

## THE COMPONENTWISE STRUCTURED AND UNSTRUCTURED BACKWARD ERRORS CAN BE ARBITRARILY FAR APART\*

SIEGFRIED M. RUMP<sup>†</sup>

**Abstract.** Given a linear system  $Ax = b$  and some vector  $\tilde{x}$ , the backward error characterizes the smallest relative perturbation of  $A, b$  such that  $\tilde{x}$  is a solution of the perturbed system. If the input matrix has some structure such as being symmetric or Toeplitz, perturbations may be restricted to perturbations within the same class of structured matrices. For normwise perturbations, the symmetric and the general backward errors are equal, and the question about the relation between the symmetric and general *componentwise* backward errors arises. In this note we show for a number of common structures in numerical analysis that for componentwise perturbations the structured backward error can be equal to 1, whilst the unstructured backward error is arbitrarily small. Structures cover symmetric, persymmetric, skewsymmetric, Toeplitz, symmetric Toeplitz, Hankel, persymmetric Hankel, and circulant matrices. This is true although the normwise condition number  $\|A^{-1}\| \|A\|$  is close to 1.

**Key words.** backward error, structured perturbations, Rigal–Gaches, Oettli–Prager

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

**DOI.** 10.1137/140985500

**1. Introduction and notation.** The following note was triggered by a problem posed by Jim Demmel at the Householder XIX meeting 2014 in Spa, Belgium.

A backward error of a square linear system  $Ax = b$  of a given  $\tilde{x}$  measures the smallest perturbations  $\Delta A$  and  $\Delta b$  such that

$$(A + \Delta A)\tilde{x} = b + \Delta b.$$

Perturbations may be *normwise* or *componentwise*, leading to the definitions

$$(1.1) \quad \begin{aligned} \eta(A, \tilde{x}, b) &:= \min\{\epsilon : (A + \Delta A)\tilde{x} = b + \Delta b, \|\Delta A\| \leq \epsilon\|A\|, \|\Delta b\| \leq \epsilon\|b\|\}, \\ \omega(A, \tilde{x}, b) &:= \min\{\epsilon : (A + \Delta A)\tilde{x} = b + \Delta b, |\Delta A| \leq \epsilon|A|, |\Delta b| \leq \epsilon|b|\}. \end{aligned}$$

The Rigal–Gaches and Oettli–Prager theorems [3] characterize the backward errors

$$(1.2) \quad \eta(A, \tilde{x}, b) = \frac{\|A\tilde{x} - b\|}{\|A\| \|\tilde{x}\| + \|b\|} \quad \text{and} \quad \omega(A, \tilde{x}, b) = \max_i \frac{|A\tilde{x} - b|_i}{(|A| |\tilde{x}| + |b|)_i}$$

with the convention  $0/0 := 0$ . Note that  $A$  may be singular.

It is well known that both quantities can be arbitrarily far apart. An example is

$$A_\delta = \begin{pmatrix} 1 & \delta \\ 0 & 1 \end{pmatrix}, \quad \tilde{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{and} \quad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Then, using the the  $\infty$ -norm,

$$\eta(A_\delta, \tilde{x}, b) = \frac{\delta}{2 + \delta} \quad \text{and} \quad \omega(A_\delta, \tilde{x}, b) = 1.$$

---

\*Received by the editors September 8, 2014; accepted for publication (in revised form) by S. C. Eisenstat January 26, 2015; published electronically April 14, 2015.

<http://www.siam.org/journals/simax/36-2/98550.html>

<sup>†</sup>Institute for Reliable Computing, Hamburg University of Technology, Hamburg 21071, Germany, and Faculty of Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan (rump@tuhh.de).

Although the normwise backward error can be arbitrarily small, the componentwise backward error is always 1. Note that  $A_\delta$  is well conditioned for all small  $\delta$ .

