henderson and S. R. Searle. The vec-permutation matrix, the vec operator and Kronecker products: A review. *Linear and Multilinear Algebra*, 9(4):271–288, 1980/81.

[3] Desmond J. Higham and Nicholas J. Higham. Backward error and condition of structured linear systems. *SIAM J. Matrix Anal. Appl.*, 13(1):162–175, 1992.

[4] Desmond J. Higham and Nicholas J. Higham. Structured backward error and condition of generalized eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 20(2):493–512, 1998.

[5] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[6] Charles Kenney and Alan J. Laub. Condition estimates for matrix functions. *SIAM J. Matrix Anal. Appl.*, 10(2):191–209, 1989.

[7] D. Steven Mackey. Private communication, December 2003.

[8] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Structured tools for structured matrices. *Electron. J. Linear Algebra*, 10:106–145, 2003.