Interval Newton Operators for Function Strips

R. KRAWCZYK

Bohlweg 2, D-3392 Clausthal-Zellerfield, West Germany

AND

A. NEUMAIER

Institut für Angewandte Mathematik, Universität Freiburg, Hermann-Herder-Strasse 10, D-7800 Freiburg, West Germany

Submitted by E. Stanley Lee

Received October 28, 1985

1. INTRODUCTION

Recently the problem of enclosing the zero set of a system of equations depending on inaccurately known data has found considerably attention, e.g., in the following references: [1, 5, 6, 8, 9, 10, 12]. In [12], common features of zero enclosing methods for data independent systems were described using the notion of the inverse of an interval matrix. There, if $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a real function, $\tilde{x} \in X \subseteq D$, and A^I denotes an inverse of a regular interval Lipschitz matrix A for f on X, the unique zero $x^* \in X$ of f (if it exists) satisfies

$$x^* \in \tilde{x} - A^I f(\tilde{x}).$$

In the present paper we extend this approach to a function strip $G(x) := [\underline{g}(x), \bar{g}(x)]$ representing an inaccurately known system of equations and show that under appropriate conditions the zero set X^* of $G(x)$ satisfies

$$X^* \subseteq \tilde{x} - A^I G(\tilde{x}).$$

Therefore, if $X^* \subseteq X_0$, the iteration

$$X_{k+1} := (\tilde{x}_k - A^I G(\tilde{x}_k)) \cap X_k, \text{ where } \tilde{x}_k \in X_k,$$

yields a sequence of nested intervals X_k containing X^*. Under suitable assumptions this sequence is shown to converge to a fixed interval $\hat{X}$; a further improvement is generally not possible by this method.

0022-247X/87 $3.00

After some preparations in Section 2 we define the function strip G and its zero set in Section 3 and prove some of its properties. In Section 4 we enclose the zero set X^* by an interval X^I and give bounds for the distance $\operatorname{rad} X^I - \operatorname{rad} \square X^*$. By applying these results to a fixed interval of the interval Newton operator defined in Section 5 we obtain quadratical convergence of the distance $\operatorname{rad} \hat{X} - \operatorname{rad} \square X^*$ for $\operatorname{rad} G(x) \to 0$. In Section 6 we give conditions for the convergence of the iterated sequence $\{X_k\}$ to a fixed interval $\hat{X}$. Finally, some examples of inverses and their influence on fixed intervals and convergence matrices are considered in Section 7.

2. Notation and Basic Concepts

The terminology of the paper follows Krawczyk and Neumaier [10] and Neumaier [11]; but for convenience of the reader, some definitions are repeated. Lower case italic letters denote real values and upper case italic letters denote sets, intervals, and maps. We denote the set of n-dimensional interval vectors and $n \times n$-interval matrices by $\mathbb{IR}^n$ and $\mathbb{IR}^{n \times n}$, respectively, use mid X, mid A (or $\check{x}$, $\check{a}$) for the midpoints, rad X, rad A for the radius, and $|X|$, $|A|$ for the absolute value of $X \in \mathbb{IR}^n$ and $A \in \mathbb{IR}^{n \times n}$, respectively.

Moreover, we set $\mathbb{I}D := \{X \in \mathbb{IR}^n \mid X \subseteq D\}$. The unit matrix is written as e, and $E := [-e, e]$. $\sigma(a)$ denotes the spectral radius of $a \in \mathbb{R}^{n \times n}$. For bounded sets $\Sigma \subseteq \mathbb{R}^n$, the symbol $\square\Sigma := [\inf \Sigma, \sup \Sigma] \in \mathbb{IR}^n$ denotes the hull of Σ. The (vector-valued) distance of two intervals $X_1 Y \in \mathbb{IR}^n$ is defined by

$$q(X_1 Y) := \inf\{\tilde{q} \in \mathbb{R}^n \mid \hat{q} \geqslant 0, X \subseteq Y + [-\tilde{q}, \tilde{q}], \quad Y \subseteq X + [-\tilde{q}, \tilde{q}]\} \tag{2.1}$$

(see (2) in [13]).

It will be assumed that the reader is familiar with interval arithmetic and the rules for radius, midpoint, and absolute value of intervals. However, we recall the equivalence

$$Y \supseteq X \Leftrightarrow |\operatorname{mid} Y - \operatorname{mid} X| \leqslant \operatorname{rad} Y - \operatorname{rad} X. \tag{2.2}$$

Moreover, we state the following

Lemma 2.1. *If $a \in A \in \mathbb{IR}^{n \times n}$ then*

$$|A| \leqslant |a| + 2 \operatorname{rad} A. \tag{2.3}$$

Proof. From $\underline{a} \leqslant a \leqslant \bar{a}$ it follows that

$$\bar{a} \leqslant a + \bar{a} - \underline{a} \quad \text{and} \quad -\underline{a} \leqslant -a + \bar{a} - \underline{a},$$

hence

$$|A| = \sup(\bar{a}, -\underline{a}) \leqslant \sup(a + \bar{a} - \underline{a}, -a + \bar{a} - \underline{a})$$
$$= \sup(a, -a) + \bar{a} - \underline{a} = |a| + 2 \operatorname{rad} A. \quad \blacksquare$$

If $A \in \mathbb{IR}^{n \times n}$ then $\langle A \rangle$ denotes the Ostrowski operator (see Sect. 2.2 in [11]). We call $A \in \mathbb{IR}^{n \times n}$ an H-matrix if $\langle A \rangle$ is an M-matrix.

A map S: $\mathbb{IR}^n \to \mathbb{IR}^n$ is called sublinear if the following axioms are valid for all $X, Y \in \mathbb{IR}^n$:

(S1) $X \subseteq Y \Rightarrow SX \subseteq SY$ (inclusion isotonicity),

(S2) $\alpha \in \mathbb{R} \Rightarrow S(X\alpha) = (SX)\alpha$ (homogeneity),

(S3) $S(X + Y) \subseteq SX + SY$ (subadditivity).

We extend S to matrix arguments by applying it to each column of $A \in \mathbb{IR}^{n \times n}$. Moreover, we put

$$\kappa(S) := Se \quad \text{and} \quad |S| := |SE|; \tag{2.4}$$

hence $\kappa(S)$, $|S| \in \mathbb{IR}^{n \times n}$, and the relations

$$Se = \kappa(S)\, e \quad \text{and} \quad SE = |S|\, E \tag{2.5}$$

are valid.

A sublinear map is called *normal*, if for all $X \in \mathbb{IR}^n$,

(S4) $\operatorname{rad}(SX) \geqslant |S| \operatorname{rad} X$;

it is called *centered* if

(S5) $X \in \mathbb{IR}^n$, $\operatorname{mid}(SX) = 0 \Rightarrow \operatorname{mid} X = 0$,

and *regular* if

(S6) $x \in \mathbb{R}^n$, $0 \in Sx \Rightarrow x = 0$.

A sublinear map A^I is called an *inverse* of the regular matrix $A \in \mathbb{IR}^{n \times n}$ if

$$a^{-1}x \in A^I X \quad \text{for all } a \in A,\ x \in X. \tag{2.6}$$

In particular, the *hull inverse* A^H of A, defined by

$$A^H X := \square\{a^{-1}x \mid a \in A,\ x \in X\}, \tag{2.7}$$

is a sublinear map [11, Example 4].

THEOREM 2.2. *Let* $S\colon \mathbb{IR}^n \to \mathbb{IR}^n$ *be a sublinear map, and define* $\bar{S}\colon \mathbb{IR}^n \to \mathbb{IR}^n$ *by*

$$\bar{S}X := S\check{x} + \kappa(S)(E \operatorname{rad} X). \tag{2.8}$$

Then $\bar{S}$ *is an normal sublinear map and the following holds:*

(i) $Sx = \bar{S}x \subseteq \bar{S}X$ *for all* $x \in X \in \mathbb{IR}^n$.