If  $A$  is structured, for example symmetric, then perturbations may be restricted to the same set of structured matrices. In this paper we consider the following structures:

$$(1.3) \quad \mathcal{S} := \{\text{sym, persym, skewsym, Toep, symToep, Hankel, persymHankel, circ}\}$$

corresponding to symmetric, persymmetric, skewsymmetric, general Toeplitz, symmetric Toeplitz, general Hankel, persymmetric Hankel, and circulant matrices.

We denote by  $\mathcal{M}_n$  the set of real  $n \times n$  matrices and by  $\mathcal{M}_n^{\text{struct}}$  the matrices in the corresponding structure. The index  $n$  is omitted when clear from the context. For  $A \in \mathcal{M}^{\text{struct}}$  this leads to the *normwise structured backward error*

$$\eta^{\text{struct}}(A, \tilde{x}, b) := \min\{\epsilon : (A + \Delta A)\tilde{x} = b + \Delta b, \quad \Delta A \in \mathcal{M}^{\text{struct}}, \\ \|\Delta A\| \leq \epsilon\|A\|, \|\Delta b\| \leq \epsilon\|b\|\}.$$

It is well known [1] that for symmetric structure  $\eta^{\text{sym}}(\tilde{x}) = \eta(\tilde{x})$ . Indeed the governing equation reads

$$\Delta A \tilde{x} = b - A \tilde{x} + \Delta b =: y,$$

so that<sup>1</sup>  $\|\Delta A\| \geq \|y\|/\|\tilde{x}\| =: \alpha$ , and for  $H$  denoting the Householder matrix mapping the unit vector  $\tilde{x}/\|\tilde{x}\|_2$  into the unit vector  $y/\|y\|_2$  it follows that  $(\alpha H)^T = \alpha H$ ,  $\alpha H \tilde{x} = y$ , and  $\|\alpha H\| = \alpha$ .

Similarly, the *componentwise structured backward error* is defined by

$$\omega^{\text{struct}}(A, \tilde{x}, b) := \min\{\epsilon : (A + \Delta A)\tilde{x} = b + \Delta b, \quad \Delta A \in \mathcal{M}^{\text{struct}}, \\ |\Delta A| \leq \epsilon|A|, |\Delta b| \leq \epsilon|b|\}.$$

Subsequently, the parameters  $A$  and  $b$  will be omitted when clear from the context. Note that by using  $\Delta A := -A$  and  $\Delta b := -b$  it follows that all backward errors, general or structured, are bounded by 1 for all  $\tilde{x}$ , the latter corresponding to a 100% change of the input data.

For a given linear system  $Ax = b$  with approximate solution  $\tilde{x}$  and with symmetric matrix  $A$ , Jim Demmel asked at the Householder XIX meeting 2014 in Spa, Belgium about the relation between the general componentwise backward error  $\omega(\tilde{x})$  and the structured componentwise backward error  $\omega^{\text{sym}}(\tilde{x})$ .

In this note we show that for symmetric and for other common structures in numerical linear algebra both quantities can be arbitrarily far apart.

**2. Main result.** We show that the following is true for all structures in (1.3): The general componentwise backward error  $\omega(\tilde{x})$  may be arbitrarily small whilst the structured componentwise backward error  $\omega^{\text{struct}}(\tilde{x})$  is equal to 1. This may happen although the traditional condition number  $\|A^{-1}\|\|A\|$  is close to 1. The proof is based on similar results on general and structured componentwise condition numbers in [4].

**THEOREM 2.1.** *For  $n \geq 6$ ,  $\epsilon > 0$ , and  $\text{struct} \in \mathcal{S}$  as in (1.3) the following is true:*

$$(2.1) \quad \exists A \in \mathcal{M}_n^{\text{struct}}, \tilde{x}, b \in \mathbb{R}^n, \quad \eta(\tilde{x}) < \epsilon, \omega(\tilde{x}) < \epsilon, \quad \text{and} \quad \omega^{\text{struct}}(\tilde{x}) = 1.$$

Moreover, for  $n \geq 7$ , provided  $n$  is even for a skewsymmetric structure, the matrix  $A$  may be chosen such that the 2-norm condition number  $\|A^{-1}\|_2\|A\|_2$  is less than 3.

