

STRUCTURED PERTURBATIONS PART I: NORMWISE DISTANCES*

SIEGFRIED M. RUMP[†]

Abstract. In this paper we study the condition number of linear systems, the condition number of matrix inversion, and the distance to the nearest singular matrix, all problems with respect to normwise structured perturbations. The structures under investigation are symmetric, persymmetric, skewsymmetric, symmetric Toeplitz, general Toeplitz, circulant, Hankel, and persymmetric Hankel matrices (some results on other structures such as tridiagonal and tridiagonal Toeplitz matrices, both symmetric and general, are presented as well). We show that for a given matrix the worst case structured condition number for all right-hand sides is equal to the unstructured condition number. For a specific right-hand side we give various explicit formulas and estimations for the condition numbers for linear systems, especially for the ratio of the condition numbers with respect to structured and unstructured perturbations. Moreover, the condition number of matrix inversion is shown to be the same for structured and unstructured perturbations, and the same is proved for the distance to the nearest singular matrix. It follows a generalization of the classical Eckart–Young theorem, namely, that the reciprocal of the condition number is equal to the distance to the nearest singular matrix for all structured perturbations mentioned above.

Key words. normwise structured perturbations, condition number, distance to singularity

AMS subject classifications. 15A12, 65F35

PII. S0895479802405732

1. Motivation. Consider a numerical problem in m input parameters producing k output parameters, that is, a function $f : \mathbb{R}^m \rightarrow \mathbb{R}^k$. An algorithm to solve the problem, i.e., to compute f , in finite precision may be considered as a function \tilde{f} . A finite precision arithmetic for general *real* numbers may be defined to produce the best finite precision approximation to the (exact) real result (with some tie-breaking strategy). This includes the definition of the arithmetic for finite precision numbers. Then, for given input data $p \in \mathbb{R}^m$, the numerical result $\tilde{f}(p)$ will in general be the same for all \tilde{p} in a small neighborhood of p . So we cannot expect more from a numerical algorithm than its producing the exact function value $f(\tilde{p})$ for some \tilde{p} near p . An algorithm with this property is commonly called *backward stable*. For example, the standard method for solving an $n \times n$ dense system of linear equations, namely, Gaussian elimination with partial pivoting, is backward stable.

But is it always possible that $\tilde{f}(p) = f(\tilde{p})$ for some \tilde{p} near p ? Consider the computation in double precision floating point arithmetic according to IEEE standard 754 [30] of the square of a matrix, for example, of $A = \begin{pmatrix} 1+u & 4 \\ 4 & -1 \end{pmatrix}$, where $u = 2^{-52}$ such that 1 and $1+u$ are adjacent floating point numbers. The result is $\tilde{B} = fl(A^2) = \begin{pmatrix} 17 & 4u \\ 4u & 17 \end{pmatrix}$. For a perturbation $\Delta A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ we obtain

$$(A + \Delta A)^2 = \begin{pmatrix} (1+u+\alpha)^2 + (4+\beta)(4+\gamma) & (4+\beta)(u+\alpha+\delta) \\ (4+\gamma)(u+\alpha+\delta) & (4+\beta)(4+\gamma) + (1-\delta)^2 \end{pmatrix}.$$

But $(A+\Delta A)^2 = \tilde{B}$ is impossible for a small perturbation ΔA because this implies, by comparing with \tilde{B}_{11} and \tilde{B}_{22} , that $(1+u+\alpha)^2 = (1-\delta)^2$, so that $u+\alpha = -\delta$ for

*Received by the editors April 17, 2002; accepted for publication (in revised form) by N. J. Higham January 13, 2003; published electronically May 15, 2003.

<http://www.siam.org/journals/simax/25-1/40573.html>

[†]Technical University of Hamburg-Harburg, Schwarzenbergstr. 95, 21071 Hamburg, Germany (rump@tu-harburg.de).

a small perturbation ΔA . But then $(A + \Delta A)_{12} = 0$. In other words, ordinary matrix multiplication yields the best double precision floating point approximation \tilde{B} to the exact result A^2 but is not backward stable. A similar behavior is not uncommon for other structured problems.

Consider, for example, a linear system $Cx = b$ with a circulant matrix C . Many algorithms take advantage of such information, in terms of computing time and storage. In this case only the first row of the matrix and the right-hand side are input to a structured solver, so $m = 2n$ input data are mapped to $k = n$ output data. By nature, a perturbation of the matrix must be a circulant perturbation.

It is easy to find examples of $Cx = b$ such that for a computed solution \tilde{x} it is *likely* that $(C + \Delta C)\tilde{x} \neq b + \Delta b$ for *all* small perturbations ΔC and Δb such that $C + \Delta C$ is a circulant. This happens although, as above, \tilde{x} may be very close to the *exact* solution of the original problem $Cx = b$. The reason is that, in contrast to general linear systems, the space of input data is not rich enough to produce perturbed input data with the desired property. Or, in other words, there is some hidden structure in the result in contradiction to a computed approximation \tilde{x} .

In such a case, about all an algorithm can do in finite precision is to produce some \tilde{x} such that $(C + \Delta C)(\tilde{x} + \Delta x) = b + \Delta b$. In our previous setting this means that for given input data p we require an algorithm \tilde{f} to produce $q = \tilde{f}(p)$ with $q + \Delta q = f(p + \Delta p)$. An algorithm \tilde{f} with this property is called *stable* (more precisely, mixed forward-backward stable) with respect to the distance measure in use [27, section 1.5]. Indeed, there are (normwise) stable algorithms to solve a linear system with circulant matrix [40]. This leads to structured perturbations and structured condition numbers.

There has been substantial interest in algorithms for structured problems in recent years (see, for example, [1, 22, 15, 19, 33, 10, 40, 5] and the literature cited therein). Accordingly, there is growing interest in structured perturbation analysis; cf. [36, 8, 24, 25, 2, 16, 4, 15, 7, 39, 37, 38, 14]. Moreover, different kinds of structured perturbations are investigated in robust and optimal control, for example, the analysis of the μ -number or structured distances [11, 13, 34, 41, 35, 29].

Particularly, many very fast structured solvers have been developed. Frequently, however, perturbation and error analysis for structured solvers are performed with respect to *general* perturbations. This is obviously improvable because usually for a structured solver nothing else but structured perturbations are *possible*.

However, structured perturbations are not as easy to handle, and a perturbation analysis of an algorithm concerning structured perturbations is generally difficult. Before investing too much into solving a problem, it seems wise to estimate its worth. In our case that means estimating the ratio between the structured and the unstructured sensitivities of a problem. For example, it is known that for a symmetric linear system and for normwise distances it makes no difference at all whether matrix perturbations are restricted to symmetric ones or not. In such a case the “usual” (unstructured) perturbation analysis is perfectly sufficient.

Explicit formulas for other structured condition numbers are known, but not too much is known about the ratio between the structured and the unstructured condition numbers. The aim of this two-part paper is to investigate this problem for a number of common (linear) perturbations for linear systems and for matrix inversion. Part I deals with normwise distances and Part II with componentwise distances.

One result of this first part is that for normwise distances, and for structures that are symmetric Toeplitz or circulant, the general (unstructured) condition number of

a linear system may be up to about the square of the structured condition number, much as it is when solving a least squares problem using normal equations rather than some numerically stable method. Although for many structures there seems currently no stable algorithm in sight, that is, stable with respect to structured perturbations, this creates a certain challenge (see also the last section of Part II of this paper).

2. Introduction and notation. Let nonsingular $A \in M_n(\mathbb{R})$ and $x, b \in \mathbb{R}^n$, $x \neq 0$ be given with $Ax = b$. The (normwise) condition number of this linear system with respect to a weight matrix $E \in M_n(\mathbb{R})$ and a weight vector $f \in \mathbb{R}^n$ is defined by

$$(2.1) \quad \kappa_{E,f}(A, x) := \limsup_{\varepsilon \rightarrow 0} \left\{ \frac{\|\Delta x\|}{\varepsilon \|x\|} : (A + \Delta A)(x + \Delta x) = b + \Delta b, \Delta A \in M_n(\mathbb{R}), \right. \\ \left. \Delta b \in \mathbb{R}^n, \|\Delta A\| \leq \varepsilon \|E\|, \|\Delta b\| \leq \varepsilon \|f\| \right\}.$$

In definition (2.1) the parameters E and f are only used as scaling factors and may be replaced by $\|E\|$ and $\|f\|$, respectively. However, in Part II of this paper we treat componentwise perturbations, and there we need the matrix and vector information in E and f . So we use the indices E, f in (2.2) to display certain similarities between normwise and componentwise perturbations.

Throughout this paper we always use the spectral norm $\|\cdot\|_2$, where we denote the matrix norm and the vector norm by the same symbol $\|\cdot\|$. It is well known [27, Theorem 7.2] that

$$(2.2) \quad \kappa_{E,f}(A, x) = \|A^{-1}\| \|E\| + \frac{\|A^{-1}\| \|f\|}{\|x\|}.$$

Note that the (unstructured) condition number does not depend on x but only on $\|x\|$. For no perturbations in the right-hand side is the condition number even independent of x . That means ill-conditioning is a matrix intrinsic property. This will change for structured perturbations.

By definition (2.1), a perturbation of size ε in the input data A and b creates a distortion of size $\kappa \cdot \varepsilon$ in the solution. Therefore, we cannot expect a numerical algorithm to produce an approximation \tilde{x} better than that; that is, $\|\tilde{x} - x\|/\|x\|$ will not be much less than $\kappa \cdot \varepsilon$. On the other hand, we may regard an algorithm to be stable if it produces an approximation \tilde{x} of this quality, i.e., $\|\tilde{x} - x\|/\|x\| \sim \kappa \cdot \varepsilon$.

In case the matrix A has an additional structure such as symmetry or Toeplitz, the structure may be utilized to improve performance of a linear system solver. For example, we have the remarkable fact that the inverse of a (symmetric) Toeplitz matrix can be calculated in $\mathcal{O}(n^2)$ operations, the time it takes to print the entries of the inverse [18, Algorithm 4.7.3].

Usually, such a specialized solver utilizes only part of the input matrix, for example, only the first row in the symmetric Toeplitz case—the other entries are assumed to be defined according to the given structure. This implies that only structured perturbations of the input matrix *are possible*. Perturbations of the input matrix are structured by nature as, for example, symmetric Toeplitz. Accordingly, perturbation theory may use a *structured condition number* defined similarly to (2.1):

$$(2.3) \quad \kappa_{E,f}^{\text{struct}}(A, x) := \limsup_{\varepsilon \rightarrow 0} \left\{ \frac{\|\Delta x\|}{\varepsilon \|x\|} : (A + \Delta A)(x + \Delta x) = b + \Delta b, \Delta A \in M_n^{\text{struct}}(\mathbb{R}), \right. \\ \left. \Delta b \in \mathbb{R}^n, \|\Delta A\| \leq \varepsilon \|E\|, \|\Delta b\| \leq \varepsilon \|f\| \right\}.$$

For other definitions of structured condition numbers see [16] and [17]. The set $M_n^{\text{struct}}(\mathbb{R})$ depicts the set of $n \times n$ real matrices with a certain structure *struct*. In this paper we will investigate the linear structures

$$(2.4) \quad \text{struct} \in \{\text{sym}, \text{persym}, \text{skewsym}, \text{symToep}, \text{Toep}, \text{circ}, \text{Hankel}, \text{persymHankel}\}$$

depicting the set of symmetric, persymmetric, skewsymmetric, symmetric Toeplitz, general Toeplitz, circulant, Hankel, and persymmetric Hankel matrices. In view of (2.3) note that for $A \in M_n^{\text{struct}}(\mathbb{R})$ for any of the structures in (2.4) it is $\Delta A \in M_n^{\text{struct}}(\mathbb{R})$ equivalent to $A + \Delta A \in M_n^{\text{struct}}(\mathbb{R})$. We will derive explicit formulas or estimations for κ^{struct} . Particularly, we will investigate the ratio $\kappa^{\text{struct}}/\kappa$.

Consider, for example, the tridiagonal matrix

$$(2.5) \quad A = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & \ddots & \\ & & & \ddots & \ddots \\ & & & & & 2 \end{pmatrix}.$$

The traditional (unstructured) condition number (2.1), (2.2) for the natural weights $E = A$ and $f = b$ satisfies

$$\kappa_{A, Ax}(A, x) > 4 \cdot 10^{11}$$

for A as in (2.5) of size 10^6 rows and columns and for arbitrary solution x , and hence arbitrary right-hand side. For the specific solution $x = (x_i)$, $x_i = \sin(y_i)$ with y_i equally spaced in the interval $[a, k\pi - a]$ for $a = 13/6000$, $k = 690$, we have

$$(2.6) \quad \kappa_{A, Ax}^{\text{symtridiagToep}}(A, x) < 9.6 \cdot 10^5,$$

where perturbations are symmetric Toeplitz and tridiagonal. Note that in this case the matrix depends only on two parameters. For $x = (1, -1, 1, -1, \dots)^T$ and no perturbations in the right-hand side we get

$$(2.7) \quad \kappa_A^{\text{symtridiagToep}}(A, x) < 0.6.$$

We will derive methods to estimate and compute structured condition numbers. We will especially focus on the ratio $\kappa^{\text{struct}}/\kappa$. We will prove (see Theorem 5.3)

$$\kappa_{E, f}^{\text{struct}}(A, x) = \kappa_{E, f}(A, x) \quad \text{for } \text{struct} \in \{\text{sym}, \text{persym}, \text{skewsym}\}$$

and all $0 \neq x$ in \mathbb{R}^n . This extends a result in [24]. By estimations and examples we show that the ratio can be significantly less than 1 for perturbations subject to the other structures in (2.4). Among others, we will prove (see Theorems 8.4, 9.2, and 10.2)

$$1 \geq \frac{\kappa_{A, Ax}^{\text{struct}}(A, x)}{\kappa_{A, Ax}(A, x)} \geq \frac{1}{2\sqrt{2}\sqrt{\|A^{-1}\| \|A\|}}$$

for $\text{struct} \in \{\text{symToep}, \text{Toep}, \text{circ}, \text{Hankel}, \text{persymHankel}\}$. On the other hand, we will show that to every structure an easy-to-calculate matrix Ψ_x is assigned, depending only on the structure and the solution x , with the surprising result that the ratio $\kappa^{\text{struct}}/\kappa$ can only become small when the smallest singular value $\sigma_{\min}(\Psi_x)$ is small. So

the ratio can only become small for certain *solutions*, independent of the (structured) matrix.

