JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **73**, 1–23 (1980)

Order-Convergence and Iterative Interval Methods

RUDOLF KRAWCZYK AND FLEMMING SELSMARK

Institut für Mathematik, Technische Universitat Clausthal, 3392 Clausthal-Zellenfeld, Federal Republic of Germany

Submitted by W. F. Ames

In the first part of this paper we introduce order-convergence in partially ordered spaces having lattice properties. Lipschitz assumptions are made for an operator equation $Tx = \Theta$, and additional operators are then derived from the Lipschitz operators. We show how to solve the operator equation by means of these operators, using iterative methods which produce interval sequences, and we state some theorems on the inclusion and the existence of a solution of the equation as well as on the convergence of the interval sequences. In the second part of the paper we show how these theorems can be used to find the solution of a real equation, a nonlinear system of equations in $\mathbb{R}^n$ and an algebraic eigenvalue problem.

INTRODUCTION

A number of iterative methods which produce a sequence of intervals, interval vectors, interval functions, etc. are already known. These interval quantities enclose generally a solution of a given equation and contract under certain conditions to the solution. Similar methods, which use interval arithmetic, have lately been developed (see [16, 17, 22–28, 32, 35]).

The inclusion of a solution of an operator equation $x = Tx$ using an interval sequence was examined by Collatz and Schröder (see [12, 13, 41]). If, for instance, an operator T can be expressed as the sum $T^+ + T^-$ of an isotone (monotone nondecreasing) operator T^+ and an antitone (monotone nonincreasing) operator T^-, then the iteration scheme

$$\bar{x}_{k+1} = T^+\bar{x}_k + T^-\underline{x}_k\,,$$
$$\underline{x}_{k+1} = T^+\underline{x}_k + T^-\bar{x}_k$$

will produce such a sequence. From the assumption $\underline{x}_0 \leqslant \underline{x}_1 \leqslant \bar{x}_1 \leqslant \bar{x}_0$ the inclusion $\underline{x}_k \leqslant \underline{x}_{k+1} \leqslant \bar{x}_{k+1} \leqslant \bar{x}_k$ for $k = 1, 2, \ldots$ follows; i.e., every interval is included in its predecessor. Alefeld [2] has shown that in the case of a linear system of equations ($Tx = Ax + b$, with $x, b \in \mathbb{R}^n$ and a real $(n \times n)$ matrix A) the above-mentioned method is identical to the interval method

$$[x_{k+1}] = \underset{\cdot}{A}[x_k] + \underset{\cdot}{b}.$$

0022-247X/80/010001-23$02.00/0

($[x_k]$ denotes an interval vector, $\underset{\cdot}{A}$ the point matrix corresponding to A, and $\underset{\cdot}{b}$ the point vector corresponding to b, and the arithmetic operations are interpreted as interval operations in the sense of [5, 7].)

As the method is of the form $x_{k+1} = Tx_k$, the generated sequence converges only if T possesses a contracting property. Furthermore, the convergence is in general linear. Methods of higher order in metric and pseudometric spaces, as, e.g., the Newton method, have been examined by Schröder [40] and several other authors.

We therefore consider here iterative methods in partially ordered spaces which produce interval sequences. As a concept of convergence we introduce the order-convergence implied by a partial ordering. Instead of completeness a certain lattice property is assumed, namely, that to every upper (lower) bounded sequence there exists a supremum (infimum). Lipschitz operators replace the Fréchet derivatives required for the Newton method. According to the choice of these Lipschitz operators and under appropriate, easily verifiable assumptions, one obtains linear, superlinear, or quadratic convergence of the iterative methods.