(ii) $|\bar{S}| = |\kappa(S)| \leqslant |S|$.

(iii) *If* S *is regular (centered) then* $\bar{S}$ *is also regular (centered).*

(iv) *In particular, if* $S = A^I$ *is an inverse of* $A \in \mathbb{IR}^{n\times n}$ *then* $A^{\bar{I}} := \bar{S}$ *is again an inverse of* A.

Proof. (S1): Let $X \subseteq Y$. By (S3) and (3.8) in [11] we have

$$S\check{x} \subseteq S\check{y} + S(\check{x} - \check{y}) \subseteq S\check{y} + \kappa(S)(\check{x} - \check{y}),$$

and in view of $\check{x} - \check{y} \in E(\operatorname{rad} Y - \operatorname{rad} X)$ we obtain

$$\begin{aligned}\bar{S}X = S\check{x} + \kappa(S)(E \operatorname{rad} X) &\subseteq S\check{y} + \kappa(S)(E(\operatorname{rad} Y - \operatorname{rad} X)) \\ &+ \kappa(S)(E \operatorname{rad} X) = S\check{y} + \kappa(S)(E \operatorname{rad} Y) = \bar{S}Y\end{aligned}$$

(because in this case the distributive law is valid).

(S2): From (S2) for S and Eqs. (2.5) and (2.8) we obtain

$$\begin{aligned}\bar{S}(X\alpha) &= S(\check{x}\alpha) + |\kappa(S)|(E \operatorname{rad}(X\alpha)) \\ &= (S\check{x})\,\alpha + |\kappa(S)|\,(E \operatorname{rad} X)\,\alpha = (\bar{S}X)\,\alpha\end{aligned}$$

since $E \operatorname{rad}(X\alpha) = E(\operatorname{rad} X|\alpha|) = (E \operatorname{rad} X)\,\alpha$.

(S3): In view of the subadditivity of S and $\kappa(S)$, (2.8) implies

$$\begin{aligned}\bar{S}(X+Y) &= S(\check{x}+\check{y}) + \kappa(S)\,E(\operatorname{rad}(X+Y)) \\ &\subseteq S\check{x} + S\check{y} + \kappa(S)(E \operatorname{rad} X + E \operatorname{rad} Y) \\ &\subseteq S\check{x} + \kappa(S)(E \operatorname{rad} X) + S\check{y} + \kappa(S)(E \operatorname{rad} Y) \\ &= \bar{S}X + \bar{S}Y.\end{aligned}$$

(S4): By (2.4) and (2.8) we get

$$|\bar{S}| = |\kappa(S)(E \operatorname{rad} E)| = |\kappa(S)\,E| = |\kappa(S)|,$$

hence

$$\begin{aligned}\operatorname{rad}(\bar{S}X) &= \operatorname{rad}(S\check{x}) + |\kappa(S)| \operatorname{rad} X \\ &\geqslant |\kappa(S)| \operatorname{rad} X = |\bar{S}| \operatorname{rad} X.\end{aligned}$$

Therefore $\bar{S}$ is sublinear and normal. Now (i) is obvious, and (ii) is a consequence of (3.7) of [11]. If (S5) holds for S then it holds for $\bar{S}$ since $\operatorname{mid} \bar{S}X = \operatorname{mid} S\check{x} = 0$ implies $\operatorname{mid} X = \check{x} = 0$, and if (S6) holds for S then it holds for $\bar{S}$ since $\bar{S}x = Sx$ for $x \in \mathbb{R}^n$. This implies (iii).

To prove (iv), let A^I be an inverse of A and suppose that $a \in A$, $x \in X$. Then by (3.8) of [11] and (2.8) we have

$$\begin{aligned}a^{-1}x \in A^I x \subseteq A^I\check{x} + A^I(x - \check{x}) &\subseteq A^I\check{x} + \kappa(A^I)(x - \check{x}) \\ &\subseteq A^I\check{x} + \kappa(A^I)(E \operatorname{rad} X) = A^{\bar{I}}X.\end{aligned}$$

Hence $A^{\bar{I}}$ is an inverse of A and (iv) holds. ∎

Later we shall need some properties of the Gauss inverse A^G, defined by Gauss elimination (see [11, Sect. 5]).

PROPOSITION 2.3. *Let* $A \in \mathbb{IR}^{n\times n}$ *be a regular matrix possessing a triangular decomposition* (L, R). *Then the Gauss inverse* A^G *is centered.*

Proof. By definition of A^G we have $Z := A^G X = R^F L^F X = R^F Y$, where $Y := L^F X$. In terms of the components, the relations

$$\begin{aligned}Y_i &= (X_i - L_{i1} Y_1 - \cdots - L_{i,i-1} Y_{i-1})/L_{ii}, \\ Z_i &= (Y_i - R_{in} Z_n - \cdots - R_{i,i+1} Z_{i+1})/R_{ii}\end{aligned}$$

hold for $i = 1, \ldots, n$. If $\operatorname{mid}(A^G X) = 0$ then $\operatorname{mid} Z_i = 0$, and by Proposition 2.2 in [10],

$$\operatorname{mid}(Y_i - R_{in} Z_n - \cdots - R_{i,i+1} Z_{i+1}) = 0.$$

Therefore $\operatorname{mid} Y_i = 0$ for $i = 1, \ldots, n$ and by the same argument $\operatorname{mid} X_i = 0$ for $i = 1, \ldots, n$. Hence $\operatorname{mid} X = 0$, and (S5) holds for $S = A^G$. ∎

LEMMA 2.4. *Suppose that*

$$A = \begin{pmatrix} D & Z^T \\ S & A' \end{pmatrix} \in \mathbb{IR}^{n\times n}, \qquad X = \begin{pmatrix} U \\ X' \end{pmatrix} \in \mathbb{IR}^n,$$

where the partitioning is such that $D, U \in \mathbb{IR}$, *and let*

$$\Sigma(A) = A' - SD^{-1}Z^T$$

be the Schur complement of A. *If* A^G *exists then* $A^G X$ *can be expressed in terms of*

$$Y := \Sigma(A)^G (X' - SD^{-1}U)$$

as

$$A^G X = \begin{pmatrix} (U - Z^T Y)/D \\ Y \end{pmatrix}.$$

Proof. By the discussion in Section 5 of [11], the triangular decomposition of A has the form (L, R) with

$$L = \begin{pmatrix} 1 & 0 \\ SD^{-1} & L' \end{pmatrix}, \qquad R = \begin{pmatrix} D & Z^T \\ 0 & R' \end{pmatrix}, \tag{2.9}$$

where (L', R') is the triangular decomposition of $\Sigma(A)$. Now it is easy to see that

$$L^F \begin{pmatrix} U \\ X' \end{pmatrix} = \begin{pmatrix} U \\ X'' \end{pmatrix}, \qquad R^F \begin{pmatrix} U \\ X'' \end{pmatrix} = \begin{pmatrix} V \\ Y \end{pmatrix},$$

where

$$X'' = L'^F(X' - SD^{-1}U),$$
$$Y = R'^F X'' = R'^F L'^F (X' - SD^{-1}U) = \Sigma(A)^G (X' - SD^{-1}U),$$
$$V = (U - Z^T Y)/D.$$

Therefore

$$A^G X = R^F L^F \begin{pmatrix} U \\ X' \end{pmatrix} = \begin{pmatrix} V \\ Y \end{pmatrix} = \begin{pmatrix} (U - Z^T Y)/D \\ Y \end{pmatrix}. \qquad \blacksquare$$

PROPOSITION 2.5. *Let* $A \in \mathbb{IR}^{n \times n}$ *be an H-matrix with* mid $A = e$. *Then* A *has a triangular decomposition* (L, R) *with*

$$\text{mid } L = \text{mid } R = e, \qquad \langle L \rangle \langle R \rangle = \langle A \rangle, \tag{2.10}$$

and the relation

$$A^G X \subseteq X + E(\langle A \rangle^{-1} - e)|X| \tag{2.11}$$

holds for all $X \in \mathbb{IR}^n$.