Furthermore, we will investigate the structured condition number for matrix inversion

$$\kappa_E^{\text{struct}}(A) := \limsup_{\varepsilon \rightarrow 0} \left\{ \frac{\|(A + \Delta A)^{-1} - A^{-1}\|}{\varepsilon \|A^{-1}\|} : \Delta A \in M_n^{\text{struct}}(\mathbb{R}), \|\Delta A\| \leq \varepsilon \|E\| \right\}.$$

The definition includes the traditional (unstructured) condition number $\kappa_E(A)$ for matrix inversion by setting $M_n^{\text{struct}}(\mathbb{R}) := M_n(\mathbb{R})$. It is well known that $\kappa_E(A) = \|A^{-1}\| \|E\|$ [27, Theorem 6.4]. Here we will show that

$$\kappa_E^{\text{struct}}(A) = \|A^{-1}\| \|E\| \quad \text{for all structures as in (2.4).}$$

In most cases this is not difficult to prove. However, for Hankel and general Toeplitz perturbations we have to show that

$$\text{for all } x \in \mathbb{R}^n \exists H \in M_n^{\text{Hankel}}(\mathbb{R}) : Hx = x \quad \text{and} \quad \|H\| \leq 1.$$

It seems natural to consider an ill-conditioned matrix to be “almost singular.” Indeed, for normwise and unstructured perturbations the distance to singularity

$$(2.8) \quad \delta_E(A) := \min \left\{ \frac{\|\Delta A\|}{\|E\|} : A + \Delta A \text{ singular} \right\}$$

is well known to be equal to the reciprocal of the condition number (with no perturbation in the right-hand side) [27, Theorem 6.5]:

$$\delta_E(A) = \kappa_E(A)^{-1} = (\|A^{-1}\| \|E\|)^{-1}.$$

We may ask whether this carries over to structured perturbations. The structured (normwise) distance to singularity is defined accordingly by

$$(2.9) \quad \delta_E^{\text{struct}}(A) := \min \left\{ \frac{\|\Delta A\|}{\|E\|} : A + \Delta A \text{ singular}, \Delta A \in M_n^{\text{struct}}(\mathbb{R}) \right\}.$$

Indeed we will show that for all structures (2.4) under consideration δ_E^{struct} is equal to $\kappa_E^{\text{struct}}(A)^{-1}$.

We will use the following notation:

$M_n(\mathbb{R})$	set of real $n \times n$ matrices
$M_n^{\text{struct}}(\mathbb{R})$	set of structured real $n \times n$ matrices, struct as in (2.4)
$\ \cdot\ $	spectral norm
$\ A\ _F$	Frobenius norm $(\sum A_{ij}^2)^{1/2}$
E	some (weight) matrix, $E \in M_n(\mathbb{R})$
f	some (weight) vector, $f \in \mathbb{R}^n$
I, I_n	identity matrix (with n rows and columns)
e	vector of all 1's, $e \in \mathbb{R}^n$
$(\mathbf{1})$	matrix of all 1's, $(\mathbf{1}) = ee^T \in M_n(\mathbb{R})$
J, J_n	permutation matrix mapping $(1, \dots, n)^T$ into $(n, \dots, 1)^T$
$\sigma_{\min}(A)$	smallest singular value of A
$\lambda_{\min}(A)$	smallest eigenvalue of symmetric A

3. Normwise perturbations. Throughout this paper we let nonsingular $A \in M_n(\mathbb{R})$ be given together with $0 \neq x \in \mathbb{R}^n$. Denote $b := Ax$ and let $E \in M_n(\mathbb{R})$, $f \in \mathbb{R}^n$.

We first prove (2.2) in a way which is suitable for general as well as structured perturbations. The standard proof [27, Theorem 7.2] for (2.2) uses the fact that $Ax = b$ and $(A + \Delta A)(x + \Delta x) = b + \Delta b$ imply

$$(3.1) \quad \Delta x = A^{-1}(-\Delta Ax + \Delta b) + \mathcal{O}(\varepsilon^2).$$

For given ΔA with $\|\Delta A\| \leq \varepsilon\|E\|$ define $\Delta b := -\frac{\|f\|}{\|E\|\|x\|}\Delta Ax$. Then $\|\Delta b\| \leq \varepsilon\|f\|$, and (3.1) implies

$$(3.2) \quad \Delta x = -A^{-1}\Delta Ax \left(1 + \frac{\|f\|}{\|E\|\|x\|}\right) + \mathcal{O}(\varepsilon^2).$$

This is satisfied for arbitrary ΔA with $\|\Delta A\| \leq \varepsilon\|E\|$, the perturbations ΔA being structured or unstructured. This gives a reason for the following definition.

DEFINITION 3.1. For nonsingular $A \in M_n(\mathbb{R})$, $0 \neq x \in \mathbb{R}^n$, and $M_n^{\text{struct}}(\mathbb{R}) \subseteq M_n(\mathbb{R})$ we define

$$\varphi^{\text{struct}}(A, x) := \sup\{\|A^{-1}\Delta Ax\| : \Delta A \in M_n^{\text{struct}}, \|\Delta A\| \leq 1\}.$$

For $M_n^{\text{struct}}(\mathbb{R}) = M_n(\mathbb{R})$ we omit the superindex *struct*: $\varphi(A, x)$.

Now the special choice of Δb that led to (3.2) and the definition (2.3) imply

$$(3.3) \quad \frac{\varphi^{\text{struct}}(A, x)}{\|x\|} \left(\|E\| + \frac{\|f\|}{\|x\|}\right) \leq \kappa_{E,f}^{\text{struct}}(A, x)$$

for all $M_n^{\text{struct}}(\mathbb{R}) \subseteq M_n(\mathbb{R})$. Furthermore, an obvious norm estimation using (2.3) and (3.1) yields

$$(3.4) \quad \kappa_{E,f}^{\text{struct}}(A, x) \leq \|A^{-1}\| \|E\| + \|A^{-1}\| \frac{\|f\|}{\|x\|},$$

again for all $M_n^{\text{struct}}(\mathbb{R}) \subseteq M_n(\mathbb{R})$. Therefore, we have equality in (3.4) if $\varphi^{\text{struct}}(A, x) = \|A^{-1}\| \|x\|$. This is true (and well known) for unstructured perturbations

$$(3.5) \quad \varphi(A, x) = \|A^{-1}\| \|x\|$$

by choosing orthogonal ΔA with $\Delta Ax = \|x\|y$ for $\|A^{-1}\| = \|A^{-1}y\|$ and $\|y\| = 1$.

THEOREM 3.2. For nonsingular $A \in M_n(\mathbb{R})$, $0 \neq x \in \mathbb{R}^n$, and $M_n^{\text{struct}}(\mathbb{R}) \subseteq M_n(\mathbb{R})$ we have

$$(3.6) \quad \frac{\varphi^{\text{struct}}(A, x)}{\|x\|} \left(\|E\| + \frac{\|f\|}{\|x\|}\right) \leq \kappa_{E,f}^{\text{struct}}(A, x) \leq \|A^{-1}\| \|E\| + \|A^{-1}\| \frac{\|f\|}{\|x\|}.$$

Particularly, $\varphi^{\text{struct}}(A, x) = \|A^{-1}\| \|x\|$ implies

$$\kappa_{E,f}^{\text{struct}}(A, x) = \kappa_{E,f}(A, x) = \|A^{-1}\| \|E\| + \|A^{-1}\| \frac{\|f\|}{\|x\|}.$$

As we will see, the latter equality is true for symmetric, skewsymmetric, and persymmetric perturbations. For other perturbations the lower bound in (3.6) is usually too weak because $\varphi^{\text{struct}}(A, x)$ can be much less than $\|A^{-1}\| \|x\|$. An immediate upper bound by (2.3) and (3.1) is

$$(3.7) \quad \kappa_{E,f}^{\text{struct}}(A, x) \leq \varphi^{\text{struct}}(A, x) \frac{\|E\|}{\|x\|} + \|A^{-1}\| \frac{\|f\|}{\|x\|}.$$

Although we are free in the perturbations Δb , the structure in ΔA may not allow equality in (3.7). However, for $u, v \in \mathbb{R}^n$ it is $\max(\|u+v\|, \|u-v\|) \geq \sqrt{\|u\|^2 + \|v\|^2} \geq 2^{-1/2}(\|u\| + \|v\|)$ such that

$$u, v \in \mathbb{R}^n \quad \text{implies} \quad \max(\|u+v\|, \|u-v\|) = c(\|u\| + \|v\|),$$

where $2^{-1/2} \leq c \leq 1$. We are free in choosing the sign of Δb , so (3.7), $u = -A^{-1}\Delta Ax$, $v = A^{-1}\Delta b$ together with (3.1) imply the following result.

THEOREM 3.3. *Let $A \in M_n(\mathbb{R})$, $0 \neq x \in \mathbb{R}^n$, and $M_n^{\text{struct}}(\mathbb{R}) \subseteq M_n(\mathbb{R})$ be given. Then the structured (normwise) condition number as defined in (2.3) satisfies*

$$(3.8) \quad \kappa_{E,f}^{\text{struct}}(A, x) = c \cdot \left[\varphi^{\text{struct}}(A, x) \frac{\|E\|}{\|x\|} + \|A^{-1}\| \frac{\|f\|}{\|x\|} \right],$$

where $2^{-1/2} \leq c \leq 1$. For no perturbations in the right-hand side we have

$$\kappa_E^{\text{struct}}(A, x) = \varphi^{\text{struct}}(A, x) \frac{\|E\|}{\|x\|}.$$

This moves our focus from analysis of structured condition numbers to the analysis of $\varphi^{\text{struct}}(A, x)$. In the following we will use Definition 3.1 of φ^{struct} together with Theorems 3.2 and 3.3 to establish formulas and bounds for structured condition numbers.

4. Condition number for general x . For general perturbations and for the natural choice $E = A$, $f = b$, we have $\frac{\|A^{-1}\| \|b\|}{\|x\|} \leq \|A^{-1}\| \|A\|$ such that (2.2) yields

$$(4.1) \quad \kappa_A(A, x) = \|A^{-1}\| \|A\| \leq \kappa_{A, Ax}(A, x) \leq 2\|A^{-1}\| \|A\|.$$

In other words, in case of general perturbations it does not make a big difference whether we allow perturbations in the right-hand side or leave it unchanged. Moreover, the general condition number $\kappa_A(A, x)$ is independent of x . So the condition is an inherent property of the matrix.

This may change in case of structured condition numbers. A first result in this respect is that for all structures (2.4) the *worst case* structured (normwise) condition number, i.e., the supremum over all x , is equal to the worst case unstructured condition number.

THEOREM 4.1. *Let nonsingular $A \in M_n(\mathbb{R})$ be given and $M^{\text{struct}} \subseteq M_n(\mathbb{R})$ such that one of the following conditions is satisfied:*

- (i) $I \in M_n^{\text{struct}}(\mathbb{R})$.
- (ii) $J \in M_n^{\text{struct}}(\mathbb{R})$.
- (iii) $\tilde{J} := \begin{pmatrix} 0 & J \\ -I & 0 \end{pmatrix} \in M_n^{\text{struct}}(\mathbb{R})$ in case n even.

Let fixed $0 < \gamma \in \mathbb{R}$ be given. Then for all $\|x\| = \gamma$,

$$(4.2) \quad \sup_{\|y\|=\gamma} \kappa_{E,f}^{\text{struct}}(A, y) = \kappa_{E,f}(A, x) = \|A^{-1}\| \|E\| + \|A^{-1}\| \frac{\|f\|}{\|x\|},$$

so the worst case structured condition number is equal to the general condition number. Equation (4.2) is especially true for all structures in (2.4). It is of course also true for $M_n^{\text{struct}}(\mathbb{R}) = M_n(\mathbb{R})$.

Remark 4.2. Note that a nonsingular skewsymmetric matrix must be of even order.

Proof. Let $\|A^{-1}\| = \|A^{-1}y\|$ with $\|y\| = 1$. Choosing $\Delta A = I$, $\Delta A = J$, or $\Delta A = \bar{J}$ in case (i), (ii), or (iii), respectively, observing $\Delta A^2 = \pm I$, and setting $x := \gamma \Delta A y$ imply $A^{-1} \Delta A x = \pm \gamma A^{-1} y$ and $\|x\| = \gamma$. Hence Definition 3.1 yields $\|A^{-1}\| \|x\| \geq \varphi^{\text{struct}}(A, x) \geq \|A^{-1}\| \|x\|$ for that choice of x , and Theorem 3.2 finishes the proof. \square

For specific x things may change significantly, at least if the structure imposes severe restrictions on ΔA . For symmetric, persymmetric, and skewsymmetric structures this is not yet the case.

5. Symmetric, persymmetric, and skewsymmetric perturbations.

In the following we will show that those perturbations do not change the condition number at all. For symmetric perturbations this was already observed in [24]; see also [8]. In other words, “worst” perturbations may be chosen in the set $M_n^{\text{sym}}(\mathbb{R})$, $M_n^{\text{persym}}(\mathbb{R})$, or $M_n^{\text{skewsym}}(\mathbb{R})$. We prove this by investigating our key to structured perturbations, the function φ^{struct} . We first prove a lemma which will be of later use. For the symmetric case this was observed in [8].

LEMMA 5.1. *Let $x, y \in \mathbb{R}^n$ be given with $\|x\| = \|y\| = 1$ and let $\text{struct} \in \{\text{sym}, \text{persym}\}$. Then there exists $A \in M_n^{\text{struct}}(\mathbb{R})$ with*

$$(5.1) \quad y = Ax \quad \text{and} \quad \|A\| = 1.$$

If, in addition, $y^T x = 0$, then there exists $A \in M_n^{\text{skewsym}}(\mathbb{R})$ with (5.1).