<sup>1</sup>If  $\tilde{x} = 0$  there is nothing to prove.

*Proof.* We present the proof for symmetric structure in more detail, the proof for the other structures follows the same pattern. Consider  $A_\delta \in \mathcal{M}^{\text{sym}}$  with  $0 < \delta < 1$  and

$$(2.2) \quad A_\delta := \begin{pmatrix} \delta & 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & -1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 1 & 0 \end{pmatrix}, \quad \tilde{x} := \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \text{and} \quad b := \begin{pmatrix} 0 \\ 0 \\ 4 \\ 0 \\ 0 \end{pmatrix}.$$

Note that  $\det(A_\delta) = 4\delta$  implies that  $A$  is ill-conditioned<sup>2</sup> for small  $\delta$ . The condition  $|\Delta A| \leq \epsilon|A|$  and  $|\Delta b| \leq \epsilon|b|$  for symmetric  $\Delta A$  is equivalent to

$$(2.3) \quad \Delta A := \begin{pmatrix} \alpha_{11}\delta & \alpha_{12} & \alpha_{13} & 0 & -\alpha_{15} \\ \alpha_{12} & 0 & \alpha_{23} & -\alpha_{24} & 0 \\ \alpha_{13} & \alpha_{23} & 0 & \alpha_{34} & \alpha_{35} \\ 0 & -\alpha_{24} & \alpha_{34} & 0 & \alpha_{45} \\ -\alpha_{15} & 0 & \alpha_{35} & \alpha_{45} & 0 \end{pmatrix} \quad \text{and} \quad \Delta b := \begin{pmatrix} 0 \\ 0 \\ 4\beta_3 \\ 0 \\ 0 \end{pmatrix}$$

subject to  $|\alpha_{ij}| \leq \epsilon$  and  $|\beta_j| \leq \epsilon$  for all  $i, j$ . A computation yields

$$(A + \Delta A)\tilde{x} - (b + \Delta b) = \begin{pmatrix} (\alpha_{11} + 1)\delta + (\alpha_{12} - \alpha_{15}) \\ \alpha_{12} - \alpha_{24} \\ \alpha_{13} + \alpha_{23} + \alpha_{34} + \alpha_{35} - 4\beta_3 \\ \alpha_{45} - \alpha_{24} \\ \alpha_{45} - \alpha_{15} \end{pmatrix}.$$

Thus  $(A + \Delta A)\tilde{x} - (b + \Delta b) = 0$  implies from the second, fourth, and fifth line that  $\alpha_{12} = \alpha_{24} = \alpha_{45} = \alpha_{15}$ , so that the first equation and  $\delta \neq 0$  give  $\alpha_{11} = -1$ . It follows

$$\omega^{\text{sym}}(\tilde{x}) = 1.$$

On the other hand

$$|A\tilde{x} - b| = \begin{pmatrix} \delta \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad |A||\tilde{x}| + |b| = \begin{pmatrix} \delta + 2 \\ 2 \\ 8 \\ 2 \\ 2 \end{pmatrix},$$

so that (1.2) implies

$$\eta(\tilde{x}) = \frac{\delta}{8} \quad \text{and} \quad \omega(\tilde{x}) = \frac{\delta}{2 + \delta}.$$

The example proves the result for symmetric structure and  $n = 5$ . The result extends to dimension  $n + k$  by appending an arbitrary symmetric  $k \times k$  matrix to the lower right of  $A_\delta$ , setting the other components in the last  $k$  rows to zero, and expanding  $\tilde{x}$  and  $b$  by  $k$  zeros. This finishes the proof for a symmetric structure.