Proof. By Alefeld [2], the triangular decomposition (L, R) of A exists. We prove (2.10) and (2.11) by induction on n.

First, suppose that $n = 1$. Since mid $A = e$ we may write $A = [1 - r, 1 + r]$ with a number $r \in \mathbb{R}$, $0 \leqslant r < 1$. We have $L = 1$, $R = A$; hence (2.10) holds trivially. To show (2.11) we observe that

$$\langle A \rangle = 1 - r, \qquad \langle A \rangle^{-1} - e = r/(1 - r).$$

Now if $|X| = |\bar{x}|$ then for a suitable sign,

$$\begin{aligned} A^G X &= [\underline{x}/(1 \pm r), \bar{x}/(1 - r)] \\ &= [\underline{x} \mp \underline{x}r/(1 \pm r), \bar{x} + \bar{x}r/(1 - r)] \\ &\subseteq [\underline{x} - \bar{x}r/(1 - r), \bar{x} + \bar{x}r/(1 - r)] \\ &= X + [-1, 1]\, r/(1 - r)|X|, \end{aligned}$$

and if $|X| = |\underline{x}|$ then a similar argument applies. Hence the proposition holds for $n = 1$.

Now we assume that $n > 1$ and the proposition holds in dimension less than n. With the notation as in Lemma 2.4 and its proof, we have $0 \notin D$ since A is an H-matrix, and mid $D = 1$, mid $B = e'$ (the $(n-1) \times (n-1)$ unit matrix), mid $S =$ mid $Z = 0$ since mid $A = e$. The Schur complement $\Sigma(A) = A' - SD^{-1}Z^T = A' + E|S|\langle D \rangle^{-1}|Z|^T$ therefore satisfies mid $\Sigma(A) = e'$, and it is an H-matrix by Proposition 6(iii) of [11]. Since (L', R') is the triangular decomposition of $\Sigma(A)$, the induction hypothesis implies

$$\text{mid } L' = \text{mid } R' = e',$$

$$\langle L' \rangle \langle R' \rangle = \langle \Sigma(A) \rangle = \langle A' \rangle - |S| \langle D \rangle^{-1} |Z|^T.$$

From (2.9) we now deduce immediately that mid $L =$ mid $R = e$ and

$$\begin{aligned} \langle L \rangle \langle R \rangle &= \begin{pmatrix} 1 & 0 \\ -|S|\langle D \rangle^{-1} & \langle L' \rangle \end{pmatrix} \begin{pmatrix} \langle D \rangle & -|Z|^T \\ 0 & \langle R' \rangle \end{pmatrix} \\ &= \begin{pmatrix} \langle D \rangle & -|Z|^T \\ -|S| & \langle \Sigma(A) \rangle + |S| \langle D \rangle^{-1} |Z|^T \end{pmatrix} = \langle A \rangle \end{aligned} \tag{2.12}$$

so that (2.10) holds.

To prove (2.11) we define

$$b := \langle \Sigma(A) \rangle^{-1}(|X'| + |S| \langle D \rangle^{-1} |U|),$$
$$a := \langle D \rangle^{-1}(|U| + |Z|^T b).$$

Then (2.12) implies

$$\langle A \rangle \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} |U| \\ |X'| \end{pmatrix} = |X|, \tag{2.13}$$

and by Lemma 2.4 and [11, Theorem 4], we have

$$\begin{aligned} |Y| &\leqslant |\Sigma(A)^G|\, |X' - SD^{-1}U| \\ &\leqslant \langle \Sigma(A) \rangle^{-1}(|X'| + |S| \langle D \rangle^{-1} |U|) = b. \end{aligned}$$

Now, by the above argument for $n=1$, we have

$$\begin{aligned}(U-Z^TY)/D = D^G(U-Z^TY) &\subseteq U - Z^TY + [-1,1](\langle D\rangle^{-1}-1)|U-Z^TY|\\ &= U+[-1,1]|Z|^T|Y| + [-1,1](\langle D\rangle^{-1}-1)(|U|+|Z|^T|Y|)\\ &\subseteq U+[-1,1](a-|U|),\end{aligned}$$

and by the induction hypothesis we have

$$\begin{aligned}Y &= \Sigma(A)^G(X'-SD^{-1}U)\\ &\subseteq X'-SD^{-1}U + E'(\langle\Sigma(A)\rangle^{-1}-e')|X'-SD^{-1}U|\\ &= X'+E'|S|\langle D\rangle^{-1}|U| + E'(\langle\Sigma(A)\rangle^{-1}-e')(|X'|+|S|\langle D\rangle^{-1}|U|)\\ &= X'+E'(b-|X'|).\end{aligned}$$

Therefore, by Lemma 2.4 and (2.13)

$$\begin{aligned}A^GX &= \begin{pmatrix}(U-Z^TY)/D\\ Y\end{pmatrix} \subseteq \begin{pmatrix}U+[-1,1](a-|U|)\\ X'+E'(b-|X'|)\end{pmatrix}\\ &= \begin{pmatrix}U\\ X'\end{pmatrix} + E\left(\begin{pmatrix}a\\ b\end{pmatrix} - \begin{pmatrix}|U|\\ |X'|\end{pmatrix}\right)\\ &= X + E(\langle A\rangle^{-1}|X|-|X|)\end{aligned}$$

which implies (2.11). ∎

COROLLARY 2.6. *Let $A \in \mathbb{IR}^{n\times n}$ be an H-matrix with* mid $A=e$. *Then*

$$\begin{aligned}\kappa(A^G) &\subseteq [2e-\langle A\rangle^{-1}, \langle A\rangle^{-1}],\\ |\kappa(A^G)| &= |A^G| = \langle A\rangle^{-1}.\end{aligned}$$

Proof. By (2.11), we have

$$\kappa(A^G) = A^Ge \subseteq e + E(\langle A\rangle^{-1}-e)|e| = [2e-\langle A\rangle^{-1}, \langle A\rangle^{-1}].$$

Moreover, using Theorem 4 of [11] and (2.10), we find

$$\begin{aligned}|\kappa(A^G)| \leqslant |A^G| &= \langle R\rangle^{-1}\langle L\rangle^{-1} = \langle A\rangle^{-1}\\ &= \underline{A}^{-1} = |\kappa(\underline{A}^G)| \leqslant |\kappa(A^G)|,\end{aligned}$$

whence $|\kappa(A^G)| = |A^G| = \langle A\rangle^{-1}$. ∎

3. A Function Strip and Its Zero Set

Let G: $D \in \mathbb{IR}^n \to \mathbb{IR}^n$ be a map which associates with each $x\in D$ an interval

$$G(x) := [\underline{g}(x), \bar{g}(x)]. \tag{3.1}$$

Such a map is called a *function strip.*

We assume that G satisfies on D an interval Lipschitz condition, i.e., the real functions $\underline{g}$ and $\bar{g}$ both satisfy the same interval Lipschitz condition

$$\begin{gathered}\underline{g}(x_1)-\underline{g}(x_2) \in A(x_1-x_2), \qquad \bar{g}(x_1)-\bar{g}(x_2)\in A(x_1-x_2)\\ \text{for all } x_1, x_2 \in D,\end{gathered} \tag{3.2}$$

where the interval Lipschitz matrix A is regular. We call

$$X^* := \{x\in D \mid \underline{g}(x) \leqslant 0 \leqslant \bar{g}(x)\} \tag{3.3}$$

the *zero set* of G (which can be empty).