Proof. For symmetric structure the Householder reflection H along $x + y$ satisfies $H = H^T$, $\|H\| = 1$, and $Hx = y$. A matrix B is persymmetric iff $B = JB^T J$. Let H be the Householder reflection along $x + Jy$ and set $A := JH$. Then $A = JA^T J$ is persymmetric, $\|A\| = 1$, and $Ax = JHx = J \cdot Jy = y$.

For skewsymmetric structure and x, y orthonormal there is orthogonal $Q \in M_n(\mathbb{R})$ with $[x|y] = Q[e_1 | -e_2]$, e_i denoting the i th column of the identity matrix. Define $D := \text{diag}(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, 0, \dots, 0)$ and $A := QDQ^T$. Then $A = -A^T$, $\|A\| = 1$, and $Ax = QDe_1 = -Qe_2 = y$. \square

LEMMA 5.2. *Let nonsingular $A \in M_n(\mathbb{R})$ and $0 \neq x \in \mathbb{R}^n$ be given. Then*

$$(5.2) \quad \varphi^{\text{struct}}(A, x) = \varphi(A, x) = \|A^{-1}\| \|x\|$$

for $\text{struct} \in \{\text{sym}, \text{persym}\}$. Relation (5.2) is also true for $\text{struct} = \text{skewsym}$ and $A \in M_n^{\text{skewsym}}$.

Proof. By Definition 3.1 and (3.5), $\varphi^{\text{struct}}(A, x) \leq \varphi(A, x) = \|A^{-1}\| \|x\|$, so it remains to show $\varphi^{\text{struct}}(A, x) \geq \|A^{-1}\| \|x\|$. Without loss of generality, assume $\|x\| = 1$ and let $\|A^{-1}\| = \|A^{-1}y\|$ for $\|y\| = 1$. It suffices to find $\Delta A \in M^{\text{struct}}$ with $\|\Delta A\| \leq 1$ and $\Delta Ax = y$. This is exactly the content of Lemma 5.1 for $\text{struct} \in \{\text{sym}, \text{persym}\}$.

For skewsymmetric structure suppose $A \in M^{\text{skewsym}}$. Eigenvalues of A are conjugate purely imaginary, and nonsingularity of A implies that n is even, and also implies that all singular values are of even multiplicity. That means there are orthogonal $y_1, y_2 \in \mathbb{R}^n$ with $\|y_1\| = \|y_2\| = 1$ and $\|A^{-1}y_1\| = \|A^{-1}y_2\| = \|A^{-1}\|$. Choose $y \in \text{span}\{y_1, y_2\}$ with $x^T y = 0$ and $\|y\| = 1$. By construction, $\|A^{-1}y\| = \|A^{-1}\|$, and Lemma 5.1 finishes the proof. \square

Together with Theorem 3.2 this proves the following.

THEOREM 5.3. *Let nonsingular $A \in M_n(\mathbb{R})$ and $0 \neq x \in \mathbb{R}^n$ be given. For $\text{struct} \in \{\text{sym}, \text{persym}, \text{skewsym}\}$ we have*

$$\kappa_{E,f}^{\text{struct}}(A, x) = \kappa_{E,f}(A, x),$$

where in case $\text{struct} = \text{skewsym}$ we suppose additionally $A \in M_n^{\text{skewsym}}(\mathbb{R})$.

The result was observed for symmetric structures in [24, 23]. As we will see, this nice fact is no longer true for the other structures. In fact, there may be quite a factor between κ^{struct} and κ .

6. Exploring the structure. Before we proceed we collect some general observations on structured condition numbers. To establish bounds for the ratio $\kappa^{\text{struct}}/\kappa$ we need a relation between $\|E\|$ and $\|f\|$. Therefore we especially investigate the natural choice $E = A$ and $f = b$. The first statement is a useful lower bound.

LEMMA 6.1. *Let nonsingular $A \in M_n(\mathbb{R})$, $0 \neq x \in \mathbb{R}^n$, and some $M_n^{\text{struct}}(\mathbb{R}) \subseteq M_n(\mathbb{R})$ be given. Suppose*

$$(6.1) \quad \varphi^{\text{struct}}(A, x) \geq \omega \|A^{-T}x\|$$

for $0 \leq \omega \in \mathbb{R}$. Then

$$\kappa_{A,Ax}^{\text{struct}}(A, x) \geq \sqrt{\frac{\omega}{2} \|A^{-1}\| \|A\|}.$$

Proof. Without loss of generality assume $\|x\| = 1$. Then

$$(6.2) \quad 1 = x^T A^{-1}Ax \leq \|x^T A^{-1}\| \|Ax\| = \|A^{-T}x\| \|Ax\|.$$

In view of (3.8) for $E = A$, $f = b$, $\|x\| = 1$, and $Ax = b$, we are finished if we can show

$$\varphi^{\text{struct}}(A, x) \|A\| + \|A^{-1}\| \|Ax\| \geq \sqrt{\omega \|A^{-1}\| \|A\|}.$$

This is true if $\|Ax\| \geq \sqrt{\omega \|A\| \|A^{-1}\|}$. On the contrary, (6.2) yields $\|A^{-T}x\| \geq \|Ax\|^{-1} > \sqrt{\omega^{-1} \|A^{-1}\| \|A\|}$, and combining this with (6.1) finishes the proof. \square

The symmetric Toeplitz matrices are related to persymmetric Hankel matrices by

$$(6.3) \quad T \in M_n^{\text{symToep}} \Leftrightarrow JT \in M_n^{\text{persymHankel}} \Leftrightarrow TJ \in M_n^{\text{persymHankel}}.$$

Similarly, (general) Toeplitz matrices are related to general Hankel matrices by

$$(6.4) \quad T \in M_n^{\text{Toep}} \Leftrightarrow JT \in M_n^{\text{Hankel}} \Leftrightarrow TJ \in M_n^{\text{Hankel}}.$$

By rewriting (3.1) into

$$\Delta x = (JA)^{-1}(-J\Delta Ax + J\Delta b) + \mathcal{O}(\varepsilon^2)$$

and

$$J\Delta x = (AJ)^{-1}(-\Delta AJ \cdot Jx + \Delta b) + \mathcal{O}(\varepsilon^2)$$

and observing $\|J\Delta A\| = \|\Delta AJ\| = \|\Delta A\|$ and $\|J\Delta b\| = \|\Delta b\|$, definition (2.3) yields the following.

THEOREM 6.2. *For nonsingular $A \in M_n(\mathbb{R})$ and $0 \neq x \in \mathbb{R}^n$ we have*

$$\kappa_{E,f}^{\text{symToep}}(A, x) = \kappa_{E,f}^{\text{persymHankel}}(JA, x) = \kappa_{E,f}^{\text{persymHankel}}(AJ, Jx)$$

and

$$\kappa_{E,f}^{\text{Toep}}(A, x) = \kappa_{E,f}^{\text{Hankel}}(JA, x) = \kappa_{E,f}^{\text{Hankel}}(AJ, Jx).$$

Therefore we will concentrate in the following on symmetric Toeplitz and Hankel structures. Every result for those is valid mutatis mutandis for persymmetric Hankel and general Toeplitz structures, respectively.

To further explore the structure we derive two-sided explicit bounds for $\varphi^{\text{struct}}(A, x)$. For linear structures in the matrix entries of $A \in M_n(\mathbb{R})$, every A_{ij} depends linearly on some k parameters. Denote by $\text{vec}(A) = (A_{11}, \dots, A_{1n}, \dots, A_{n1}, \dots, A_{nn})^T \in \mathbb{R}^{n^2}$ the vector of stacked columns of A . Then for every dimension there is some fixed structure matrix $\Phi^{\text{struct}} \in M_{n^2, k}(\mathbb{R})$ such that

$$(6.5) \quad A \in M_n^{\text{struct}}(\mathbb{R}) \Leftrightarrow \exists p \in \mathbb{R}^k : \text{vec}(A) = \Phi^{\text{struct}} \cdot p.$$

This idea was developed in [24]. For our structures (2.4) the number of independent parameters k is as shown in Table 6.1.

TABLE 6.1
Number of independent parameters.

Structure	sym	persym	skewsym	circ	symToep	Toep	Hankel	persymHankel
k	$(n^2 + n)/2$	$(n^2 + n)/2$	$(n^2 - n)/2$	n	n	$2n - 1$	$2n - 1$	n

For the structures in (2.4) the structure matrix Φ^{struct} is sparse with entries 0/1 except for skewsymmetric matrices with entries 0/+1/-1. We can make Φ^{struct} unique by defining the parameter vector ‘‘columnwise’’; i.e., $p \in \mathbb{R}^k$ is the unique vector of the first k independent components in $\text{vec}(A)$.

It is important to note that Φ^{struct} defines for every dimension n a one-to-one mapping between \mathbb{R}^k and $M_n^{\text{struct}}(\mathbb{R})$. To compute bounds on φ^{struct} we relate the matrix norm $\|A\|_2$ to the vector norm $\|p\|_2$.

LEMMA 6.3. *Let $A \in M_n^{\text{struct}}(\mathbb{R})$ and $p \in \mathbb{R}^k$ be given such that $\text{vec}(A) = \Phi^{\text{struct}} p$. Then*

$$(6.6) \quad \alpha \|A\| \leq \|p\| \leq \beta \|A\|$$

with constants α, β according to the following table:

Structure	α	β
circ	$1/\sqrt{n}$	1
symToep	$1/\sqrt{2n-2}$	1
Toep	$1/\sqrt{n}$	$\sqrt{2}$
Hankel	$1/\sqrt{n}$	$\sqrt{2}$
persymHankel	$1/\sqrt{2n-2}$	1

All upper bounds and the lower bound for circulants are sharp, and the other lower bounds are sharp up to a factor $\sqrt{2}$.

Proof. For $A \in M_n^{\text{circ}}$ we have

$$\|p\| = \|Ae_1\| \leq \|A\| \leq \|A\|_F = \left(n \sum p_i^2\right)^{1/2} = \sqrt{n}\|p\|.$$

The left and right estimations are sharp for $A = I$ and $A = (\mathbf{1})$, respectively. For $A \in M_n^{\text{symToep}}$,

$$\|p\| = \|Ae_1\| \leq \|A\| \leq \|A\|_F \leq \left((2n-2) \sum p_i^2\right)^{1/2} = \sqrt{2n-2}\|p\|.$$

For $A = I$ it is $\|A\| = \|p\| = 1$, and for $A = (\mathbf{1})$ it is $\|A\| = \sqrt{n}\|p\| = n$. For $A \in M_n^{\text{Hankel}}$ we have

$$\|p\|^2 \leq 2 \max(\|Ae_1\|^2, \|e_1^T A\|^2) \leq 2\|A\|^2$$

and

$$\|A\| \leq \|A\|_F \leq \left(n \sum p_i^2\right)^{1/2} = \sqrt{n}\|p\|.$$

For $A = (\mathbf{1})$ it is $\|A\| = n = \frac{n}{\sqrt{2n-1}}\|p\|$, and for the Hankel matrix with $A_{11} = A_{nn} = 1$ and zero entries elsewhere it is $\|p\| = \sqrt{2} = \sqrt{2}\|A\|$. The other estimations follow by (6.3) and (6.4). \square

The bounds for circulants are noted for completeness; we will derive better methods to estimate $\kappa_{E,f}^{\text{circ}}$ in the next section. The difficulty in estimating $\varphi^{\text{struct}} = \sup\{\|A^{-1}\Delta Ax\| : \Delta A \in M^{\text{struct}}, \|\Delta A\| \leq 1\}$ is that the supremum is taken only over structured matrices ΔA . With Lemma 6.3 this can be rewritten to the supremum over *all* parameter vectors $\Delta p \in \mathbb{R}^k$, $\|\Delta p\| \leq \text{const}$, where k is the number of independent parameters according to Table 6.1 and const follows by Lemma 6.3. We have

$$(6.7) \quad \begin{aligned} & \{\Delta A \in M_n(\mathbb{R}) : \text{vec}(\Delta A) = \Phi^{\text{struct}} \Delta p, \quad \Delta p \in \mathbb{R}^k, \|\Delta p\| \leq \alpha\} \\ & \subseteq \{\Delta A \in M_n^{\text{struct}}(\mathbb{R}) : \|\Delta A\| \leq 1\} \\ & \subseteq \{\Delta A \in M_n(\mathbb{R}) : \text{vec}(\Delta A) = \Phi^{\text{struct}} \Delta p, \quad \Delta p \in \mathbb{R}^k, \|\Delta p\| \leq \beta\}, \end{aligned}$$

where Δp varies freely in a norm ball of the \mathbb{R}^k . So (6.7) is the key to obtaining computable lower and upper bounds for the structured condition number, the bounds not being far apart.