---

<sup>2</sup>This is true in that example because  $|\det(A_\delta)| = \prod_{i=1}^5 \sigma_i(A_\delta) \geq \sigma_1(A_\delta)\sigma_5(A_\delta)^4 \geq \sigma_5(A_\delta)^4$  and therefore  $\text{cond}_2(A_\delta) = \sigma_1/\sigma_5 \geq (4\delta)^{-1/4}$ .

Denote by  $J \in \mathcal{M}_n$  the “flip” matrix, that is the permutation matrix mapping  $(1, 2, \dots, n)^T$  into  $(n, \dots, 2, 1)^T$ . Replacing  $A_\delta$  and  $b$  in (2.2) by  $JA$  and  $Jb$ , respectively, proves the theorem for a persymmetric structure.

For a skewsymmetric structure consider  $A_\delta \in \mathcal{M}^{\text{skewsym}}$  with  $0 < \delta < 1$  and

$$(2.4) \quad A_\delta := \begin{pmatrix} 0 & \delta & 0 & 0 \\ -\delta & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & 1 & 0 \end{pmatrix}, \quad \tilde{x} := \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \text{and} \quad b := \begin{pmatrix} \delta \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The matrix is ill-conditioned for small  $\delta$  as  $\det(A_\delta) = \delta^2$ . Defining  $\Delta A$  and  $\Delta b$  similarly to (2.3) we obtain

$$(A + \Delta A)\tilde{x} - (b + \Delta b) = \begin{pmatrix} (\alpha_{12} - \beta_1) \delta \\ (-\alpha_{12} - 1) \delta + (\alpha_{24} - \alpha_{23}) \\ \alpha_{23} - \alpha_{34} \\ \alpha_{34} - \alpha_{24} \end{pmatrix} = 0.$$

The third and fourth equations imply  $\alpha_{23} = \alpha_{34} = \alpha_{24}$ , so that the second equation yields  $\alpha_{12} = -1$ . Hence  $\omega^{\text{skewsym}}(\tilde{x}) = 1$ . On the other hand,

$$|A\tilde{x} - b| = \begin{pmatrix} 0 \\ \delta \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad |A||\tilde{x}| + |b| = \begin{pmatrix} 2\delta \\ \delta + 2 \\ 2 \\ 2 \end{pmatrix},$$

so that (1.2) implies

$$\eta(\tilde{x}) = \frac{\delta}{2 + 2\delta} \quad \text{and} \quad \omega(\tilde{x}) = \frac{\delta}{2 + \delta}.$$

Appending an arbitrary  $k \times k$  skewsymmetric matrix to the lower right of  $A_\delta$  and expanding  $\tilde{x}$  and  $b$  by  $k$  zeros proves the result for dimension  $n + k$ . Note that for odd dimension a real skewsymmetric matrix is necessarily singular.

For  $0 < \delta < 1$  define

$$(2.5) \quad \begin{aligned} a_\delta &:= (0 \ 0 \ -1 \ 1 \ 1 \ \delta), \\ \tilde{x} &:= (1 \ 0 \ -1 \ -1 \ 0 \ 1)^T, \\ b &:= (0 \ 2 \ 0 \ 0 \ 2 \ 0)^T, \end{aligned}$$

and let  $A_\delta$  be the symmetric Toeplitz matrix with first row  $a_\delta$ . The matrix is ill-conditioned for small  $\delta$  as  $\det(A_\delta) = -8\delta - \delta^2$ . Proceeding as before we obtain

$$(A + \Delta A)\tilde{x} - (b + \Delta b) = \begin{pmatrix} (\alpha_{16} + 1) \delta + (\alpha_{13} - \alpha_{14}) \\ \alpha_{13} + \alpha_{15} - 2\beta_2 \\ \alpha_{14} - \alpha_{13} \\ \alpha_{14} - \alpha_{13} \\ \alpha_{13} + \alpha_{15} - 2\beta_5 \\ (\alpha_{16} + 1) \delta + (\alpha_{13} - \alpha_{14}) \end{pmatrix} = 0,$$