In the following we shall need several properties of the functions

$$g(x,d) := d\underline{g}(x) + (e-d)\,\bar{g}(x), \tag{3.4}$$

where $d \in [0,e]$ is a nonnegative diagonal matrix with diagonal entries in $[0, 1]$.

PROPOSITION 3.1. (i) *For all $x\in D$ and all diagonal matrices $d\in[0,e]$,*

$$g(x,d)\in G(x).$$

(ii) *For all $x\in D$ and $\tilde{g}\in G(x)$ there is a diagonal matrix $d\in[0,e]$ such that $\tilde{g}=g(x,d)$.*

(iii) *For all x_1, $x_2\in D$ and diagonal matrices $d\in[0,e]$ there is a matrix $a\in A$ with*

$$g(x_1,d)-g(x_2,d) = a(x_1-x_2). \tag{3.5}$$

(iv) *For each diagonal matrix $d\in[0,e]$, $g(x,d)$ has at most one zero $x_d^*\in D$.*

(v) *$x^*\in D$ belongs to the zero set X^* of G iff*

$$g(x^*,d)=0 \qquad \text{for some } d\in[0,e].$$

Proof. (i)

$$\begin{aligned} \underline{g}(x) &= d\underline{g}(x) + (e-d)\,\underline{g}(x) \\ &\leqslant d\underline{g}(x) + (e-d)\,\bar{g}(x) = g(x,d) \qquad \text{(by (3.4))} \\ &\leqslant d\bar{g}(x) + (e-d)\,\bar{g}(x) = \bar{g}(x). \end{aligned}$$

(ii) Let $\tilde{g} \in G(x)$. Then $\underline{g}(x) - \tilde{g} \leqslant 0$ and $\bar{g}(x) - \tilde{g} \geqslant 0$; therefore a diagonal matrix $d \in [0, e]$ exists with $d(\underline{g}(x) - \tilde{g}) + (e-d)(\bar{g}(x) - \tilde{g}) = 0 = d\underline{g}(x) + (e-d)\,\bar{g}(x) - \tilde{g}$. Hence by (3.4) we have $\tilde{g} = g(x, d)$.

(iii) $g(x_1, d) - g(x_2, d) = d(\underline{g}(x_1) - \underline{g}(x_2)) + (e-d)(\bar{g}(x_1) - \bar{g}(x_2)) \in d\ A(x_1 - x_2) + (e-d)\ A(x_1 - x_2) = A(x_1 - x_2)$, and by Lemma 3 in [13] the relation $g(x_1, d) - g(x_2, d) \in A(x_1 - x_2)$ is equivalent to (3.5).

(iv) Suppose x_1^*, x_2^* are two zeros of $g(x, d)$. Then (3.5) implies that $0 = g(x_1^*, d) - g(x_2^*, d) = a(x_1^* - x_2^*)$ with $a \in A$. Since a is regular we have $x_1^* = x_2^*$.

(v) Let $x^* \in X^*$. Then by (3.3), $\underline{g}(x^*) \leqslant 0$ and $\bar{g}(x^*) \geqslant 0$, hence there exists a diagonal matrix $d \in [0, e]$ with $d\underline{g}(x^*) + (e-d)\,\bar{g}(x^*) = 0$. By (3.4) we obtain $g(x^*, d) = 0$. Conversely, if $g(x^*, d) = 0$ then (i) implies $0 \in G(x^*)$ whence $x^* \in X^*$. ∎

PROPOSITION 3.2. *If x_1, $x_2 \in D$ then*

$$G(x_1) \subseteq A(x_1 - x_2) + G(x_2). \tag{3.6}$$

Proof. Let $\tilde{g} \in G(x_1)$. Then by Proposition (3.1)(i), (ii), and (iii), $\tilde{g} = g(x_1, d) = a(x_1 - x_2) + g(x_2, d) \in A(x_1 - x_2) + G(x_2)$. ∎

4. ENCLOSING THE ZERO SET OF A FUNCTION STRIP

In this section we find enclosures for the zero set X^* of a function strip $G\colon D \subseteq \mathbb{R}^n \to \mathbb{IR}^n$ satisfying a Lipschitz condition (3.2) as discussed in Section 3. Moreover, we discuss the quality of these enclosures.

LEMMA 4.1. *Suppose that A^H defined by (2.7) is regular and*

$$\tilde{X} := \tilde{x} - A^H G(\tilde{x}) \subseteq D \qquad \textit{for some } \tilde{x} \in D. \tag{4.1}$$

Then for every diagonal matrix $d \in [0, e]$, the function $g(x, d)$ defined by (3.4) has a unique zero $x_d^* \in D$, and

$$X^* = \{x_d^* \mid d \in [0, e]\}. \tag{4.2}$$

Proof. In view of Proposition 3.1(i) and (S1) we have for all $d \in [0, e]$ the inclusion

$$\tilde{x} - A^H g(\tilde{x}, d) \subseteq \tilde{X},$$

so that by the theorem in [14], $g(x, d)$ has a zero $x_d^* \in X$. By Proposition 3.1(iv) and (v), x_d^* is the unique zero of $g(x, d)$ in D and $X^* = \{x_d^* \mid d \in [0, e]\}$. ∎

THEOREM 4.2. *Let $\tilde{x} \in D$ and $\tilde{X} := \tilde{x} - A^H G(\tilde{x}) \subseteq D$. Then*

$$X^* \subseteq \{\tilde{x} - \tilde{a}^{-1}\tilde{g} \mid \tilde{a} \in A,\ \tilde{g} \in G(\tilde{x})\} \subseteq \tilde{X}. \tag{4.3}$$

Moreover, if A^H is regular and $\tilde{X} \subseteq D$ then

$$0 \leqslant \operatorname{rad} \tilde{X} - \operatorname{rad} \square X^* \leqslant 2 \operatorname{rad}(A^{-1})\,|G(\tilde{x})|; \tag{4.4}$$

here $A^{-1} = \square\{\tilde{a}^{-1} \mid \tilde{a} \in A\}$.

Proof. By Proposition 3.1(v), each $x^* \in X^*$ satisfies $g(x^*, d) = 0$ for some $d \in [0, e]$. Hence by Proposition 3.1(iii), $g(\tilde{x}, d) = g(\tilde{x}, d) - g(x^*, d) = a(\tilde{x} - x^*)$ for some $a \in A$, so that $x^* = \tilde{x} - a^{-1} g(\tilde{x}, d) \in \tilde{X}$. This implies (4.3). To prove (4.4) we define

$$\Sigma := \{\tilde{x} - \tilde{a}^{-1}\tilde{g} \mid \tilde{a} \in A,\ \tilde{g} \in G(\tilde{x})\}.$$

Then $X^* \subseteq \Sigma$ and $\tilde{X} = \square\Sigma$ by definition of A^H (see (2.7)). Let $z \in \Sigma$. Then $z = \tilde{x} - \tilde{a}^{-1}\tilde{g}$ with $\tilde{a} \in A$, $\tilde{g} \in G(\tilde{x})$, and by Proposition 3.1(ii), $\tilde{g} = g(\tilde{x}, d)$ for some $d \in [0, e]$. With the zero x_d^* of $g(x, d)$ guaranteed by Lemma 4.1, Proposition 3.1(iii) implies $x_d^* = \tilde{x} - a^{-1}\tilde{g}$, and

$$\begin{aligned} |z - x_d^*| &= (\tilde{a}^{-1} - a^{-1})\,\tilde{g}| \leqslant |\tilde{a}^{-1} - a^{-1}|\,|\tilde{g}| \\ &\leqslant 2 \operatorname{rad}(A^{-1})\,|G(\tilde{x})|. \end{aligned}$$