To estimate $\varphi^{\text{struct}}(A, x)$ we use the following ansatz as in [24]. Note that $\Delta A \cdot x = (x^T \otimes I) \text{vec}(\Delta A)$, \otimes denoting the Kronecker product. For $\text{vec}(\Delta A) = \Phi^{\text{struct}} \Delta p$ this implies

$$(6.8) \quad \Delta A \cdot x = (x^T \otimes I) \Phi^{\text{struct}} \cdot \Delta p.$$

The matrix $(x^T \otimes I) \Phi^{\text{struct}} \in M_{n,k}(\mathbb{R})$ depends only on x for every dimension. This leads us to the definition

$$(6.9) \quad \Psi_x^{\text{struct}} := (x^T \otimes I) \Phi^{\text{struct}} \in M_{n,k}(\mathbb{R}),$$

the dimension k as in Table 6.1. This definition holds for every linear structure. For the structures in (2.4), the matrices Ψ_x^{struct} can be calculated explicitly. For example, for the Hankel matrix

$$H = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$$

we have $\Phi^{\text{Hankel}} \in M_{n^2, k} = M_{9,5}$, a column block matrix with n blocks $\Phi_i \in M_{n, k}$, $1 \leq i \leq n$, and

$$\Phi_i = \begin{pmatrix} 0 & \dots & 0 & 1 & & & 0 & \dots & 0 \\ & & & & \ddots & & & & \\ 0 & \dots & 0 & & & 1 & & & \\ \underbrace{0 & \dots & 0}_{i-1} & & & & & \underbrace{0 & \dots & 0}_{n-i} \end{pmatrix} \in M_{3,5},$$

so that $\Psi_x^{\text{struct}} = (x^T \otimes I)\Phi^{\text{struct}}$ implies

$$(6.10) \quad \Psi_x^{\text{Hankel}} = \sum x_i \Phi_i = \begin{pmatrix} x_1 & x_2 & x_3 & & & \\ & x_1 & x_2 & x_3 & & \\ & & x_1 & x_2 & x_3 & \\ & & & x_1 & x_2 & x_3 \end{pmatrix} \in M_{n, k}.$$

We mention

$$(6.11) \quad \begin{aligned} \Psi_x^{\text{circ}} &= \text{circ}(x^T)^T, \\ \Psi_x^{\text{symToep}} &= T(x) + JT(Jx) - xe_1^T, \\ \Psi_x^{\text{Hankel}} &= \text{Toeplitz}([x_1, z], [x^T, z]), \end{aligned}$$

where $z = \text{zeros}(1, n-1)$ and $T(x) := \text{Toeplitz}(x, [x_1, z])$ in Matlab [32] notation; that is, $\text{Toeplitz}(c, r)$ denotes the Toeplitz matrix with first column c and first row r . With this we have explicit bounds for φ^{struct} .

LEMMA 6.4. *Let nonsingular $A \in M_n(\mathbb{R})$ and $0 \neq x \in \mathbb{R}^n$ be given. Let struct be one of the structures mentioned in Lemma 6.3. Then*

$$\varphi^{\text{struct}}(A, x) = \gamma \|A^{-1} \Psi_x^{\text{struct}}\|,$$

where $\alpha \leq \gamma \leq \beta$ and α, β as in Lemma 6.3.

Proof. Combining (6.9), (6.8), and (6.7) with Definition 3.1 yields

$$\begin{aligned} \alpha \|A^{-1} \Psi_x^{\text{struct}}\| &= \sup\{\|A^{-1} \Psi_x^{\text{struct}} \Delta p\| : \Delta p \in \mathbb{R}^k, \|\Delta p\| \leq \alpha\} \\ &\leq \sup\{\|A^{-1} \Delta A x\| : \Delta A \in M_n^{\text{struct}}(\mathbb{R}), \|\Delta A\| \leq 1\} \\ &= \varphi^{\text{struct}}(A, x) \\ &\leq \sup\{\|A^{-1} \Psi_x^{\text{struct}} \Delta p\| : \Delta p \in \mathbb{R}^k, \|\Delta p\| \leq \beta\} \\ &= \beta \|A^{-1} \Psi_x^{\text{struct}}\|. \quad \square \end{aligned}$$

Combining this with Theorem 3.3 yields computable bounds for the structured condition number.

THEOREM 6.5. *Let nonsingular $A \in M_n(\mathbb{R})$ and $0 \neq x \in \mathbb{R}^n$ be given. Let struct be one of the structures mentioned in Lemma 6.3. Then*

$$(6.12) \quad \kappa_{E, f}^{\text{struct}}(A, x) = c \frac{\gamma \|A^{-1} \Psi_x^{\text{struct}}\| \|E\| + \|A^{-1}\| \|f\|}{\|x\|},$$

where $2^{-1/2} \leq c \leq 1$ and $\alpha \leq \gamma \leq \beta$ for α, β as in Lemma 6.3. In case of no perturbations in the right-hand side,

$$(6.13) \quad \kappa_E^{\text{struct}}(A, x) = \gamma \frac{\|A^{-1}\Psi_x^{\text{struct}}\|}{\|x\|} \|E\|.$$

This implies the following remarkable property of the ratio between the structured and unstructured condition numbers.

COROLLARY 6.6. *Let nonsingular $A \in M_n(\mathbb{R})$ and $0 \neq x \in \mathbb{R}^n$ be given. Let struct be one of the structures mentioned in Lemma 6.3. Then*

$$(6.14) \quad \frac{\kappa_{E,f}^{\text{struct}}(A, x)}{\kappa_{E,f}(A, x)} \geq 2^{-1/2} \frac{\alpha \|A^{-1}\| \frac{\sigma_{\min}(\Psi_x^{\text{struct}})}{\|x\|} \|E\| + \|A^{-1}\| \frac{\|f\|}{\|x\|}}{\|A^{-1}\| \|E\| + \|A^{-1}\| \frac{\|f\|}{\|x\|}},$$

for α as in Lemma 6.3. Moreover, for no perturbations in the right-hand side,

$$(6.15) \quad 2^{1/2} \frac{\kappa_{E,f}^{\text{struct}}(A, x)}{\kappa_{E,f}(A, x)} \geq \frac{\kappa_E^{\text{struct}}(A, x)}{\kappa_E(A, x)} \geq \alpha \frac{\sigma_{\min}(\Psi_x^{\text{struct}})}{\|x\|}.$$

Proof. We have $\Psi_x^{\text{struct}} \in M_{n,k}(\mathbb{R})$ with $k \geq n$; therefore $\|A^{-1}\Psi_x^{\text{struct}}\| \geq \|A^{-1}\| \sigma_{\min}(\Psi_x^{\text{struct}})$. Now (2.2) and Theorem 6.5 finish the proof. \square

This result allows us to estimate the minimum ratio of $\kappa^{\text{struct}}/\kappa$ independent of the matrix A only by examining the smallest singular value of Ψ_x^{struct} , where the latter can be computed, for example, by (6.11). So we have the surprising result that a small ratio $\kappa^{\text{struct}}/\kappa$ is only possible for certain *solutions* x , independent of the (structured) matrix. It also shows that for fixed x an arbitrarily small ratio of $\kappa^{\text{struct}}/\kappa$ is only possible if $\text{rank}(\Psi_x^{\text{struct}}) < n$. From a practical point of view this means that standard unstructured perturbation analysis suffices at least for all cases where $\sigma_{\min}(\Psi_x^{\text{struct}})$ is not too small.

The statistics in Table 6.2 show how often a small ratio $\kappa^{\text{struct}}/\kappa$ can occur. Note that this is a lower estimate of the ratio for all matrices A ; it need not be attained for a specific matrix A . Table 6.2 shows the minimum and median of $\tau(x) := \sigma_{\min}(\Psi_x^{\text{struct}})/\|x\|$ for some 10^4 samples of x with entries uniformly distributed within $[-1, 1]$. Also note that, in order to obtain the lower estimate for $\kappa^{\text{struct}}/\kappa$, by (6.15) the displayed numbers have to be multiplied by α according to the table in Lemma 6.3.

TABLE 6.2
Minimum value and median of $\tau(x) = \sigma_{\min}(\Psi_x^{\text{struct}})/\|x\|$.

	Symmetric Toeplitz		Circulant	
	$\min(\tau(x))$	$\text{median}(\tau(x))$	$\min(\tau(x))$	$\text{median}(\tau(x))$
n	$2.4 \cdot 10^{-7}$	$4.9 \cdot 10^{-2}$	$1.3 \cdot 10^{-5}$	$2.6 \cdot 10^{-1}$
10	$5.0 \cdot 10^{-7}$	$2.2 \cdot 10^{-2}$	$3.9 \cdot 10^{-5}$	$2.0 \cdot 10^{-1}$
20	$1.1 \cdot 10^{-6}$	$8.8 \cdot 10^{-3}$	$3.1 \cdot 10^{-5}$	$1.4 \cdot 10^{-1}$
50	$1.5 \cdot 10^{-6}$	$4.3 \cdot 10^{-3}$	$7.5 \cdot 10^{-5}$	$1.0 \cdot 10^{-1}$

Table 6.2 shows that small ratios are possible but seem to be rare. We mention that rank-deficient Ψ_x^{struct} is possible, for example, for $x = (1, \dots, 1)^T$ and $n \geq 2$. That means that for this solution vector x the ratio $\kappa^{\text{struct}}/\kappa$ may become arbitrarily small. This is indeed the case, as we will see in the following sections. However, it changes for Hankel structures, as we will show in section 10.

Therefore Theorem 6.5 implies

$$|(c+d)^{-1}| \leq \kappa_A^{\text{symtridiagToep}}(A, x) \leq \sqrt{10}|(c+d)^{-1}|.$$

Note that this is true for every x of the structure as defined above. For the matrix as in (2.5) this means

$$1 \leq \kappa_A^{\text{symtridiagToep}}(A, x) \leq \sqrt{10}$$

for every x as above, whereas, for $d+2c=0$,

$$\kappa_A(A, x) = \|A^{-1}\| \|A\| \sim n^2.$$

For a general tridiagonal Toeplitz matrix A with diagonal element d and off-diagonal elements c, e we have $\|p\| = \sqrt{c^2 + d^2 + e^2}$. Furthermore, $\|Ax\| \leq (|c| + |d| + |e|)\|x\|$ for every $x \in \mathbb{R}^n$ and therefore

$$\frac{\|A\|}{\|p\|} \leq \frac{|c| + |d| + |e|}{\sqrt{c^2 + d^2 + e^2}} \leq \sqrt{3}.$$

The estimation is asymptotically sharp for $c = d = e = 1$.

For x being the second column of the identity matrix and $n \geq 3$ it follows that

$$\frac{\|A\|}{\|p\|} \geq \frac{\sqrt{c^2 + d^2 + e^2}}{\|p\|} = 1.$$

This estimation is sharp for $A = I$. For $n = 2$ it is $A = \begin{pmatrix} d & e \\ c & d \end{pmatrix}$ and $\|A\|/\|p\| \geq 1/\sqrt{2}$. This estimation is sharp for $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

For symmetric tridiagonal A with diagonal elements d_ν and off-diagonal elements c_ν , we have $\|p\| = \sqrt{\|c\|^2 + \|d\|^2}$ and $\|A\|_F = \sqrt{2\|c\|^2 + \|d\|^2}$. This implies

$$\frac{\|A\|}{\|p\|} \geq \frac{1}{\sqrt{n}} \frac{\|A\|_F}{\|p\|} \geq \frac{1}{\sqrt{n}}$$

and

$$\frac{\|A\|}{\|p\|} \leq \max_{0 \leq x, y \leq 1} \frac{\sqrt{2x^2 + y^2}}{\sqrt{x^2 + y^2}} \leq \sqrt{2}.$$

The first estimation is sharp for $A = I$, the second up to a small factor. Finally, for general tridiagonal A we have $\|A\|_F = \|p\|$, so

$$\frac{1}{\sqrt{n}} \|p\| \leq \|A\| \leq \|p\|.$$

The estimations are sharp for $A = I$ and the matrix with $A_{11} = 1$ and $A_{ij} = 0$ elsewhere, respectively.

Summarizing, we have the following result.

THEOREM 7.1. *Let $A \in M_n^{\text{struct}}(\mathbb{R})$ and $p \in \mathbb{R}^k$ be given such that $\text{vec}(A) = \Phi^{\text{struct}} p$. Then*

$$\alpha \|p\| \leq \|A\| \leq \beta \|p\|$$

with constants α, β according to the following table:

Structure	α	β
<i>symtridiagToep</i>	1	$\sqrt{5}$
<i>tridiagToep</i>	$\frac{1}{\sqrt{2}}$	$\sqrt{3}$ <i>for</i> $n \neq 2$ <i>for</i> $n = 2$
<i>symtridiag</i>	$1/\sqrt{n}$	1
<i>tridiag</i>	$1/\sqrt{n}$	1

All lower bounds are sharp; all upper bounds are sharp up to a small constant factor.

Using the constants α, β and Theorem 6.5 the structured condition numbers are easily calculated.

Also, linear structures in the right-hand side can be treated by an augmented linear system of dimension $n + 1$. Such structures appear, for example, in the Yule-Walker problem [18, section 4.7.2].

But more can be said, especially about $\kappa^{\text{struct}}/\kappa$. Things are particularly elegant for circulant matrices.

8. Circulant matrices. Circulant matrices are of the form

$$C = \begin{pmatrix} c_0 & c_1 & c_2 & c_3 \\ c_3 & c_0 & c_1 & c_2 \\ c_2 & c_3 & c_0 & c_1 \\ c_1 & c_2 & c_3 & c_0 \end{pmatrix}$$

and do have a number of remarkable properties [9]. Denote by P the permutation matrix mapping $(1, \dots, n)^T$ into $(2, \dots, n, 1)^T$. Then a circulant can be written as

$$C = \text{circ}(c_0, \dots, c_{n-1}) = \sum_{\nu=0}^{n-1} c_\nu P^\nu \in M_n^{\text{circ}}.$$

From this polynomial representation it follows that circulants commute. Therefore, for $A \in M_n^{\text{circ}}$, Definition 3.1 implies

$$\varphi^{\text{circ}}(A, x) = \sup\{\|\Delta A \cdot A^{-1}x\| : \Delta A \in M_n^{\text{circ}}, \|\Delta A\| \leq 1\} \leq \|A^{-1}x\|,$$

and observing $\Delta A := I \in M_n^{\text{circ}}$ it follows that

$$(8.1) \quad \varphi^{\text{circ}}(A, x) = \|A^{-1}x\|.$$

THEOREM 8.1. *Let a nonsingular circulant $A \in M_n^{\text{circ}}(\mathbb{R})$ and $0 \neq x \in \mathbb{R}^n$ be given. Then*

$$(8.2) \quad \kappa_{E,f}^{\text{circ}}(A, x) = c \frac{\|A^{-1}x\| \|E\| + \|A^{-1}\| \|f\|}{\|x\|}$$

with $2^{-1/2} \leq c \leq 1$. In particular, for no perturbations in the right-hand side we have

$$(8.3) \quad \kappa_E^{\text{circ}}(A, x) = \frac{\|A^{-1}x\| \|E\|}{\|x\|}$$

and

$$(8.4) \quad \frac{\kappa_E^{\text{circ}}(A, x)}{\kappa_E(A, x)} = \frac{\|A^{-1}x\|}{\|A^{-1}\| \|x\|} \geq \frac{1}{\|A^{-1}\| \|A\|}.$$

The inequality is sharp.