and the third and first equations yield  $\alpha_{16} = -1$ . Hence  $\omega^{\text{symToep}}(\tilde{x}) = 1$ . On the other hand,

$$|A\tilde{x} - b| = \begin{pmatrix} \delta \\ 0 \\ 0 \\ 0 \\ 0 \\ \delta \end{pmatrix} \quad \text{and} \quad |A||\tilde{x}| + |b| = \begin{pmatrix} \delta + 2 \\ 4 \\ 2 \\ 2 \\ 4 \\ \delta + 2 \end{pmatrix},$$

so that (1.2) implies

$$\eta(\tilde{x}) = \frac{\delta}{5 + \delta} \quad \text{and} \quad \omega(\tilde{x}) = \frac{\delta}{2 + \delta}.$$

An example with a symmetric Toeplitz matrix of dimension  $6 + k$  is constructed as follows:  $a_\delta$  in (2.5) is appended by arbitrary  $k$  real numbers,  $\tilde{x}$  is appended by  $k$  zeros, and the last  $k$  components of  $b$  are adapted according to a perturbation  $\Delta A$  realizing the minimum.

To prove Theorem 2.1 for a persymmetric Hankel structure, we replace the quantities  $A_\delta, \tilde{x}$ , and  $b$  from (2.5) by  $JA_\delta, \tilde{x}$ , and  $Jb$ , respectively, using  $J$  as before.

To treat general Hankel structures define

$$(2.6) \quad \begin{aligned} c_\delta &:= (\delta \quad 1 \quad 1 \quad -1 \quad 0 \quad 0), \\ r_\delta &:= (0 \quad 0 \quad -1 \quad 1 \quad 1 \quad 0), \\ \tilde{x} &:= (1 \quad 0 \quad -1 \quad -1 \quad 0 \quad 1)^T, \\ b &:= (0 \quad 2 \quad 0 \quad 0 \quad 2 \quad 0)^T \end{aligned}$$

with  $0 < \delta < 1$ , and let  $A_\delta$  be the Hankel matrix with first column  $c_\delta^T$  and last row  $r_\delta$ . The matrix is ill-conditioned for small  $\delta$  as  $\det(A_\delta) = 4\delta$ . Proceeding as before we obtain

$$(A + \Delta A)\tilde{x} - (b + \Delta b) = \begin{pmatrix} (\alpha_{11} + 1)\delta + (\alpha_{41} - \alpha_{31}) \\ \alpha_{21} + \alpha_{41} - 2\beta_2 \\ \alpha_{31} - \alpha_{63} \\ \alpha_{64} - \alpha_{41} \\ \alpha_{63} + \alpha_{65} - 2\beta_5 \\ \alpha_{63} - \alpha_{64} \end{pmatrix},$$

and the third, last, and fourth equations yield  $\alpha_{31} = \alpha_{63} = \alpha_{64} = \alpha_{41}$ , so that  $\alpha_{11} = -1$  by the first equation. Hence  $\omega^{\text{Hankel}}(\tilde{x}) = 1$ . On the other hand,

$$|A\tilde{x} - b| = \begin{pmatrix} \delta \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad |A||\tilde{x}| + |b| = \begin{pmatrix} \delta + 2 \\ 4 \\ 2 \\ 2 \\ 4 \\ 2 \end{pmatrix},$$

and (1.2) implies

$$\eta(\tilde{x}) = \frac{\delta}{5 + \delta} \quad \text{and} \quad \omega(\tilde{x}) = \frac{\delta}{2 + \delta}.$$

An example with a Hankel matrix of dimension  $6 + k$  is constructed as follows:  $c_\delta$  in (2.6) is preceded by arbitrary  $k$  real numbers,  $r_\delta$  in (2.5) is appended by arbitrary  $k$  real numbers,  $\tilde{x}$  is preceded by  $k$  zeros, and the last  $k$  components of  $b$  are adapted as in the case of symmetric Toeplitz matrices.