Since $X^* \subseteq \Sigma$, the definiton (2.1) of the distance and its properties (see Lemma 1 in [13]) imply $\operatorname{rad} \tilde{X} - \operatorname{rad} \square X^* \leqslant q(\tilde{X}, \square X^*) = q(\square\Sigma, \square X^*) \leqslant 2 \operatorname{rad}(A^{-1})|G(\tilde{x})|$. This gives the upper bound in (4.4), and the lower bound follows from $\square X^* \subseteq \tilde{X}$. ∎

In practice it is difficult to compute $A^H G(\tilde{x})$, and one uses instead the coarser enclosure by $A^I G(\tilde{x})$, where A^I is a computable inverse of A (2.6). In this case the bound (4.4) becomes slightly bigger:

THEOREM 4.3. *Let A^I be a regular inverse of A, and suppose that $\tilde{x} \in D$ is such that*

$$X^I := \tilde{x} - A^I G(\tilde{x}) \subseteq D.$$

If A^H is regular then, for every $a \in A$, we have

$$0 \leqslant \operatorname{rad} X^I - \operatorname{rad} \square X^*$$
$$\leqslant (2 \operatorname{rad}(A^{-1}) + \sup\{\operatorname{rad} \kappa(A^I), |A^I| - |a^{-1}|\}) \, |G(\tilde{x})|.$$

Proof. In view of [11], Theorem 1, we have

$$\begin{aligned}\operatorname{rad} X^I &= \operatorname{rad}(A^I G(\tilde{x})) \\ &\leqslant \operatorname{rad}(\kappa(A^I) \operatorname{mid} G(\tilde{x}) + |A^I|(E \operatorname{rad} G(\tilde{x}))) \\ &= \operatorname{rad}(\kappa(A^I)) |\operatorname{mid} G(\tilde{x})| + |A^I| \operatorname{rad} G(\tilde{x}).\end{aligned}$$

Now for every $a \in A$, $\operatorname{rad} \tilde{X} = \operatorname{rad}(A^H G(\tilde{x})) \geqslant \operatorname{rad}(a^{-1} G(\tilde{x})) \geqslant |a^{-1}| \operatorname{rad} G(\tilde{x})$, since $A^H G(\tilde{x}) \supseteq a^H G(\tilde{x}) = a^{-1} G(\tilde{x})$, and we get

$$\begin{aligned}\operatorname{rad} X^I - \operatorname{rad} \tilde{X} &\leqslant \operatorname{rad}(\kappa(A^I)) |\operatorname{mid} G(\tilde{x})| + (|A^I| - |a^{-1}|) \operatorname{rad} G(\tilde{x}) \\ &\leqslant \sup\{\operatorname{rad} \kappa(A^I), |A^I| - |a^{-1}|\} |G(\tilde{x})|,\end{aligned}$$

since $|a^{-1}| \leqslant |A^I|$ and $|\operatorname{mid} G(\tilde{x})| + \operatorname{rad} G(\tilde{x}) = |G(\tilde{x})|$. Now apply Theorem 4.2. ∎

By the above theorem, the quality of the enclosure is good if:

(i) $|G(\tilde{x})|$ is small. This holds, e.g., if G is a narrow function strip ($\operatorname{rad} G(x)$ small) and $\tilde{x} \in X^*$, since then $0 \in G(\tilde{x})$, $|G(\tilde{x})| \leqslant 2 \operatorname{rad} G(\tilde{x})$.

(ii) $\operatorname{rad}(A^{-1})$ and $|A^I| - |a^{-1}|$ are small. Since A usually becomes larger with D, a small D improves the quality. Indeed, quite often $\operatorname{rad} A^{-1} = O(|\check{A}^{-1}| \operatorname{rad} D)$. A good inverse A^I which encloses the hull with little overestimation also improves the quality.

Remark. One might hope to get

$$\operatorname{rad} \tilde{X} \geqslant |A^{-1}| \operatorname{rad} G(\tilde{x}) \tag{4.5}$$

so that in the theorem, a^{-1} can be replaced by A^{-1}. However, (4.5) only holds if A^H is normal, and the example

$$A = \begin{pmatrix} [-1, 7] & -1 \\ 3 & 1 \end{pmatrix}, \qquad X = \begin{pmatrix} [-1, 1] \\ [-1, 1] \end{pmatrix}, \qquad A^H X = \begin{pmatrix} [-1, 1] \\ [-2, 2] \end{pmatrix},$$
$$\operatorname{rad}(A^H X) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \not\geqslant |A^H| \operatorname{rad} X = \begin{pmatrix} 0.5 & 0.5 \\ 1.5 & 0.7 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2.2 \end{pmatrix}$$

shows that A^H need not be normal.

Nonetheless, the bound in Theorem 4.3 is not optimal since for $A^I = A^H$ it does not reduce to that of Theorem 4.2.

5. Interval Newton Operators and Their Fixed Intervals

Let G be a function strip defined by (3.1) which satisfies the Lipschitz condition (3.2) with a regular matrix A, and let $A^I \colon \mathbb{IR}^n \to \mathbb{IR}^n$ be an inverse of A (see (2.6)). Then we call the operator N defined by

$$N(X) := \check{x} - A^I G(\check{x}), \qquad \text{where } \check{x} = \operatorname{mid} X, \tag{5.1}$$

an *interval Newton operator* for G. An interval $\hat{X} := \hat{x} + E \operatorname{rad} \hat{X}$ is called *fixed interval* of N if

$$\hat{X} = N(\hat{X}) = \hat{x} - A^I G(\hat{x}). \tag{5.2}$$

If A^I is a centered inverse, i.e., if (S5) is fulfilled, then (5.2) implies

$$\operatorname{mid} G(\hat{x}) = 0, \tag{5.3}$$

and by (3.8) of Neumaier [10],

$$\operatorname{rad} \hat{X} = \operatorname{rad} N(\hat{X}) = \operatorname{rad}(A^I G(\hat{x})) \leqslant |A^I| \operatorname{rad} G(\hat{x}). \tag{5.4}$$

Therefore, if the centered inverse A^I is also normal, i.e., if A^I fulfills in addition (S4), then

$$\operatorname{rad} \hat{X} = |A^I| \operatorname{rad} G(\hat{x}). \tag{5.5}$$

Let an interval operator N be defined by (5.1), where A^I is a centered inverse of A. In order to improve the radius of a fixed interval $\hat{X}$ of N we may use the inverse $A^{\tilde{I}}$ defined in Theorem 2. instead of A^I in (5.1). Then by Theorem 2.3(ii), (iii), and (5.5) we obtain for a corresponding fixed interval $\hat{X}'$ the relation

$$\operatorname{rad} \hat{X}' = |\kappa(A^I)| \operatorname{rad} G(\hat{x}) \leqslant \operatorname{rad} \hat{X}. \tag{5.6}$$

Now we consider the question how good the enclosure of X^* by a fixed interval $\hat{X}$ of N can be. The answer is given by

THEOREM 5.1. *Let N be defined by (5.1) with an inverse A^I of A, and let X^* denote the zero set of G (see (3.3)). Then*

(i) *$X^* \subseteq X$ implies $X^* \subseteq N(X)$.* (5.7)

(ii) *If a fixed interval $\hat{X}$ of N exists, then*

$$X^* \subseteq \hat{X}. \tag{5.8}$$

(iii) *If, in addition, A^I is regular and centered, then for every $a \in A$,*

$$0 \leqslant \operatorname{rad} \hat{X} - \operatorname{rad} \square X^* \leqslant (2 \operatorname{rad}(A^{-1}) + |A^I| - |a^{-1}|) \operatorname{rad} G(\hat{x}). \tag{5.9}$$

(iv) *Moreover, if for some* $a \in A$,

$$|A^I| \leqslant |a^{-1}| + 2\,\mathrm{rad}(\kappa(A^I)), \tag{5.10}$$

then

$$0 \leqslant \mathrm{rad}\,\hat{X} - \mathrm{rad}\,\square X^* \leqslant 2(\mathrm{rad}(A^{-1}) + \mathrm{rad}(\kappa(A^I)))\,\mathrm{rad}\,G(\check{x}). \tag{5.11}$$

Proof. (i) follows from Theorem 4.2(i) with $\check{x} = \tilde{x}$ and $\tilde{X} \subseteq \check{x} - A^I G(\check{x}) = N(X)$.