Proof. The assertions follow by Theorem 3.3 and (8.1), where the last inequality stems from $\|x\| \leq \|A\| \|A^{-1}x\|$. Choosing x such that $\|A^{-1}x\| = \sigma_{\min}(A^{-1})\|x\| = \|A\|^{-1}\|x\|$ finishes the proof. \square

So for no perturbations in the right-hand side we have $\kappa_A^{\text{circ}}(A, x) = 1$ for every circulant A and any x chosen such that $\|A^{-1}x\| = \sigma_{\min}(A^{-1})\|x\| = \|A\|^{-1}\|x\|$. Note that $\kappa_A(A, x) = \|A^{-1}\| \|A\|$ for every x . Also note that the ratio in (8.4) applies to general weight matrices E .

These are, however, extreme cases. Formula (8.2) also shows that, in general, $\kappa_{E,f}^{\text{circ}}$ and $\kappa_{E,f}$ are not too far apart because, in general, the same is true for $\|A^{-1}x\|$ and $\|A^{-1}\| \|x\|$.

To analyze the ratio $\kappa^{\text{circ}}/\kappa$ including perturbations in the right-hand side we need again a relation between $\|E\|$ and $\|f\|$. Therefore we switch to the natural choice $E = A$ and $f = b$. Furthermore, we need more details on circulants.

Every circulant is diagonalized by the scaled Fourier matrix $F \in M_n(\mathbb{C})$, $F_{ij} = \omega^{(i-1)(j-1)}/\sqrt{n}$, for ω denoting the n th root of unity [9]. Note that F is unitary and symmetric. So every circulant C is represented by $C = F^H D F$ for some diagonal $D \in M_n(\mathbb{C})$. We need some auxiliary results which will also be useful for Hankel matrices.

LEMMA 8.2. *Let $A \in M_n(\mathbb{C})$, $z \in \mathbb{C}^n$, and a circulant $C \in M_n^{\text{circ}}(\mathbb{C})$ be given. Then*

$$\|AC\| = \|AC^H\| \quad \text{and} \quad \|Cz\| = \|C^H z\|.$$

Proof. Let $C = F^H D F$ for diagonal $D \in M_n(\mathbb{C})$. There is diagonal $S \in M_n(\mathbb{C})$ with $|S_{ii}| = 1$ for all i and $D = S D^H = D^H S$. Since F and S are unitary we obtain

$$\begin{aligned} \|AC^H\| &= \|AF^H D^H F\| = \|AF^H D^H\| = \|AF^H D^H S\| = \|AF^H D\| \\ &= \|AF^H D F\| = \|AC\| \end{aligned}$$

and

$$\|C^H z\| = \|F^H D^H F z\| = \|D^H F z\| = \|S^H D F z\| = \|F^H D F z\| = \|C z\|. \quad \square$$

The next lemma characterizes real circulants. This result is definitely known; however, the only reference we found contains typos and is without proof. So we repeat the short proof.

LEMMA 8.3. *Every circulant C is equal to $F^H D F$ for (complex) diagonal D . Let P denote the permutation matrix mapping $(1, \dots, n)^T$ into $(1, n, \dots, 2)^T$. Then C is real iff $D = P D^H P$.*

Proof. The matrix C is real iff it is equal to its conjugate \overline{C} . Now the definitions of F and $F = F^T$ imply $\overline{F} = F^H$ and $F^H = P F = F P$; we get the latter equality because F, F^H , and P , are symmetric. Hence

$$\overline{C} = F D^H F^H = F^H \cdot P D^H P \cdot F$$

proves the assertion. \square

For $A = F^H D F \in M_n^{\text{circ}}(\mathbb{R})$ being a circulant, $A^{-1} = F^H D^{-1} F$ is a circulant as well, so (8.1) and Lemma 8.2 show

$$\varphi^{\text{circ}}(A, x) = \|A^{-1}x\| = \|A^{-T}x\|.$$

Combining this with Lemma 6.1 yields

$$\kappa_{A,Ax}^{\text{circ}}(A, x) \geq 2^{-1/2} \sqrt{\|A^{-1}\| \|A\|},$$

and (4.1) implies

$$\frac{\kappa_{A,Ax}^{\text{circ}}(A, x)}{\kappa_{A,Ax}(A, x)} \geq \frac{1}{2\sqrt{2} \cdot \sqrt{\|A^{-1}\| \|A\|}}.$$

We give an explicit $n \times n$ example, $n \geq 5$, showing that this inequality is sharp up to a small constant factor. For $m \geq 0$ and $0 < \varepsilon < 1$ define

$$(8.5) \quad A = F^H \text{diag}(1, v, \varepsilon, \varepsilon^{-1}, [1,]\varepsilon^{-1}, \varepsilon, v)F =: F^H D F,$$

where v denotes a row vector of m ones and $[1,]$ indicates that this diagonal element 1 may be left out. Accordingly, A is a circulant of dimension $n = 2m + 5$ or $n = 2m + 6$, depending on whether the diagonal element 1 is left out or not. In either case A is real by Lemma 8.3. The eigenvalues of A are the D_{ii} with corresponding columns of F^H as eigenvectors. Particularly, e is an eigenvector to $D_{11} = 1$, so in our case $Ae = A^{-1}e = e$. Furthermore, $\|A\| = \varepsilon^{-1} = \|A^{-1}\|$. For $x = e/\sqrt{n}$ we have $\|b\| = \|Ax\| = \|A^{-1}x\| = \|x\| = 1$. So (2.2) gives

$$\kappa_{A,Ax}(A, x) = \varepsilon^{-2} + \varepsilon^{-1},$$

and Theorem 8.1 implies

$$\kappa_{A,Ax}^{\text{circ}}(A, x) = c \frac{\varepsilon^{-1} + \varepsilon^{-1}}{1} \leq 2\varepsilon^{-1}$$

for $2^{-1/2} \leq c \leq 1$. Summarizing, we have the following result for circulants.

THEOREM 8.4. *For a nonsingular circulant $A \in M_n^{\text{circ}}(\mathbb{R})$ and $0 \neq x \in \mathbb{R}^n$ we have*

$$1 \geq \frac{\kappa_{A,Ax}^{\text{circ}}(A, x)}{\kappa_{A,Ax}(A, x)} \geq \frac{1}{2\sqrt{2} \sqrt{\|A^{-1}\| \|A\|}}.$$

As by the matrix (8.5) the second estimation is sharp up to a factor $4\sqrt{2}$ for all $n \geq 5$.

Finally, we remark that in case of unstructured perturbations, allowing or not allowing perturbations in the right-hand side may alter the condition number by at most a factor of 2; see (4.1). This changes dramatically for circulant structured perturbations (and also for other structures). Following along the lines of example (8.5), define

$$(8.6) \quad A = F^H \text{diag}(1, v, \varepsilon, [1,]\varepsilon, v)F$$

with v denoting a row vector of $m \geq 0$ ones. Thus, A is of dimension $n = 2m + 3$ or $n = 2m + 4$, depending on whether the diagonal element is left out or not. The same arguments as before apply to $x = e/\sqrt{n}$, and $\|A\| = \|A^{-1}x\| = \|Ax\| = 1$, $\|A^{-1}\| = \varepsilon^{-1}$, (2.2), and Theorem 8.1 yield

$$\kappa_A(A, x) = \varepsilon^{-1}, \quad \kappa_{A,Ax}^{\text{circ}}(A, x) \geq 2^{-1/2}(1 + \varepsilon^{-1}), \quad \text{but} \quad \kappa_A^{\text{circ}}(A, x) = 1.$$

For a discussion of stability of a numerical algorithm for solving a linear system it seems inappropriate to ignore perturbations in the right-hand side. So (8.3) may

be of more theoretical interest. However, Theorem 8.4 shows that a linear system may be beyond the scope of a numerical algorithm which is only stable with respect to general perturbations, whereas it may be solved to some precision by a special circulant solver.

Notice that the ratio $\kappa^{\text{circ}}/\kappa$ may only become small for ill-conditioned matrices. This is also true for componentwise perturbations, as we will see in Part II of this paper (Theorem 7.2). In fact, this is the only structure out of (2.4) for which this statement is true.

9. Symmetric Toeplitz and persymmetric Hankel matrices. With Theorem 6.5 and (6.11) we already have computable bounds for κ^{symToep} and, therefore, in view of (6.3), for $\kappa^{\text{persymHankel}}$. More can be said about κ^{symToep} and also about how small the ratio $\kappa^{\text{symToep}}/\kappa$ can be.

Let $\tilde{J} \in \{+J, -J\}$, $\tilde{J} \in M_n(\mathbb{R})$, and $x \in \mathbb{R}^n$ be given such that $x = \tilde{J}x$. Then $A \in M_n^{\text{symToep}}(\mathbb{R})$ implies $A = \tilde{J}A\tilde{J}$ and $\tilde{J}Ax = \tilde{J}A\tilde{J}x = Ax$. That means every $A \in M_n^{\text{symToep}}(\mathbb{R})$ maps $X := \{x \in \mathbb{R}^n : x = \tilde{J}x\}$ into itself. For nonsingular A , the mapping $A : X \rightarrow X$ is bijective. Assume for the moment that n is even, set $m = n/2$, and split A into

$$(9.1) \quad A = \begin{pmatrix} T & U \\ U^T & T \end{pmatrix} \quad \text{with } T, U \in M_m(\mathbb{R}).$$

Accordingly, split \tilde{J} into $\tilde{J} = \begin{pmatrix} 0 & \bar{J} \\ \bar{J} & 0 \end{pmatrix}$ such that $|\bar{J}| = J_m$ is the ‘‘flip’’-matrix of dimension m . Then $A = \tilde{J}A\tilde{J}$ implies $U^T = \bar{J}U\bar{J}$. For $x \in X$ this means $x = \begin{pmatrix} \bar{x} \\ \bar{J}\bar{x} \end{pmatrix}$ with $\bar{x} \in \mathbb{R}^m$ and therefore

$$Ax = A \begin{pmatrix} \bar{x} \\ \bar{J}\bar{x} \end{pmatrix} = \begin{pmatrix} (T + U\bar{J})\bar{x} \\ \bar{J}(T + U\bar{J})\bar{x} \end{pmatrix}.$$

Thus nonsingularity of A implies nonsingularity of $T + U\bar{J}$.

To estimate φ^{symToep} let nonsingular $A \in M^{\text{symToep}}$, $\Delta A \in M^{\text{symToep}}$, and $x = \tilde{J}x = \begin{pmatrix} \bar{x} \\ \bar{J}\bar{x} \end{pmatrix} \in \mathbb{R}^n$ be given. Then

$$\Delta Ax = \begin{pmatrix} \bar{y} \\ \bar{J}\bar{y} \end{pmatrix} \quad \text{and} \quad A^{-1}\Delta Ax = A^{-1} \begin{pmatrix} \bar{y} \\ \bar{J}\bar{y} \end{pmatrix} = \begin{pmatrix} \bar{z} \\ \bar{J}\bar{z} \end{pmatrix}$$

for some $\bar{y}, \bar{z} \in \mathbb{R}^m$, where $\bar{y} = (T + U\bar{J})\bar{z}$. Therefore

$$\|A^{-1}\Delta Ax\| = \left\| \begin{pmatrix} (T + U\bar{J})^{-1}\bar{y} \\ \bar{J}(T + U\bar{J})^{-1}\bar{y} \end{pmatrix} \right\| \leq \|(T + U\bar{J})^{-1}\| \left\| \begin{pmatrix} \bar{y} \\ \bar{J}\bar{y} \end{pmatrix} \right\|.$$

Moreover,

$$\left\| \begin{pmatrix} \bar{y} \\ \bar{J}\bar{y} \end{pmatrix} \right\| = \left\| \begin{pmatrix} \bar{y} \\ \bar{J}\bar{y} \end{pmatrix} \right\| = \|\Delta Ax\| \leq \|\Delta A\| \cdot \|x\|$$

and therefore

$$(9.2) \quad \varphi^{\text{symToep}}(A, x) \leq \|(T + U\bar{J})^{-1}\| \|x\|.$$

The same analysis, only more technical, is possible for odd n . In this case $m := (n + 1)/2$ and

$$\pm \bar{J} = \begin{pmatrix} & & & 1 & 0 \\ & & & \vdots & \\ & & \cdot & & \\ 1 & & & & 0 \end{pmatrix} \in M_{m-1,m}(\mathbb{R}).$$

Note that $x = \tilde{J}x$ implies $x_m = 0$ in case $\tilde{J} = -J$. For the splitting

$$(9.3) \quad A = \begin{pmatrix} T_1 & U \\ U^T & T_2 \end{pmatrix}, \quad T_1 \in M_m^{\text{symToep}}, \quad T_2 \in M_{m-1}^{\text{symToep}}, \quad U \in M_{m,m-1}(\mathbb{R}),$$

we obtain $\bar{J}T_1\bar{J}^T = T_2$. In a similar way as before one can show

$$(9.4) \quad \varphi^{\text{symToep}}(A, x) \leq \|(T_1 + U\bar{J})^{-1}\| \|x\|.$$

The steps are technical and omitted. Combining (9.2) and (9.4) with Theorem 3.3 we obtain the following result.