The case of general Toeplitz structures follows by using  $JA_\delta, \tilde{x}$ , and  $Jb$  with the data from the previous example.

For circulant structure and  $n \geq 3$  define for  $0 < \delta < 1$

$$(2.7) \quad \begin{aligned} a_\delta &:= \left( -1 \quad 1 + \delta \quad 0 \quad \dots \quad 0 \right)^T, \\ \tilde{x} &:= \left( 1 + \delta \quad 1 \quad \dots \quad 1 \right)^T, \\ b &:= \left( 0 \quad 2\delta \quad 0 \quad \dots \quad 0 \right)^T, \end{aligned}$$

where the values left and right of the dots are repeated so that  $a_\delta, \tilde{x}, b \in \mathbb{R}^n$ . Let  $A_\delta$  be the circulant matrix with first column  $a_\delta$ . Then  $\det(A_\delta)$  is of order  $n\delta$ , so that  $A_\delta$  is ill-conditioned for small  $n$  and  $\delta$ . Proceeding as before we obtain

$$(2.8) \quad \begin{aligned} [(A + \Delta A)\tilde{x} - (b + \Delta b)]_1 &= (\alpha_{1n} - \alpha_{11}) (1 + \delta), \\ [(A + \Delta A)\tilde{x} - (b + \Delta b)]_3 &= (\alpha_{1n} - \alpha_{11}) + \delta(1 + \alpha_{1n}), \end{aligned}$$

so that  $\alpha_{1n} = -1$ . It follows  $\omega^{\text{circ}}(\tilde{x}) = 1$ . On the other hand,

$$|A\tilde{x} - b| = \begin{pmatrix} 0 \\ \delta^2 \\ \delta \\ \dots \\ \delta \end{pmatrix} \quad \text{and} \quad |A||\tilde{x}| + |b| = \begin{pmatrix} 2 + 2\delta \\ 2 + 4\delta + \delta^2 \\ 2 + \delta \\ \dots \\ 2 + \delta \end{pmatrix},$$

and (1.2) implies

$$\eta(\tilde{x}) = \frac{\delta}{2 + 5\delta + \delta^2} \quad \text{and} \quad \omega(\tilde{x}) = \frac{\delta}{2 + \delta}.$$

This finishes the proof of the first part of Theorem 2.1. It remains to show that for  $n \geq 7$  there exist similar examples for all structures with small condition number  $\|A^{-1}\| \|A\|$ .

Note that for all structures except circulants, the presented examples (2.2), (2.4), (2.5), (2.6) of small dimension can be extended by arbitrary real numbers as long as the structure is preserved. Let  $A_\delta \in \mathcal{M}_5$  be the symmetric matrix in (2.2), and define

$$B_\delta = \begin{pmatrix} A_\delta & z \\ z^T & 0 \end{pmatrix} \in \mathcal{M}_6, \quad \tilde{y} = \begin{pmatrix} \tilde{x} \\ 0 \end{pmatrix}, \quad \text{and} \quad c = \begin{pmatrix} b \\ 0 \end{pmatrix}.$$

We choose  $z := (1, -1, 0, 1, -1)^T$ , which is close to a singular vector of  $A_\delta^{-1}$ . Then  $B_\delta$  has singular values  $0.5(\sqrt{16 + \delta^2} \pm \delta)$  plus a 4-fold singular value 2. Thus

$$\kappa_2(B_\delta) = \|B_\delta\|_2 \|B_\delta^{-1}\|_2 = \frac{4}{\sqrt{16 + \delta^2} - \delta} < \frac{4}{4 - \delta}.$$

To obtain examples of larger dimension,  $B_\delta$  is extended to the lower right by an identity matrix, whereas  $\tilde{y}$  and  $c$  are filled with zeros. This covers the persymmetric structure as well.