(ii) is a consequence of $N(\hat{X}) = \hat{X}$.

(iii) follows as in Theorem 4.3 since mid $G(\check{x}) = 0$ by (5.3).

(iv) follows directly from (5.9) and (5.10). ∎

Remarks. (1) If $B \supseteq A^{-1}$ and $A^I = B^M$ (B^M means matrix multiplication, see Example 1 in Sect. 3 of [11]) then the assumption (5.10) is satisfied by Lemma 2.1.

(2) If $A^{\hat{I}}$ is defined as in Theorem 2.2 then $|A^{\hat{I}}|$ fulfills (5.10) because of Lemma 2.1.

(3) If rad $G(\check{x})$, $\mathrm{rad}(A^{-1})$, and rad $\kappa(A^I)$ are of order $O(\varepsilon)$ then rad $\hat{X}$ − rad $\square X^*$ in (5.11) is of order $O(\varepsilon^2)$. This means quadratical convergence, if rad $G(\check{x}) \to 0$.

(4) In [9] the interval operator N was called *inclusion preserving* if (5.7) holds, and *normal* if (5.8) holds.

6. A Convergence Theorem for Interval Iteration

In this section we use an interval Newton operator N to determine an interval sequence $\{X_k\}$, and we consider the question under which assumptions this sequence converges to a fixed interval of N. For this purpose we use the following propositions.

Proposition 6.1. *Let* $X, Y \subseteq D$ *and* N *be defined by* (5.1). *Then*

$$|\mathrm{rad}\,N(X) - \mathrm{rad}\,N(Y)| \leqslant \mathrm{rad}(A^I A)|\check{x} - \check{y}|. \tag{6.1}$$

Proof. By Proposition 3.2 and Proposition 1 in [12], we have

$$\begin{aligned} \mathrm{rad}\,N(X) = \mathrm{rad}(A^I G(\check{x})) &\leqslant \mathrm{rad}(A^I(A(\check{x} - \check{y}) + G(\check{y}))) \\ &\leqslant \mathrm{rad}(A^I A)|\check{x} - \check{y}| + \mathrm{rad}(A^I G(\check{y})) \\ &= \mathrm{rad}(A^I A)|\check{x} - \check{y}| + \mathrm{rad}\,N(Y), \end{aligned}$$

hence rad $N(X)$ − rad $N(Y) \leqslant \mathrm{rad}(A^I A)|\check{x} - \check{y}|$. By interchanging X and Y we also get rad $N(Y)$ − rad $N(X) \leqslant \mathrm{rad}(A^I A)|\check{x} - \check{y}|$, which implies (6.1). ∎

Proposition 6.2. *Let* N *be defined by* (5.1). *If there is a fixed interval* $\hat{X}$ *of* N *and*

$$\sigma(\mathrm{rad}(A^I A)) < 1, \tag{6.2}$$

then

$$N(X) \supseteq X \supseteq \hat{X} \quad \textit{implies} \quad X = \hat{X}. \tag{6.3}$$

Proof. Let $N(X) \supseteq X \supseteq \hat{X}$. Then it follows by (2.2) and Proposition 6.1 that

$$\begin{aligned} \mathrm{rad}\,X - \mathrm{rad}\,\hat{X} \leqslant \mathrm{rad}\,N(X) - \mathrm{rad}\,\hat{X} &\leqslant \mathrm{rad}(A^I A)|\check{x} - \hat{x}| \\ &\leqslant \mathrm{rad}(A^I A)(\mathrm{rad}\,X - \mathrm{rad}\,\hat{X}), \end{aligned}$$

hence

$$(e - \mathrm{rad}(A^I A))(\mathrm{rad}\,X - \mathrm{rad}\,\hat{X}) \leqslant 0.$$

Since $(e - \mathrm{rad}(A^I A))^{-1} \geqslant 0$ because of (6.2), this implies

$$\mathrm{rad}\,X - \mathrm{rad}\,\hat{X} \leqslant 0 \text{ and therefore } X = \hat{X}. \; ∎$$

An interval vector X_0 is called *stable* with respect to N, if

$$X_0 \subseteq X \quad \text{implies} \quad X_0 \subseteq N(X). \tag{6.4}$$

Note, that by (5.7), $\square X^*$ is stable.

Now we will improve the enclosure $X \supseteq X^*$ iteratively.

Theorem 6.3. *Let* $G\colon D \subseteq \mathbb{R}^n \to \mathbb{IR}^n$ *be a function strip satisfying the Lipschitz condition* (3.2) *with regular* A, *and let* X^* *be the zero set of* G *defined by* (3.3). *Let* $N(X)$ *be defined by* (5.1) *with an inverse* A^I *of* A. *Suppose* $\{X_k\}_{k=0}^{\infty}$ *is an interval sequence with* $X_0 \in \mathbb{I}D$, $X_k \in \mathbb{I}D \cup \{\varnothing\}$ $(k = 1, 2, \ldots)$ *such that* $X_{k+1} = \varnothing$ *if* $X_k = \varnothing$, *and otherwise*

$$X_{k+1} := X_k \cap N(X_k). \tag{6.5}$$

Then

$$X_{k+1} \subseteq X_k \subseteq \cdots \subseteq X_0, \tag{6.6}$$

$X_\infty := \lim_{n\to\infty} X_k$ *exists, and the following assertions hold:*

(i) *If* $X_k = \varnothing$ *for some* $k > 0$ *then* $X_0 \cap X^* = \varnothing$.

(ii) *If* $X^* \neq \varnothing$ *and* $X_0 \supseteq X^*$ *then* $X_k \supseteq X^*$ *for all* $k \in N$,

$$X_\infty \supseteq X^* \qquad \text{and} \qquad N(X_\infty) \supseteq X_\infty.$$

Moreover, if there exists a stable fixed interval $\hat{X}$ *of* N *then:*

(iii) *If* $X_0 \supseteq \hat{X}$ *then*

$$\operatorname{rad} X_{k+1} - \operatorname{rad} \hat{X} \leqslant \operatorname{rad}(A^I A)(\operatorname{rad} X_k - \operatorname{rad} \hat{X}),$$

(iv) *If* $X_0 \supseteq \hat{X}$ *and* $\sigma(\operatorname{rad}(A^I A)) < 1$ *then*

$$X_\infty = \hat{X}.$$

Proof. (6.6) is obvious. (i) as well as (ii) are consequences of Theorem 5.1(i), and $N(X_\infty) \supseteq X_\infty$ holds because of the continuity of A^I and N, respectively. Now assume that $\hat{X}$ is a stable fixed interval. By $X_0 \supseteq \hat{X}$ and (6.4) it follows that $X_k \supseteq \hat{X}$ for all $k \in N$, hence by (2.2) $|\operatorname{mid} X_k - \operatorname{mid} \hat{X}| \leqslant \operatorname{rad} X_k - \operatorname{rad} \hat{X}$. By applying Proposition 6.1 we obtain (iii). (iv) is a consequence of $N(X_\infty) \supseteq X_\infty \supseteq \hat{X}$ and Proposition 6.2. ∎

We call $\operatorname{rad}(A^I A)$ a *convergence matrix* of N with respect to A^I and $\sigma(\operatorname{rad}(A^I A))$ the associated *convergence factor.*

Remark. In [9] an interval operator N was called *strong*, if the property (6.3) holds, and *fixed interval preserving*, if the condition (6.4) is fulfilled.