I. Abstract Theorems

1. *Notation and Basic Concepts*

Let $H = (H, \leqslant)$ be a linear partially ordered set over the field of real numbers $\mathbb{R}$, and let Θ denote the null element of H. H^+ denotes the positive cone, i.e., $H^+ = \{x \in H \mid \Theta \leqslant x\}$. An element $x \in H^+$ is called positive. If two elements $\underline{x}, \bar{x} \in H$ satisfy the relation $\underline{x} \leqslant \bar{x}$, then the subset $\{x \in H \mid \underline{x} \leqslant x \leqslant \bar{x}\}$ is called an *interval* and is denoted by $[\underline{x}, \bar{x}]$ or, for simplicity, by $[x]$. $\mathbb{I}(H)$ denotes the set of all intervals over H. Furthermore we define the *radius*

$$\rho[x] = \tfrac{1}{2}(\bar{x} - \underline{x}) \in H^+ \tag{1.1}$$

and the *mean value*

$$\mu[x] = \tfrac{1}{2}(\bar{x} + \underline{x}) \tag{1.2}$$

of any interval $[x] \in \mathbb{I}(H)$.

Addition in H implies addition in $\mathbb{I}(H)$ by the rule

$$[\underline{x}, \bar{x}] + [\underline{y}, \bar{y}] = [\underline{x} + \underline{y}, \bar{x} + \bar{y}].^1 \tag{1.3}$$

[1] This type of addition does not in general correspond to the usual complex addition. However, in lattices both additions are identical.

Multiplication with a constant $c \in \mathbb{R}$ follows accordingly:

$$\begin{aligned} c[\underline{x}, \bar{x}] &= [c\underline{x}, c\bar{x}] \quad \text{if} \quad c \geqslant 0, \\ &= [c\bar{x}, c\underline{x}] \quad \text{if} \quad c < 0. \end{aligned} \tag{1.4}$$

$\mathbb{I}(H)$ is a quasilinear space with respect to the above definitions of addition and multiplication (see [30]).

Furthermore, from (1.1) and (1.3) follows

LEMMA 1. *If* $[x], [y] \in \mathbb{I}(H)$, *then* $\rho[[x] \pm [y]] = \rho[x] + \rho[y]$.

An interval $[\underline{x}, \bar{x}]$, with $\underline{x} = \bar{x}$, is called a *point interval*, which we also denote by x. An interval $[x]$ is called a *null interval* if $\Theta \in [x]$.

Let $\{x_k\}_{k=1}^{\infty}$ (hereafter denoted $\{x_k\}$) be an indefinite sequence with $x_k \in H \ \forall\, k \in \mathbb{N}$, where $\mathbb{N}$ is the set of natural numbers. The sequence $\{x_k\}$ is called *upper bounded* (*lower bounded*) if an element $y \in H$ exists such that $x_k \leqslant y$ $(x_k \geqslant y)$ $\forall\, k \in \mathbb{N}$.

Furthermore, let H be a conditionally complete σ-lattice; i.e., a lattice with the following property: To every upper (lower) bounded sequence $\{x_k\}$, a unique element $\bar{x} \in H$ $(\underline{x} \in H)$ exists which satisfies the conditions

$$x_k \leqslant \bar{x} \ (x_k \geqslant \underline{x}) \qquad \forall\, k \in \mathbb{N}, \tag{1.5}$$

$$y \in H \text{ with } x_k \leqslant y \ (x_k \geqslant y) \ \forall\, k \in \mathbb{N} \text{ implies } \bar{x} \leqslant y \ (\underline{x} \geqslant y). \tag{1.6}$$

$\bar{x}$ is called the supremum and $\underline{x}$ the infimum of the sequence, and we write $\bar{x} = \sup\{x_k\}$, $\underline{x} = \inf\{x_k\}$.

For future reference we mention (without proof):

LEMMA 2. *Let* $[x], [y] \in \mathbb{I}(H)$ *with* $[x] \cap [y] \neq \varnothing$. *Then* $[x] \cap [y]$ *is an interval,* $[x] \cap [y] = [\sup(\underline{x}, \underline{y}), \inf(\bar{x}, \bar{y})]$, *and* $\rho[[x] \cap [y]] \leqslant \rho[x]$, $\rho[[x] \cap [y]] \leqslant \rho[y]$.