For a skewsymmetric structure we define (the zeros are of proper dimension)

$$B_\delta := \begin{pmatrix} A_\delta & Z \\ -Z^T & 0 \end{pmatrix} \in \mathcal{M}_6, \tilde{y} := \begin{pmatrix} \tilde{x} \\ 0 \end{pmatrix}, c := \begin{pmatrix} b \\ 0 \end{pmatrix} \text{ with } Z := \begin{pmatrix} \sqrt{3} & \sqrt{3} \\ -1 & 1 \\ -1 & 1 \\ -1 & 1 \end{pmatrix},$$

using the matrix  $A_\delta \in \mathcal{M}_4$  as in (2.4). The characteristic polynomial of  $B_\delta$  computes to  $(x^2 + 3)(x^2 + 6)^2$ , so that  $\kappa_2(B_\delta) = \sqrt{2}$ . Examples of larger dimension are constructed as before.

Using MATLAB notation and  $z_k := \text{zeros}(1, k)$ , consider symmetric Toeplitz matrices  $T_m \in \mathcal{M}_m^{\text{symToep}}$  for  $m \in \{2k, 2k + 1\}$  with first row

$$\begin{bmatrix} z_k & 1 & z_{k-1} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} z_k & 1 & z_{k-1} & 1 \end{bmatrix},$$

respectively. Then  $T_m$  is a permutation matrix with  $\kappa_2(T_{2k}) = 1$ , and  $T_{2k+1}$  has a singular value 2 and  $2k$  singular values 1, so  $\kappa_2(T_{2k+1}) = 2$ . For  $m \geq 7$ , the upper left  $(6 \times 6)$ -block of  $T_m$  is zero. Replacing  $T_m$  by  $\beta T_m$  for large  $\beta$ , and replacing the upper left block by  $A_\delta \in \mathcal{M}_6^{\text{symToep}}$  as in (2.5) produces, using a perturbation argument, a symmetric Toeplitz matrix with condition numbers arbitrarily close to 1 and 2, respectively. Adapting the approximate solution  $\tilde{x}$  and right-hand side  $b$  as before produces the desired examples for  $n \geq 12$ .

For  $7 \leq n \leq 11$  extending the first row  $a_\delta$  by

$$1, [0.75 \ -1], [1.5 \ -1 \ 0], [1 \ -0.5 \ 0.5 \ 1.5], [0.75 \ -0.5 \ 0.5 \ 1 \ -1],$$

respectively, produces matrices with (2.1) and condition number less than 3. This covers the persymmetric Hankel structure as well.

Consider Hankel matrices  $H_m \in \mathcal{M}_m^{\text{Hankel}}$  for  $m \in \{2k, 2k + 1\}$  with first column and last row

$$\begin{bmatrix} v & 0 \end{bmatrix}, \begin{bmatrix} 0 & v \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & v & 0 \end{bmatrix}, \begin{bmatrix} 0 & v & 1 \end{bmatrix},$$

respectively, where  $v = [\text{zeros}(1, k - 1) \ 1 \ \text{zeros}(1, k - 1)]$ . As before it is verified that the corresponding condition numbers are 1 and 2, respectively, and that the right upper  $(6 \times 6)$ -block is zero. Then we proceed as before to produce the desired examples for  $n \geq 12$ .

For  $7 \leq n \leq 11$  we precede the first column  $c_\delta$  in (2.6) by

$$1, [-1 \ 1], [0 \ 0 \ 2], [1.5 \ 1 \ -0.5 \ 0.5] \ [-0.5 \ 1 \ 0 \ -0.25 \ 0.75]$$

and append

$$1, [1 \ -1], [1 \ -1 \ 0], [1 \ -0.5 \ 0 \ 0.5], [0.75 \ -0.5 \ 0.75 \ 1.25 \ 0.25]$$

to the last row  $r_\delta$ , respectively. This produces matrices with (2.1) and condition number less than 3. That covers the general Toeplitz structure as well.