7. Particular Inverses and Newton Operators

In this section we consider several Newton operators associated with specific inverses and compare them with respect to inclusion, absolute value, and convergence matrix.

Example 1. Let $A^I := A^G$ be the Gauss inverse of a regular matrix A possessing a triangular decomposition (L, R). The iteration with the corresponding Newton operator

$$N_G(X) := \check{x} - A^G G(\check{x})$$

was considered for real functions $G(x)$ by Alefeld [4] who uses the notation $IGA(A, Z)$ for $A^G Z$.

Properties of the Inverse A^G.

(i) A^G *is normal and centered.*

(ii) $|A^G| \operatorname{rad} A \leqslant \operatorname{rad}(A^G A) \leqslant \operatorname{rad}(\kappa(A^G))|\check{A}| + |A^G| \operatorname{rad} A.$

(iii) $|\kappa(A^G)| \leqslant |A^G| = \langle R \rangle^{-1} \langle L \rangle^{-1}.$

Proof. (i) follows from Theorem 3 of [11] and Proposition 2.3. (ii) is an application of (3.8) of [11] and (S4), and (iii) follows from (3.7) and (5.13) of [11]. ∎

By Theorem 2.2 we can improve A^G by considering instead the inverse $A^{\bar{G}}$; we then have $|A^{\bar{G}}| \leqslant |A^G|$. However, we do not know how to compare the corresponding convergence matrices $\operatorname{rad}(A^{\bar{G}} A)$ and $\operatorname{rad}(A^G A)$. The argument leading to (ii) only gives

$$|A^{\bar{G}}| \operatorname{rad} A \leqslant \operatorname{rad}(A^{\bar{G}} A) \leqslant \operatorname{rad}(\kappa(A^{\bar{G}}))|\check{A}| + |A^{\bar{G}}| \operatorname{rad} A,$$

so that both the lower and upper bound for $\operatorname{rad}(A^{\bar{G}} A)$ are at least as good as the corresponding bounds for $\operatorname{rad}(A^G A)$.

Unless $n = 2$ or A has special properties the interval Gauss algorithm very likely breaks down or yields a pessimistic inclusion for $A^H G(x)$. The usually recommended remedy is preconditioning with the midpoint inverse; see [11] and the citations there. For the remaining examples we therefore assume the following situation:

$$\begin{aligned} &A \in \mathbb{IR}^{n\times n}, && c := (\operatorname{mid} A)^{-1} \quad \text{exists}, \\ &r := |c| \operatorname{rad} A && \text{satisfies } \sigma(r) < 1, \\ &q := (e - r)^{-1} r \ (\geqslant 0). \end{aligned} \tag{7.1}$$

It is well known that under this assumption A is regular. We shall call A *strongly regular* if (7.1) holds.

Example 2. If A is strongly regular then by Eq. (4.4) in [12], the map A^K defined by

$$A^K Z := cZ + (qE)(cZ) \qquad \text{for} \quad Z \in \mathbb{IR}^n \tag{7.2}$$

is an inverse of A. The corresponding Newton operator is

$$N_K(X) := \check{x} - cG(\check{x}) + (qE)(cG(\check{x}))$$

and was introduced in [8].

Properties of the Inverse A^K.

(i) A^K *is normal and centered.*

(ii) $\operatorname{rad}(A^K A) = 2q.$

(iii) $|\kappa(A^K)| = |A^K| = (e + q)|c| = (e - r)^{-1}|c|.$

(iv) $\kappa(A^K) = [c - q|c|, c + q|c|].$

(v) $|A^K| \leqslant |a^{-1}| + 2 \operatorname{rad}(\kappa(A^K))$ *for all* $a \in A$.

(vi) *If* $\sigma(q) < 1$ *then* A^K *is regular.*

Proof. (2.4) and (7.2) imply

$$\kappa(A^K) = A^K e = c + (qE)\,c = [c - q|c|,\ c + q|c|];$$

hence (iv) holds. Since

$$|A^K| = |A^K E| = |c| + q|c| = (e+q)|c| = (e-r)^{-1}|c|$$

and

$$|\kappa(A^K)| = \sup(-(c - q|c|),\ c + q|c|)$$
$$= \sup(-c, c) + q|c| = |c| + q|c|,$$

(iii) holds as well. Now axiom (S4) holds since

$$\operatorname{rad}(A^K Z) = \operatorname{rad}(cZ) + q|cZ| \geqslant \operatorname{rad}(cZ) + q\operatorname{rad}(cZ)$$
$$= (e+q)\operatorname{rad}(cZ) = (e+q)|c|\operatorname{rad} Z = |A^K|\operatorname{rad} Z,$$

and (S5) holds since by regularity of c, $0 = \operatorname{mid}(A^K Z) = \operatorname{mid}(cZ) = c\operatorname{mid} Z$ implies $\operatorname{mid} Z = 0$. Hence (i) holds. Since $cA = e + [-r, r]$ (see (4.1) in [8]), (ii) follows from

$$\operatorname{rad}(A^K A) = \operatorname{rad}(cA + (qE)(cA)) = r + q(e+r) = 2q.$$

Now (ii) and (iv) imply $|A^K| = (e-r)|c| = q|c| + |c| \leqslant q|c| + |c - a^{-1}| + |a^{-1}| \leqslant 2q|c| + |a^{-1}|$, giving (v). Finally, if $\sigma(q) < 1$ then (see [11]) the matrix $[e-q, e+q]$ is an H-matrix, hence regular, so that $x \in \mathbb{R}^n$, $0 \in A^K x = cx + (qE)(cx) = [e-q, e+q]\,cx$ implies $cx = 0$ and therefore $x = 0$. Hence (S6) holds, and (vi) follows. ∎

By (5.5), a fixed interval $\hat{X}_K$ of N_K has radius

$$\operatorname{rad}\hat{X}_K = (e-r)^{-1}|c|\operatorname{rad} G(\hat{x}_K).$$

By Theorem 5.4 and 5.5 of [9], the condition $\sigma(q) < 1$ already implies that a fixed interval $\hat{X}_K$ is stable and that $\lim_{k\to\infty} X_k = \hat{X}_K$ for the iteration (6.5) with $X_0 \supseteq \hat{X}_K$. This is much better than we can deduce from Theorem 6.3(iv), where we have to *assume* that $\hat{X}_K$ is stable and $\sigma(2q) < 1$ to have convergence. This suggests that there might be a stronger version of Theorem 6.3.

The use of the inverse $A^{\bar{K}}$ (constructed in Theorem 2.2) in place of A^K gives no improvement since

$$A^K Z = c\check{z} + |c|\,E\operatorname{rad} Z + (qE)(c\check{z} + |c|\,E\operatorname{rad} Z)$$
$$= c\check{z} + qE(c\check{z}) + (e+q)|c|\,E\operatorname{rad} Z$$
$$= A^K\check{z} + |\kappa(A^K)|(E\operatorname{rad} Z) = A^{\bar{K}} Z$$

for all $Z \in \mathbb{IR}^n$.