THEOREM 9.1. *Let nonsingular $A \in M_n^{\text{symToep}}$, and for $\tilde{J} = sJ, s \in \{-1, 1\}$, let $0 \neq x \in \mathbb{R}^n$ be given with $x = \tilde{J}x$. Set $m := \lceil n/2 \rceil$ and define*

$$\bar{J} = \begin{pmatrix} & & & s \\ & & \cdot & \\ & & & \\ s & & & \end{pmatrix} \in M_m(\mathbb{R}) \quad \text{for } n \text{ even}$$

and

$$\bar{J} = \begin{pmatrix} & & s & 0 \\ & & \vdots & \\ & & & \\ s & & & 0 \end{pmatrix} \in M_{m-1}(\mathbb{R}) \quad \text{for } n \text{ odd}.$$

Then for $T := A[1 : m, 1 : m]$ and $U := A[1 : m, m + 1 : n]$ we have

$$(9.5) \quad \kappa_{E,f}^{\text{symToep}}(A, x) \leq \|(T + U\bar{J})^{-1}\| \|E\| + \|A^{-1}\| \frac{\|f\|}{\|x\|}.$$

Particularly for no perturbations in the right-hand side, we obtain

$$(9.6) \quad \frac{\kappa_E^{\text{symToep}}(A, x)}{\kappa_E(A, x)} \leq \frac{\|(T + U\bar{J})^{-1}\|}{\|A^{-1}\|}.$$

Note that the upper bound for $\kappa_E^{\text{symToep}}(A, x)$ is only true for x with $x = \tilde{J}x$. The ratio in the right-hand side of (9.6) may become arbitrarily small as for

$$A = \text{Toeplitz}(1, 0, \dots, 0, 1 + \varepsilon) \quad \text{and} \quad x = e/\sqrt{n}.$$

Again we use Matlab notation; that is, $\text{Toeplitz}(c)$ denotes the symmetric Toeplitz matrix with first column c . In this case $T + UJ = \text{diag}(2 + \varepsilon, 1, \dots, 1)$ and $\kappa_E^{\text{symToep}}(A, x) \leq \|E\|$. On the other hand, $y = (-1, 0, \dots, 0, 1)^T$ is an eigenvector of A to the eigenvalue ε , so $\kappa_E(A, x) = \|A^{-1}\| \|E\| \geq \varepsilon^{-1} \|E\|$. However, allowing perturbations in the right-hand side, we obtain for the natural choice $E = A, f = b$

$$\kappa_{A, Ax}^{\text{symToep}}(A, x) \geq (2\sqrt{2\varepsilon})^{-1} \|A\|,$$

which is almost the same as $\kappa_A(A, x)$. Indeed, allowing perturbations in the right-hand side, the ratio $\kappa_{A, Ax}^{\text{symToep}} / \kappa_{A, Ax}$ depends on the condition number $\kappa_{A, Ax}$. It can only become small for ill-conditioned matrices. The ratio can be estimated as before using Lemma 6.1.

THEOREM 9.2. *Let nonsingular $A \in M_n^{\text{symToep}}(\mathbb{R})$ and $0 \neq x \in \mathbb{R}^n$ be given. Then*

$$1 \geq \frac{\kappa_{A, Ax}^{\text{symToep}}(A, x)}{\kappa_{A, Ax}(A, x)} \geq \frac{1}{2\sqrt{2} \cdot \sqrt{\|A^{-1}\| \|A\|}}.$$

Proof. We have $I \in M_n^{\text{symToep}}$, so $\varphi^{\text{symToep}}(A, x) \geq \|A^{-1}x\| = \|A^{-T}x\|$, and Lemma 6.1 and (4.1) finish the proof. \square

The lower bound in Theorem 9.2 seems not far from being sharp. Consider

$$A = \text{Toeplitz}(1, -1 - \varepsilon, 1 - \varepsilon, -1 + \varepsilon) \quad \text{and} \quad x = e,$$

the symmetric Toeplitz matrix with first row $[1, -1 - \varepsilon, 1 - \varepsilon, -1 + \varepsilon]$. Then

$$\kappa_{A, Ax}(A, x) \sim 16\varepsilon^{-2} \quad \text{and} \quad \kappa_{A, Ax}^{\text{symToep}}(A, x) < 11\varepsilon^{-1}.$$

Unfortunately, we do not have a generic $n \times n$ example. However, it is numerically easy to find examples of larger dimension. Therefore, we expect the second inequality in Theorem 9.2 to be sharp up to a small constant for all n .

Additional algebraic properties such as positive definiteness of the matrix do not improve the situation. An example is the symmetric positive definite Toeplitz matrix A with first row $(1 + \varepsilon^2, -1 + \varepsilon, 1 - \varepsilon, -1 + 3\varepsilon)$ and $x := e$. One computes $\lambda_{\min}(A) = 0.75\varepsilon^2 + \mathcal{O}(\varepsilon^3)$, and (2.2) and Theorem 9.1 yield

$$\kappa_{A, Ax}(A, x) > 5.33\varepsilon^{-2} + \mathcal{O}(\varepsilon^{-1}) \quad \text{and} \quad \kappa_{A, Ax}^{\text{symToep}}(A, x) < 7\varepsilon^{-1} + \mathcal{O}(1).$$

Note that the estimation in Theorem 9.1 is only valid for $x = \tilde{J}x$, $\tilde{J} = sJ$, $s \in \{+1, -1\}$. Let general $x \in \mathbb{R}^n$ be given and split $x = \begin{pmatrix} u \\ v \end{pmatrix}$ into $u \in \mathbb{R}^m, v \in \mathbb{R}^{n-m}$. Define \bar{J} as in Theorem 9.1 with $s = 1$, and set $\bar{y} := \frac{1}{2}(u + \bar{J}v)$ and $\bar{z} := \frac{1}{2}(u - \bar{J}v)$. Then for $y := \begin{pmatrix} \bar{y} \\ \bar{J}\bar{y} \end{pmatrix} \in \mathbb{R}^n$ and $z := \begin{pmatrix} \bar{z} \\ -\bar{J}\bar{z} \end{pmatrix} \in \mathbb{R}^n$ we have

$$Jy = y, \quad -Jz = z, \quad \text{and} \quad x = y + z.$$

For $\Delta A \in M_n^{\text{symToep}}$ and $\|\Delta A\| \leq 1$ we can apply (9.2) and (9.4) to conclude that

$$\|A^{-1}\Delta Ax\| = \|A^{-1}\Delta A(y + z)\| \leq \|(T + U\bar{J})^{-1}\| \|y\| + \|(T - U\bar{J})^{-1}\| \|z\|.$$

COROLLARY 9.3. *For nonsingular $A \in M_n^{\text{symToep}}$, $0 \neq x \in \mathbb{R}^n$, and T, U, \bar{J} , y and z as defined above we have*

$$\kappa_{E, f}^{\text{symToep}}(A, x) \leq (\|(T + U\bar{J})^{-1}\| \|y\| + \|(T - U\bar{J})^{-1}\| \|z\|) \frac{\|E\|}{\|x\|} + \|A^{-1}\| \frac{\|f\|}{\|x\|}.$$

Obviously $\|y\| \leq \|x\|$ and $\|z\| \leq \|x\|$, so one may replace the expression in the parentheses by $\mu\|x\|$ with $\mu := \max(\|(T + U\bar{J})^{-1}\|, \|(T - U\bar{J})^{-1}\|)$. However, such an approach does not give additional information. Let $A^{-1}w = \lambda w$ for $0 \neq w \in \mathbb{R}^n$. Then $A^{-1} \cdot Jw = JA^{-1}J \cdot Jw = \lambda Jw$, such that $A^{-1}(w \pm Jw) = \lambda(w \pm Jw)$. At

least one of $w \pm Jw$ is nonzero, so we conclude that to every eigenvalue of A^{-1} there is an eigenvector w such that $w = sJw$ for $s \in \{-1, +1\}$. For $w, \|w\| = 1$ being an eigenvector to $|\lambda| = \|A^{-1}\|$ and for the splitting $w = \begin{pmatrix} \bar{w} \\ \bar{w} \end{pmatrix}$ it follows that

$$\|A^{-1}\| = \|A^{-1}w\| = \left\| \begin{pmatrix} (T + sU\bar{J})^{-1}\bar{w} \\ \bar{J}(T + sU\bar{J})^{-1}\bar{w} \end{pmatrix} \right\| \leq \mu \cdot \left\| \begin{pmatrix} \bar{w} \\ \bar{w} \end{pmatrix} \right\| = \mu,$$

so that the above approach only verifies $\kappa_{E,f}^{\text{symToep}} \leq \kappa_{E,f}$.

We also see from this how to construct examples with small ratio $\kappa^{\text{symToep}}/\kappa$. If A is ill conditioned, at least one of the matrices $T + sU\bar{J}$ must be equally ill conditioned. Small ratios may occur if one of them, say for $s = 1$, is well conditioned and x is chosen with big part $y = Jy$ but small $z = -Jz$ in the splitting $x = y + z$.

Finally, note that Theorem 9.1 and Corollary 9.3 give upper bounds for κ^{symToep} . We do not know how sharp estimation (9.5) is. Numerical experience suggests that the overestimation is small. Can that be proved? Again, all statements in this section are valid mutatis mutandis for $A \in M_n^{\text{persymHankel}}$.

10. Hankel and general Toeplitz matrices. With Theorem 6.5 and (6.11) we already have computable bounds for the (normwise) Hankel condition number and therefore, in view of (6.4), for κ^{Toep} . In the following we investigate how small the ratio $\kappa^{\text{Hankel}}/\kappa$ can be. We first show a lower bound in the spirit of Theorems 8.4 and 9.2.

Suppose $A^T = A \in M_n(\mathbb{R})$, not necessarily $A \in M_n^{\text{Hankel}}(\mathbb{R})$. By definition,

$$\varphi^{\text{Hankel}}(A, x) = \sup\{\|A^{-1}\Delta Ax\| : \Delta A \in M_n^{\text{Hankel}}(\mathbb{R}), \|\Delta A\| \leq 1\}.$$

Hankel matrices are symmetric. So if we can show that for every $0 \neq x \in \mathbb{R}^n$ there is a Hankel matrix ΔA with $\|\Delta A\| \leq 1$ and $\Delta Ax = x$, then

$$(10.1) \quad \varphi^{\text{Hankel}}(A, x) \geq \|A^{-1}x\| = \|A^{-T}x\|,$$

and Lemma 6.1 delivers the desired bound. This is indeed true, as shown by the following lemma. We will prove it for the real and complex cases, the latter being needed in sections 11 and 12.

LEMMA 10.1. *Let $x \in \mathbb{C}^n$ be given. Then there exists $H \in M_n^{\text{Hankel}}(\mathbb{C})$ with $Hx = \bar{x}$ and $\|H\| \leq 1$, where \bar{x} denotes the complex conjugate of x . In case $x \in \mathbb{R}^n$, H can be chosen real so that $Hx = x$.*

Proof. The expression (6.5) is of course also true for complex Hankel matrices because Φ^{Hankel} is a 0/1-matrix. So we are looking for a parameter vector $p \in \mathbb{C}^{2n-1}$ such that the Hankel matrix H with $\text{vec}(H) = \Phi^{\text{Hankel}}p$ satisfies the assertions of the lemma. Then

$$(10.2) \quad Hx = \Psi_x^{\text{Hankel}}p$$

for Ψ_x^{Hankel} as in (6.9), (6.10), and (6.11). We discuss the following for $n = 3$, which will give enough information for the general case. We first embed $\Psi_x := \Psi_x^{\text{Hankel}}$ into the circulant C_x with the first row identical to that of Ψ_x , i.e.,

$$C_x := \begin{pmatrix} x_1 & x_2 & x_3 & 0 & 0 \\ 0 & x_1 & x_2 & x_3 & 0 \\ 0 & 0 & x_1 & x_2 & x_3 \\ x_3 & 0 & 0 & x_1 & x_2 \\ x_2 & x_3 & 0 & 0 & x_1 \end{pmatrix}.$$

Then the matrix of the first n rows of C_x is equal to Ψ_x . Define

$$(10.3) \quad C := C_x^+ C_x^H,$$

with C_x^+ denoting the pseudoinverse of C_x . For $C_x = F^H D F$, the pseudoinverse $C_x^+ = F^H D^+ F$ is also a circulant, and we have

$$C_x C = F^H D F \cdot F^H D^+ F \cdot F^H D^H F = F^H D D^+ D^H F = F^H D^H F = C_x^H.$$

But Ψ_x comprises the first n rows of C_x , so $\Psi_x C$ is equal to the matrix of the first n rows of C_x^H . Define

$$(10.4) \quad p := C e_1,$$

with e_1 denoting the first column of I_{2n-1} . The first n rows of $C_x^H e_1$ form the vector \bar{x} , so by (10.2),

$$Hx = \Psi_x p = \Psi_x C e_1 = \bar{x}$$

for the Hankel matrix H defined by the parameter vector $p = C e_1$. Note that by (10.3) and (10.4) C , and therefore H , is real for real x so that $Hx = x$ in that case.

It remains to estimate the matrix norm of H . Denote the first column of the circulant C by $(c_1, \dots, c_{2n-1})^T$. For $n = 3$, the definitions (10.4) and (10.2) imply

$$H = \begin{pmatrix} c_1 & c_2 & c_3 \\ c_2 & c_3 & c_4 \\ c_3 & c_4 & c_5 \end{pmatrix} \quad \text{and} \quad HJ = \begin{pmatrix} c_3 & c_2 & c_1 \\ c_4 & c_3 & c_2 \\ c_5 & c_4 & c_3 \end{pmatrix}.$$

The matrix HJ is the lower left $n \times n$ submatrix of C . So by Lemma 8.2 it follows that

$$\|H\| = \|HJ\| \leq \|C\| = \|C_x^+ C_x^H\| = \|C_x^+ C_x\| = \|D^+ D\| = 1. \quad \square$$

Combining Lemma 10.1 with (10.1), Lemma 6.1, and (4.1) proves the following lower bounds. Note that only symmetry of A was used in (10.1).