The case of circulant structures remains. For  $n \geq 7$  we define

$$\begin{aligned} (2.9) \quad A_\delta &:= -I + \beta P^3 + (1 + \delta)P^{n-1}, \\ \tilde{x} &:= (1 + \delta)e_1 + e_2 + e_3 + e_n, \\ b &:= 2\delta e_2 + e_4 + \beta e_{n-3} + (1 + \delta)\beta e_{n-2} + \beta e_{n-1} + (\beta - 1)e_n, \end{aligned}$$

where  $\beta \in \mathbb{R}$  is large positive,  $I \in \mathcal{M}_n$  is the identity matrix with columns  $e_i$ , and  $P$  denotes the permutation matrix mapping  $(1, \dots, n)$  into  $(2, \dots, n, 1)$ . A computation shows  $A_\delta \tilde{x} - b = \delta^2 e_2 + \delta(e_3 + e_n)$ , so that  $\|A_\delta \tilde{x} - b\|_\infty = \delta$  for small  $\delta$ .

Furthermore, (2.8) is satisfied again. Note that this is only true for  $n \geq 7$ , where  $e_4 = e_{n-3}$  for  $n = 7$ . Hence  $\omega^{\text{circ}}(\tilde{x}) = 1$ , and it follows

$$\eta^{\text{circ}}(\tilde{x}) = \frac{\delta}{(2 + 2\beta + \delta)(1 + \delta)} \quad \text{and} \quad \omega^{\text{circ}}(\tilde{x}) = \frac{\delta}{2 + \delta}.$$

That shows (2.1). For large  $\beta$  the matrix  $A_\delta$  is a small perturbation of a scaled permutation matrix, so that its condition number is arbitrarily close to 1.

That finishes the proof.  $\square$

**3. Open questions.** For several structures and any dimension  $n \geq 6$  examples of linear systems  $Ax = b$  were presented such that for given  $\tilde{x}$  the *general* componentwise backward error  $\omega(\tilde{x})$  is arbitrarily small, whilst the *structured* componentwise backward error  $\omega^{\text{struct}}(\tilde{x})$  is equal to 1. That may be true for a perfectly well-conditioned input matrix in the usual normwise sense.

Are there (computable) conditions on the matrix and/or the structure to bound  $\omega^{\text{struct}}(\tilde{x})/\omega(\tilde{x})$ ?

Linear systems with structured matrix are usually solved with specialized algorithms, often resulting in a considerable reduction of the computational effort. For example [2], the inverse of a symmetric positive definite Toeplitz matrix can be computed in  $\mathcal{O}(n^2)$  operations, the time to print the entries.

Often such algorithms are analyzed w.r.t. general perturbations. Might there be, as Mario Arioli asked, a “structured iterative refinement” method providing an update of  $\tilde{x}$  with satisfactory small structured componentwise backward error?

There are several results [5, 6] on general and structured (normwise) perturbations based on scalar product spaces. Regarding componentwise perturbations, are there results related to the existence of a corresponding multiplicative Lie group?

**Acknowledgments.** I would like to thank Florian Bünger, Marko Lange, Marko Lindner, and Jens Zemke for fruitful and inspiring discussions. Moreover, my dearest thanks to the anonymous referees for their constructive and fruitful remarks.

#### REFERENCES

- [1] J.R. BUNCH, J.W. DEMMEL, AND C.F. VAN LOAN, *The strong stability of algorithms for solving symmetric linear systems*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 494–499.
- [2] G.H. GOLUB AND CH. VAN LOAN, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.
- [3] N. J. HIGHAM, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, 2002.
- [4] S.M. RUMP, *Structured perturbations part II: Componentwise distances*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 31–56.
- [5] F. TISSEUR AND S. GRAILLAT, *Structured condition numbers and backward errors in scalar product spaces*, Electron. J. Linear Algebra, 15 (2006), pp. 159–177.
- [6] F. TISSEUR AND N.J. HIGHAM, *Structured pseudospectra for polynomial eigenvalue problems, with applications*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 187–208.