EXAMPLE 3. If A is strongly regular then by Eq. (4.18) in [10], the map A^V defined by

$$A^V Z := [e-q, e+q](cZ) \quad \text{for} \quad Z \in \mathbb{IR}^n \tag{7.3}$$

is an inverse of A; by the subdistributive law we have

$$A^V Z \subseteq A^K Z, \qquad \text{with equality if} \quad \operatorname{rad} Z = 0. \tag{7.4}$$

The corresponding Newton operator is

$$N_V(X) := \check{x} - [e-q, e+q](cG(\check{x}))$$

and was introduced in [10] as an improvement over $N_K(X)$.

PROPERTIES OF THE INVERSE A^V.

- (i) *A^V is normal and centered.*
- (ii) $q \leqslant \operatorname{rad}(A^V A) \leqslant 2q$.
- (iii) $|\kappa(A^V)| = |A^V| = (e+q)|c| = (e-r)^{-1}|c|$.
- (iv) $\kappa(A^V) = [c - q|c|,\ c + q|c|]$.
- (v) $|A^V| \leqslant |a^{-1}| + 2\operatorname{rad}(\kappa(A^V))$ *for all* $a \in A$.
- (vi) *If* $\sigma(q) < 1$ *then* A^V *is regular.*

Proof. (ii), (iv), and (v) follow immediately from (7.4) and the corresponding properties of A^K. Now (iii) follows from

$$(e+q)|c| = |\kappa(A^V)| \leqslant |A^V| \leqslant |A^K| = (e+q)|c|$$

and the lower bound in (ii) follows from (i), (iii), and (71) since $\operatorname{rad}(A^V A) \geqslant |A^V|\operatorname{rad} A = (e-r)^{-1}|c|\operatorname{rad} A = q$. It remains to prove (i). But A^V is normal by Section 4 of [11] and centered by Proposition 2.2 of [10]. ∎

We remark that $A^{P}Z = A^{\bar{K}} Z = A^K Z \supseteq A^V Z$ since $A^V\check{z} = A^K\check{z}$ and $\kappa(A^V) = \kappa(A^K)$; thus the construction of Theorem 2.2 does not improve the inverse.

EXAMPLE 4. If A is strongly regular then

$$cA = [e-r, e+r] \tag{7.5}$$

is an H-matrix, and by [2], the triangular decomposition of cA and the Gauss inverse $(cA)^G$ exist. Therefore the map A^P defined by

$$A^P Z := (cA)^G(cZ) \qquad \text{for} \quad Z \in \mathbb{IR}^n \tag{7.6}$$

is defined and is an inverse of A, the *preconditioned Gauss inverse.* By Proposition 2.5 we have

$$A^P Z = (cA)^G (cZ) \subseteq cZ + E((e-r)^{-1} - e)|cZ|$$
$$= cZ + Eq|cZ| = cZ + (qE)(cZ)$$

so that

$$A^P Z \subseteq A^K Z \qquad \text{for all } Z \in \mathbb{IR}^n. \tag{7.7}$$

However, in contrast to (7.4), already simple examples with $n=1$ show that equality need not hold for rad $Z=0$. It is easy to see that $A^P Z \subseteq A^V Z$ for $n=1$; it would be interesting to know whether this also holds for higher dimensions.

PROPERTIES OF THE INVERSE A^P.

(i) *A^P is normal and centered.*

(ii) $q \leqslant \operatorname{rad}(A^P A) \leqslant 2q$.

(iii) $|\kappa(A^P)| \leqslant |A^P| = (e+q)|c| = (e-r)^{-1}|c|$.

(iv) $\kappa(A^P) \subseteq [c - q|c|,\ c + q|c|]$.

(v) *If $\sigma(q) < 1$ then A^P is regular.*

Proof. By Theorem 3 and Proposition 2 of [11], A^P is normal and by Proposition 2.3, A^P is centered since c is nonsingular. Hence (i) holds. The lower bound in (ii) holds since by (S4), Corollary 2.6 and 7.5 we have

$$\operatorname{rad}(A^P A) = \operatorname{rad}((cA)^G (cA)) \geqslant |(cA)^G| \operatorname{rad}(cA) = (e-r)^{-1} r = q.$$

The upper bound in (ii), and (iv), (v) follow immediately from (7.7) and the properties of A^K. Finally (iii) holds since

$$|A^P| = |A^P E| = |(cA)^G (cE)|$$
$$= \langle cA \rangle^{-1} |c| = (e-r)^{-1}|c| = (e+q)|c|$$

by rule (R8) of [11]. ∎

Remarks. 1. The hypothesis (5.10) in Theorem 5.1(iv) holds for A^K and A^V, so that (5.11) is valid for these inverses. We did not manage to decide whether (5.10) also holds for A^G and A^P.

2. Since $|A^P| = |A^V| = |A^K| = (e-r)^{-1}|c|$, the discussion in Section 5, in particular equation (5.5) implies that the fixed intervals of N_P coincide with those of N_V and N_K. However, the convergence matrix of N_P satisfies $\operatorname{rad}(A^P A) \leqslant 2q = \operatorname{rad}(A^K A)$ whence N_P is preferable to N_K. The above results do not allow a comparison of the convergence matrices of N_P and N_V. Also, it is not clear whether the use of the inverse A^P constructed in Theorem 2.2 is advisable. We have $A^P Z \subseteq A^R Z = A^K Z$ so that A^P is preferable to A^K; however, the real question is whether it improves on A^P.

REFERENCES

1. E. ADAMS, On sets of solutions of collections of nonlinear systems in $\mathbb{IR}^n$, *in* "Interval Mathematics 1980" (K. Nickel, Ed.), Academic Press, New York/London, 1980, pp. 247–256.
2. G. ALEFELD, Über die Durchführbarkeit des Gaußschen Algorithmus bei Gleichungen mit Intervallen als Koeffizienten, *Computing Suppl.* 1 (1977), 15–19.
3. G. ALEFELD AND J. HERZBERGER, "Introduction to Interval Computations," Academic Press, New York/London, 1983.
4. G. ALEFELD, On the convergence of some interval arithmetic modifications of Newton's method, *SIAM J. Numer. Anal.* **21** (1984), 363–372.
5. J. GARLOFF AND R. KRAWCZYK, "Optimal inclusion of a solution set," *SIAM J. Numer. Anal.* **23** (1986), 217–226.
6. D. M. GAY, Computing perturbation bounds for nonlinear algebraic equations, *SIAM J. Num. Anal.* **20** (1983), 638–651.
7. R. KRAWCZYK, Centered forms and interval operators, *Computing* **34** (1985), 243–259.
8. R. KRAWCZYK, Interval iterations for including a set of solutions, *Computing* **32** (1984), 13–31.
9. R. KRAWCZYK, "Properties of interval operations," *Computing* **37** (1986), 227–245.
10. R. KRAWCZYK AND A. NEUMAIER, An improved interval Newton operator, *J. Math. Anal. Appl.* **118** (1986), 194–207.
11. A. NEUMAIER, New techniques for the analysis of linear interval equations, *Linear Algebra Appl.* **58** (1984), 273–325.
12. A. NEUMAIER, "Interval Iteration for Zeros of Systems of Equations," *BIT* **25** (1985), 256–273.
13. A. NEUMAIER, "Overestimation in Linear Interval Equations," *SIAM J. Numer. Anal.*, in press.
14. A. NEUMAIER, "Existence of Solutions of Piecewise Differentiables Systems of Equations," *Archiv. Math.* **47** (1986), 443–447.

Printed by the St. Catherine Press Ltd., Tempelhof 41, Bruges, Belgium