THEOREM 10.2. *Let nonsingular symmetric $A \in M_n(\mathbb{R})$, and let $0 \neq x \in \mathbb{R}^n$ be given. Then*

$$(10.5) \quad \kappa_{A, Ax}^{\text{Hankel}}(A, x) \geq 2^{-1/2} \sqrt{\|A^{-1}\| \|A\|}$$

and therefore

$$(10.6) \quad 1 \geq \frac{\kappa_{A, Ax}^{\text{Hankel}}(A, x)}{\kappa_{A, Ax}(A, x)} \geq \frac{1}{2\sqrt{2} \cdot \sqrt{\|A^{-1}\| \|A\|}}.$$

The lower bound (10.6) is a severe underestimation, in fact, it is independent of A . By Corollary 6.6 we know that

$$2^{1/2} \frac{\kappa_{E, f}^{\text{Hankel}}(A, x)}{\kappa_{E, f}(A, x)} \geq \frac{\kappa_E^{\text{Hankel}}(A, x)}{\kappa_E(A, x)} \geq n^{-1/2} \frac{\sigma_{\min}(\Psi_x^{\text{struct}})}{\|x\|}.$$

If Ψ_x^{struct} were rank-deficient, this would imply that every minor of size n is zero. Then the minor of first n columns of Ψ_x^{Hankel} as in (6.10) implies $x_1 = 0$, and continuing

TABLE 10.1
Mean value and standard deviation of $\tau(x) = \sigma_{\min}(\Psi_x^{\text{Hankel}})/\|x\|$.

n	Uniform x_i		Normal x_i	
	mean($\tau(x)$)	std($\tau(x)$)	mean($\tau(x)$)	std($\tau(x)$)
10	0.49	0.135	0.50	0.136
20	0.42	0.110	0.42	0.111
50	0.35	0.084	0.35	0.084
100	0.31	0.069	0.31	0.069

with the minors of columns i to $i + n - 1$ we conclude that $x = 0$. This implies $\sigma_{\min}(\Psi_x^{\text{struct}}) > 0$ for all $x \neq 0$ such that for fixed x there is a minimum ratio of the structured Hankel and the unstructured condition number.

Extensive numerical statistics on $\tau(x) := \sigma_{\min}(\Psi_x^{\text{Hankel}})/\|x\|$ suggest that this minimum is in general not too far from 1. In Table 10.1 we list the mean value and standard deviation of $\tau(x)$ for some 10^6 samples of x with entries uniformly distributed in $[-1, 1]$ and for entries of x with normal distribution with mean 0 and standard deviation 1.

We mention that the numbers in the two rightmost columns in Table 10.1 are almost the same for solution vectors x such that $x_i = s \cdot y_i$ with random sign $s \in \{-1, 1\}$ and uniform y_i with mean 1 and standard deviation 1.

Note again that this is a statistic on *solution vectors* x showing a lower bound for the ratio in Corollary 6.6 between the Hankel and the traditional (unstructured) condition numbers. This ratio applies to *every* matrix A regardless of its condition number.

Small values of $\tau(x) = \sigma_{\min}(\Psi_x^{\text{Hankel}})/\|x\|$ seem rare, but they are possible. Particularly, small values seem to occur for positive x and $x = Jx$. Statistically the means in Table 10.1 drop by about a factor of 2 to 3 for such randomly chosen x . A specific choice of x proposed by Heinig [21] is comprised of the coefficients of $(t + 1)^{n-1}$. For this x we obtain

$$\tau(x) \sim 2.5^{-n}.$$

This generates a lower bound for $\kappa_{A, Ax}^{\text{Hankel}}(A, x)$. We indeed managed to find Hankel matrices with $\|A^{-1}\Psi_x^{\text{Hankel}}\|/\|x\| < 2^{-n}\|A^{-1}\|$ for that x and dimensions up to 15. That means for the unperturbed right-hand side it is $\kappa_A^{\text{Hankel}}(A, x)/\kappa_A(A, x) < 2^{-n}$. We could neither construct generic $n \times n$ matrices A with this property nor find examples with the ratio of condition numbers $\kappa_{A, Ax}^{\text{Hankel}}(A, x)/\kappa_{A, Ax}(A, x)$ (allowing perturbations in the right-hand side) getting significantly less than one. This includes in particular positive definite Hankel matrices which are known to be generally ill-conditioned [3].

An open problem is how small $\tau(x)$ can be; that is, what is the smallest possible value of $\sigma_{\min}(\Psi_x^{\text{Hankel}})$ for Ψ_x^{Hankel} as in (6.10) and $\|x\| = 1$? Based on that, how small may $\kappa_A^{\text{Hankel}}(A, x)/\kappa_A(A, x)$ become?

For general (normwise) perturbations in the matrix *and* the right-hand side we conjecture that Hankel structured and unstructured stabilities differ only by a small factor, supposedly only mildly or not at all, depending on n . In other words, $\kappa_{A, Ax}^{\text{Hankel}}(A, x)/\kappa_{A, Ax}(A, x) \geq \gamma$ for γ not much less than one.

Meanwhile Böttcher and Grudsky give a partial answer to that [6]. They show, based on a deep result by Konyagin and Schlag [31], that there exist universal constants $n_0 \in \mathbb{N}$ and $\varepsilon > 0$ such that the following is true. Let $x = (x_1, \dots, x_n) \in \mathbb{R}^n$,

$n \geq n_0$, comprise independent standard normal or independent Rademacher variables (recall that Rademacher variables are random with value 1 or -1 each with probability $1/2$). Then, for all $A \in M_n^{\text{Hankel}}(\mathbb{R})$,

$$\text{probability} \left(\frac{\kappa_A^{\text{Hankel}}(A, x)}{\kappa_A(A, x)} \geq \frac{\varepsilon}{n^{3/2}} \right) > \frac{99}{100}.$$

11. Inversion of structured matrices. Similarly to the structured condition number for linear systems, the structured condition number for matrix inversion is defined by

(11.1)

$$\kappa_E^{\text{struct}}(A) := \limsup_{\varepsilon \rightarrow 0} \left\{ \frac{\|(A + \Delta A)^{-1} - A^{-1}\|}{\varepsilon \|A^{-1}\|} : \Delta A \in M_n^{\text{struct}}(\mathbb{R}), \|\Delta A\| \leq \varepsilon \|E\| \right\}.$$

For $M_n^{\text{struct}}(\mathbb{R}) = M_n(\mathbb{R})$ this is the usual (unstructured) condition number which is well known [27, Theorem 6.4] to be

$$\kappa_E(A) = \|A^{-1}\| \|E\|.$$

Surprisingly, the same is true for all of the linear structures in (2.4). A reasoning is that by Theorem 4.1 the worst case condition number of a linear system maximized over all right-hand sides is equal to the unstructured condition number. So in some way the set of columns of the identity matrix is general enough to achieve the worst case.

THEOREM 11.1. *Let nonsingular $A \in M_n^{\text{struct}}(\mathbb{R})$ be given for $\text{struct} \in \{\text{sym}, \text{persym}, \text{skewsym}, \text{symToep}, \text{Toep}, \text{circ}, \text{Hankel}, \text{persymHankel}\}$. Then*

$$\kappa_E^{\text{struct}}(A) = \|A^{-1}\| \|E\|.$$

Proof. As in the unstructured case we use the expansion

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

Therefore, the result is proved if we can show that

$$(11.2) \quad \omega^{\text{struct}}(A) := \sup\{\|A^{-1} \Delta A A^{-1}\| : \Delta A \in M^{\text{struct}}, \|\Delta A\| \leq 1\} \geq \|A^{-1}\|^2$$

because this obviously implies $\omega^{\text{struct}}(A) = \|A^{-1}\|^2$. Let $x, y \in \mathbb{R}^n$, $\|x\| = \|y\| = 1$ be given with $A^{-1}x = \|A^{-1}\|y$. Then Definition 3.1 implies

$$\omega^{\text{struct}}(A) \geq \sup\{\|A^{-1} \Delta A A^{-1}x\| : \Delta A \in M^{\text{struct}}, \|\Delta A\| \leq 1\} = \|A^{-1}\| \varphi^{\text{struct}}(A, y).$$

Therefore Lemma 5.2 proves (11.2) for $\text{struct} \in \{\text{sym}, \text{persym}, \text{skewsym}\}$. For normal $A \in M_n^{\text{struct}}(\mathbb{R})$, it is $A^{-1}x = \lambda x$ with $\|x\| = 1$ and $|\lambda| = \|A^{-1}\|$. Hence (11.2) is also proved for symmetric Toeplitz and circulant structures by using $\Delta A := I$. For $A \in M_n^{\text{persymHankel}}(\mathbb{R})$, $AJ \in M_n^{\text{symToep}}(\mathbb{R})$ and $JA^{-1}x = \lambda x$ with $\|x\| = 1$, and $|\lambda| = \|JA^{-1}\| = \|A^{-1}\|$ proves (11.2) by using $\Delta A := J \in M_n^{\text{persymHankel}}(\mathbb{R})$. For Hankel matrices again $A^{-1}x = \lambda x$ for $\|x\| = 1$ and $|\lambda| = \|A^{-1}\|$, and Lemma 10.1 yields existence of $\Delta A \in M_n^{\text{Hankel}}(\mathbb{R})$ with $\|\Delta A\| \leq 1$ and $\Delta Ax = x$, and for $A \in M_n^{\text{Toep}}(\mathbb{R})$ we have $AJ \in M_n^{\text{Hankel}}(\mathbb{R})$. \square

The theorem shows that among the worst case perturbations for the inverse of a structured matrix there are always perturbations of the same structure, the same result (cf. Theorem 5.3) as for linear systems with fixed right-hand side and $\text{struct} \in \{\text{sym}, \text{persym}, \text{skewsym}\}$.

The proof basically uses the fact that A or JA is normal. It also can be extended to the complex case. Here the structure is still strong enough, although the singular values need not coincide with the absolute values of the eigenvalues. We have the following result.

THEOREM 11.2. *Let nonsingular $A \in M_n^{\text{struct}}(\mathbb{C})$ be given for struct being Hermitian, skew-Hermitian, Toeplitz, circulant, or Hankel. Then*

$$\kappa_E^{\text{struct}}(A) = \|A^{-1}\| \|E\|.$$

Proof. We proceed as in the proof of Theorem 11.1 and have to show $\omega^{\text{struct}}(A) \geq \|A^{-1}\|^2$ for the $\omega^{\text{struct}}(A)$ defined in (11.2). For normal A , there is $A^{-1}x = \lambda x$ with $\|x\| = 1$ and $|\lambda| = \|A^{-1}\|$. So the theorem is proved for the Hermitian and circulant cases by using $\Delta A = I$, and for the skew-Hermitian case by using $\Delta A = \sqrt{-1}I$.

For A being Hankel, A is especially (complex) symmetric. So a result by Takagi [28, Corollary 4.4.4] implies $A = U\Sigma U^T$ for nonnegative diagonal Σ and unitary U . For x denoting the n th column of U we have $A\bar{x} = \sigma_{\min}(A)x$, and therefore $A^{-1}x = \|A^{-1}\|\bar{x}$. By Lemma 10.1 there exists $\Delta A \in M_n^{\text{Hankel}}(\mathbb{C})$ with $\|\Delta A\| \leq 1$ and $\Delta A\bar{x} = x$, so that $A^{-1}\Delta AA^{-1}x = \|A^{-1}\|^2x$ and

$$\omega^{\text{struct}}(A) \geq \|A^{-1}\Delta AA^{-1}x\| = \|A^{-1}\|^2.$$

Finally, for complex Toeplitz A , $H := JA$ is Hankel and, as above, we conclude that there is x and ΔH with $H^{-1}\Delta HH^{-1}x = \|H^{-1}\|^2x$. Then $\Delta A := J\Delta H$ is Toeplitz with $\|\Delta A\| \leq 1$, and $y := Jx$ with $\|y\| = 1$ yields

$$\omega^{\text{struct}}(A) \geq \|A^{-1}\Delta AA^{-1}y\| = \|H^{-1}\Delta HH^{-1}x\| = \|H^{-1}\|^2 = \|A^{-1}\|^2. \quad \square$$

One might conjecture that the result in Theorems 11.1 and 11.2 is true for all linear structures. This is, however, not the case, for example, for (general) tridiagonal Toeplitz matrices or, more generally, for (general) tridiagonal matrices. Consider

$$(11.3) \quad A = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}$$

for small $\alpha > 0$. Then $\|A\| \sim 1$ and $\|A^{-1}\| \sim \alpha^{-3}$. For general $\Delta A \in M_3^{\text{tridiag}}(\mathbb{R})$ with $\|\Delta A\| \leq 1$ one computes $\|A^{-1}\Delta AA^{-1}\| = \mathcal{O}(\alpha^{-5})$, so that $\omega^{\text{struct}}(A)$ defined in (11.2) is of the order $\alpha\|A^{-1}\|^2$. This implies that $\kappa_E^{\text{tridiag}}(A)$ is of the order $\alpha\|A^{-1}\|\|E\|$ instead of $\|A^{-1}\|\|E\|$. The same applies for general tridiagonal Toeplitz perturbations. Nevertheless one may ask: Is Theorem 11.1 true for other structures?

Usually linear systems are not solved by multiplying the right-hand side by a computed inverse. For structured matrices with small ratio $\kappa_{A, Ax}^{\text{struct}}/\kappa_{A, Ax}$, lack of stability is yet another reason for that.

12. Distance to singularity. The condition number $\kappa(A) = \|A^{-1}\|\|A\|$ of a matrix is infinite iff the matrix is singular. Therefore it seems plausible that the distance to singularity of a matrix is inversely proportional to its condition number. Define

$$\delta_E^{\text{struct}}(A) := \min\{\alpha : \Delta A \in M_n^{\text{struct}}(\mathbb{R}), \|\Delta A\| \leq \alpha\|E\|, A + \Delta A \text{ singular}\},$$

where $M_n^{\text{struct}}(\mathbb{R}) \subseteq M_n(\mathbb{R})$. For $M_n^{\text{struct}}(\mathbb{R}) = M_n(\mathbb{R})$ this number $\delta_E(A)$ is the traditional (normwise) distance to the nearest singular matrix with respect to unstructured perturbations. A classical result [27, Theorem 6.5] by Eckart and Young [12] for the 2-norm with generalizations by Gastinel and Kahan to other norms is

$$(12.1) \quad \delta_E(A) = \{\|A^{-1}\| \|E\|\}^{-1} = \kappa_E(A)^{-1}$$

for general perturbations $M_n^{\text{struct}}(\mathbb{R}) = M_n(\mathbb{R})$. Thus the distance to singularity for general perturbations is not only inversely proportional to but *equal* to the reciprocal of the condition number. Note that the distance to singularity as well as the condition number may change with diagonal scaling, the former being contrary to componentwise perturbations (cf. Part II, section 9).

There are a number of results on some blockwise structured distance to singularity and on the so-called μ -number (cf. [11, 13, 34, 41, 35]). There also are results on distance to singularity with respect to certain symmetric structures [29]. The question remains of whether a result similar to (12.1) can be obtained for the structured condition number and distance to singularity. It was indeed shown by D. Higham [23] that (12.1) is also true for symmetric perturbations.

In the previous section we have seen that the structured condition number $\kappa_E^{\text{struct}}(A)$ is equal to the unstructured condition number $\|A^{-1}\| \|E\|$ for any E and for *all* structures in (2.4).

We conclude with the remarkable fact that the reciprocal of the condition number is equal to the structured distance to the nearest singular matrix for *all* structures in (2.4).

THEOREM 12.1. *Let nonsingular $A \in M_n^{\text{struct}}(\mathbb{R})$ for $\text{struct} \in \{\text{sym}, \text{persym}, \text{skewsym}, \text{symToep}, \text{Toep}, \text{circ}, \text{Hankel}, \text{persymHankel}\}$ be given. Then*

$$(12.2) \quad \delta_E(A) = \delta_E^{\text{struct}}(A) = \kappa_E^{\text{struct}}(A)^{-1} = \kappa_E(A)^{-1} = \{\|A^{-1}\| \|E\|\}^{-1}.$$

Proof. Without loss of generality we may assume $\|E\| = 1$. Then obviously $\delta_E^{\text{struct}}(A) \geq \delta_E(A) = \sigma_{\min}(A)$, and it remains to show $(A + \Delta A)x = 0$ for some $0 \neq x \in \mathbb{R}^n$ and $\Delta A \in M_n^{\text{struct}}(\mathbb{R})$ with $\|\Delta A\| = \sigma_{\min}(A)$.

For symmetric matrices there is real λ and $0 \neq x \in \mathbb{R}^n$ with $Ax = \lambda x$ and $|\lambda| = \sigma_{\min}(A)$. If $I \in M_n^{\text{struct}}(\mathbb{R})$, then $\Delta A = -\lambda I$ does the job. This proves (12.2) for $\text{struct} \in \{\text{sym}, \text{symToep}\}$. For $\text{struct} \in \{\text{persym}, \text{persymHankel}\}$ and $A \in M_n^{\text{struct}}(\mathbb{R})$, JA is symmetric and $JAx = \lambda x$ for $0 \neq x \in \mathbb{R}^n$ and $|\lambda| = \sigma_{\min}(JA) = \sigma_{\min}(A)$. Therefore $\det(J(A + \Delta A)) = 0 = \det(A + \Delta A)$ for $\Delta A := -\lambda J \in M_n^{\text{struct}}(\mathbb{R})$.

For nonsingular skewsymmetric A we conclude as in the proof of Lemma 5.2 that all singular values have even multiplicity and that there are $u, v \in \mathbb{R}^n$ with $\|u\| = \|v\| = 1$, $u^T v = 0$, and $Av = \sigma_{\min}(A)u$. By Lemma 5.1 we find $\Delta A \in M_n^{\text{skewsym}}(\mathbb{R})$ with $\Delta Av = u$ and $\|\Delta A\| = 1$, so that $A - \sigma_{\min}(A)\Delta A$ is singular.

For a given real circulant $A = F^H D F$ there is $Ax = \lambda x$ with $0 \neq x \in \mathbb{C}^n$ and $|\lambda| = \sigma_{\min}(A)$. If λ is real, $\Delta A = -\lambda I \in M_n^{\text{circ}}(\mathbb{R})$ yields $\det(A + \Delta A) = 0$. For complex λ , define diagonal $\tilde{D} \in M_n(\mathbb{C})$ with all entries zero except the two entries λ and $\bar{\lambda}$ in the same position as in D . Define $\Delta A := F^H \tilde{D} F$. Then Lemma 8.3 implies that ΔA is a real circulant. Moreover, we have $\|\Delta A\| = \max |\tilde{D}_{\nu\nu}| = |\lambda| = \sigma_{\min}(A)$, and $A - \Delta A = F^H (D - \tilde{D}) F$ is singular.

For Hankel matrices there is $Ax = \lambda x$, $\|x\| = 1$, and $|\lambda| = \sigma_{\min}(A)$, and Lemma 10.1 proves this part. Finally, $A \in M_n^{\text{Toep}}(\mathbb{R})$ implies $AJ \in M_n^{\text{Hankel}}$ and we proceed as before. \square

As in the previous section we can formulate this theorem also for complex structures. For nonnormal matrices such as complex Hankel and Toeplitz matrices the key is again the complex part of Lemma 10.1.

THEOREM 12.2. *Let nonsingular $A \in M_n^{\text{struct}}(\mathbb{C})$ be given for struct being Hermitian, skew-Hermitian, Toeplitz, circulant, or Hankel. Then*

$$\delta_E(A) = \delta_E^{\text{struct}}(A) = \kappa_E^{\text{struct}}(A)^{-1} = \kappa_E(A)^{-1} = \{\|A^{-1}\| \|E\|\}^{-1}.$$

Proof. The proof of Theorem 12.1 obviously carries over to the normal case, that is, to complex Hermitian, skew-Hermitian, and circulant matrices. For a Hankel matrix A we use [28, Corollary 4.4.4] the factorization $A = U\Sigma U^T$ with nonnegative diagonal Σ and unitary U as in the previous section. For x denoting the n th column of U we have $A\bar{x} = \sigma_{\min}(A)x$. By Lemma 10.1 there exists $\Delta H \in M_n^{\text{Hankel}}(\mathbb{C})$ with $\|\Delta H\| \leq 1$ and $\Delta H\bar{x} = x$. Obviously, $-\Delta H \in M_n^{\text{Hankel}}(\mathbb{C})$ as well, so that $\Delta A := \sigma_{\min}(A)\Delta H$, $\|\Delta A\| = \sigma_{\min}(A)$, and $(A + \Delta A)x = 0$ finish this part of the proof. For A being Toeplitz, JA is Hankel and we proceed as in the proof of Theorem 12.1. \square

So our results are a structured version of the Eckart–Young theorem, valid for all of our structures in (2.4) including the complex case. Does the result extend to other structures?

13. Conclusion. We proved that for some problems and structures it makes no, or not much, difference whether perturbations are structured or not; for other problems and structures we showed that the sensitivity with respect to structured (normwise) perturbations may be much less than with respect to unstructured perturbations. This was especially true for the important cases of linear systems with a symmetric Toeplitz or circulant matrix. Surprisingly, it turned out that the ratio $\kappa^{\text{struct}}/\kappa$ can only become small for certain *solutions*, independent of the matrix.

The results show that a small ratio $\kappa^{\text{struct}}/\kappa$ seems not typical. So our results may be used to rely on the fact that unstructured and structured sensitivities are, in general, not too far apart. However, it may also define the challenge to design numerical algorithms to solve problems with structured data being stable not only with respect to unstructured perturbations but *being stable with respect to the corresponding structured perturbations*. There exists a result in that direction for normwise perturbations and circulant matrices [40], [27, Theorem 24.3]. However, structured analysis for circulants is assisted by the fact that circulants commute. Beyond that, there are similar results for nonlinear structures such as Cauchy or Vandermonde-like matrices (see the last section in Part II of this paper). We hope our results stimulate further research in that direction for other structures.

Acknowledgments. I wish to thank many colleagues for inspiring discussions, especially Chris Beattie, Gene Golub, Ilse Ipsen, and Nick Trefethen. Also my thanks to the students of a winter course 2001/02 on this subject, especially to Frank Blömeling and Christian Keil, for many helpful discussions, remarks, and for their indulgence during the development of the results. My special thanks go to Albrecht Böttcher from TU Chemnitz for his thorough reading and most constructive and valuable comments. He also granted me the privilege of having a part of this paper appear in his forthcoming book [6]. Last but not least, many thanks to the two anonymous referees for their thorough reading and most constructive and useful comments.

REFERENCES

- [1] G. S. AMMAR AND W. B. GRAGG, *Superfast solution of real positive definite Toeplitz systems*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 61–76.
- [2] S. G. BARTELS AND D. J. HIGHAM, *The structured sensitivity of Vandermonde-like systems*, Numer. Math., 62 (1992), pp. 17–33.
- [3] B. BECKERMANN, *The condition number of real Vandermonde, Krylov and positive definite Hankel matrices*, Numer. Math., 85 (2000), pp. 553–577.
- [4] A. W. BOJANCZYK, R. P. BRENT, AND F. R. DE HOOG, *Stability analysis of a general Toeplitz systems solver*, Numer. Algorithms, 10 (1995), pp. 225–244.
- [5] T. BOROS, T. KAILATH, AND V. OLSHEVSKY, *Pivoting and backward stability of fast algorithms for solving Cauchy linear equations*, Linear Algebra Appl., 343/344 (2001), pp. 63–99.
- [6] A. BÖTTCHER AND S. GRUDSKY, *Special Characteristics of Toeplitz Band Matrices*, to appear.
- [7] R. P. BRENT, *Stability of Fast Algorithms for Structured Linear Systems*, Technical Report TR-CS-97-18, Department of Computer Science, Australian National University, Canberra, 1997.
- [8] 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.
- [9] P. J. DAVIS, *Circulant Matrices*, John Wiley, New York, 1979.
- [10] J. DEMMEL, *Accurate singular value decompositions of structured matrices*, SIAM J. Matrix Anal. Appl., 21 (1999), pp. 562–580.
- [11] J. C. DOYLE, *Analysis of feedback systems with structured uncertainties*, Proc. IEEE-D, 129 (1982), pp. 242–250.
- [12] C. ECKART AND G. YOUNG, *The approximation of one matrix by another of lower rank*, Psychometrika, 1 (1936), pp. 211–218.
- [13] M. K. H. FAN, A. L. TITS, AND J. C. DOYLE, *Robustness in presence of mixed parametric uncertainty and unmodeled dynamics*, IEEE Trans. Automat. Control, 36 (1991), pp. 25–38.
- [14] V. FRAYSSÉ, S. GRATTON, AND V. TOUMAZOU, *Structured backward error and condition number for linear systems of the type $A * Ax = b$* , BIT, 40 (2000), pp. 74–83.
- [15] I. GOHBERG, T. KAILATH, AND V. OLSHEVSKY, *Fast Gaussian elimination with partial pivoting for matrices with displacement structure*, Math. Comput., 64 (1995), pp. 1557–1576.
- [16] I. GOHBERG AND I. KOLTRACHT, *Mixed, componentwise, and structured condition numbers*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 688–704.
- [17] I. GOHBERG AND I. KOLTRACHT, *Structured condition numbers for linear matrix structures*, in Linear Algebra for Signal Processing, A. Bojanczyk and G. Cybenko, eds., IMA Vol. Math. Appl. 69, Springer-Verlag, New York, 1995, pp. 17–26.
- [18] G. H. GOLUB AND C. VAN LOAN, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.
- [19] M. GU, *Stable and Efficient Algorithms for Structured Systems of Linear Equations*, Technical Report LBL-37690, Lawrence Berkeley National Laboratory, Berkeley, CA, 1995.
- [20] W. HAGER, *Condition estimates*, SIAM J. Sci. Stat. Comput., 5 (1984), pp. 311–316.
- [21] G. HEINIG, *private communication*, Kuwait University, 2000.
- [22] G. HEINIG, *Inversion of generalized Cauchy matrices and other classes of structured matrices*, in Linear Algebra for Signal Processing, A. Bojanczyk and G. Cybenko, eds., IMA Vol. Math. Appl. 69, Springer-Verlag, New York, 1994, pp. 63–81.
- [23] D. J. HIGHAM, *Condition numbers and their condition numbers*, Linear Algebra Appl., 214 (1995), pp. 193–213.
- [24] 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.
- [25] D. J. HIGHAM AND N. J. HIGHAM, *Componentwise perturbation theory for linear systems with multiple right-hand sides*, Linear Algebra Appl., 174 (1992), pp. 111–129.
- [26] N. J. HIGHAM, *Experience with a matrix norm estimator*, SIAM J. Sci. Stat. Comput., 11 (1990), pp. 804–809.
- [27] N. J. HIGHAM, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, 2002.
- [28] R. A. HORN AND CH. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
- [29] T. HU AND L. QIU, *On structured perturbation of Hermitian matrices*, Linear Algebra Appl., 275/276 (1998), pp. 287–314.
- [30] ANSI/IEEE 754-1985, *Standard for Binary Floating-Point Arithmetic*, IEEE, Piscataway, NJ, 1985.

- [31] S. V. KONYAGIN AND W. SCHLAG, *Lower bounds for the absolute value of random polynomials on a neighborhood of the unit circle*, Trans. Amer. Math. Soc., 351 (1999), pp. 4963–4980.
- [32] D. J. HIGHAM AND N. J. HIGHAM, *MATLAB Guide*, SIAM, Philadelphia, 2000.
- [33] V. OLSHEVSKY, *Pivoting for structured matrices with applications*, Linear Algebra Appl., to appear.
- [34] A. PACKARD AND J. DOYLE, *The complex structured singular value*, Automatica, 29 (1993), pp. 71–109.
- [35] L. QIU, B. BERNHARDSSON, A. RANTZER, E. J. DAVISON, AND J. C. DOYLE, *A formula for computation of the real stability radius*, Automatica, 31 (1995), pp. 879–890.
- [36] J. ROHN, *A new condition number for matrices and linear systems*, Computing, 41 (1989), pp. 167–169.
- [37] S. M. RUMP, *Structured perturbations and symmetric matrices*, Linear Algebra Appl., 278 (1998), pp. 121–132.
- [38] S. M. RUMP, *Ill-conditionedness need not be componentwise near to ill-posedness for least squares problems*, BIT, 39 (1999), pp. 143–151.
- [39] J.-G. SUN, *Bounds for the structured backward errors of Vandermonde systems*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 45–59.
- [40] P. Y. YALAMOV, *On the almost strong stability of the circular deconvolution algorithm*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 358–363.
- [41] K. ZHOU, J. DOYLE, AND K. GLOVER, *Robust and Optimal Control*, Prentice–Hall, Englewood Cliffs, NJ, 1995.