Moreover every element $x \in H$ can be expressed in a unique way as the difference of two positive elements; i.e.,

$$x = x^+ - x^- \qquad \text{with} \quad x^+, x^- \in H^+, \tag{1.7}$$

where

$$x^+ = \sup(x, \Theta) \qquad \text{and} \qquad x^- = \sup(-x, \Theta). \tag{1.8}$$

Furthermore, every $x \in H$ can be associated with an *absolute element* $|\, x \,| \in H^+$ defined by

$$|\, x \,| = x^+ + x^-. \tag{1.9}$$

$| x |$ can also be expressed as

$$| x | = \sup(x, -x), \tag{1.10}$$

and obviously

$$-x_0 \leqslant x \leqslant x_0 \Leftrightarrow | x | \leqslant x_0 \in H^+. \tag{1.11}$$

Let $\mathfrak{L}$ be a linear space of linear operators, where every element $L \in \mathfrak{L}$ maps the space H into itself. We introduce a partial ordering in $\mathfrak{L}$ as follows: If two operators $\underline{L}$, $\bar{L} \in \mathfrak{L}$ satisfy the relation $\underline{L}x \leqslant \bar{L}x \; \forall\, x \in H^+$, then $\underline{L} \leqslant \bar{L}$.[2] The operator $P = \bar{L} - \underline{L}$ is then a positive operator; i.e., we have $Px \in H^+$ for every $x \in H^+$. Corresponding to the definition of an interval, the subset $\{L \in \mathfrak{L} \mid \underline{L} \leqslant L \leqslant \bar{L}\}$ is denoted by $[\underline{L}, \bar{L}]$ or $[L]$, which we also call an *interval operator*. The set of all linear interval operators is denoted by $\mathbb{I}(\mathfrak{L})$. The radius and the mean value of an interval in $\mathbb{I}(\mathfrak{L})$ are defined analogously to (1.1) and (1.2), respectively.

Furthermore, let $\mathfrak{L}$ be a lattice. The supremum and infimum of two operators $L_1, L_2 \in \mathfrak{L}$ is denoted by $\sup(L_1, L_2)$ and $\inf(L_1, L_2)$, respectively.

An interval operator $[L]$ is called a *null interval operator* if $\mathfrak{O} \in [L]$, where $\mathfrak{O}$ denotes the null operator.

The use of a null interval operator $[L]$ on a null interval $[x]$ is defined as follows:

$$[L][x] = [-\sup(-\underline{L}, \bar{L}) \sup(-\underline{x}, \bar{x}), \sup(-\underline{L}, \bar{L}) \sup(-\underline{x}, \bar{x})]. \tag{1.12}$$

In the case where $\bar{L} = -\underline{L}$ and $\bar{x} = -\underline{x}$, (1.12) simplifies to

$$[L][x] = [-\bar{L}\bar{x}, \bar{L}\bar{x}]. \tag{1.13}$$

If $H = \mathbb{R}^n$, then we have for (1.12), $[L][x] \supseteq [L] \times [x]$, where $[L] \times [x]$ denotes the product of an interval matrix and an interval vector. (This is consistent with [7].) In the case where $\underline{L} = -\bar{L}$, we have $[L][x] = [L] \times [x]$.

2. *Convergence*

We now introduce an order-convergence in the space H. For this purpose we need some definitions and basic concepts:

Definition 1. The sequence $\{r_k\}$ is called *isotone* (*antitone*), and denoted by $\{r_k\} \nearrow$ ($\{r_k\} \searrow$), if $r_k \leqslant r_{k+1}$ ($r_k \geqslant r_{k+1}$) $\forall k \in \mathbb{N}$.

Definition 2. The sequence $\{r_k\} \nearrow$ (or $\{r_k\} \searrow$) is called *isotone convergent* (*antitone convergent*) to the limit $\hat{r} \in H$, and denoted by $\{r_k\} \nearrow \hat{r}$ ($\{r_k\} \searrow \hat{r}$), if $\sup\{r_k\} = \hat{r}$ ($\inf\{r_k\} = \hat{r}$).

[2] For convenience we also denote this relation by $\leqslant$.

DEFINITION 3. The sequence $\{r_k\}$ is called *order-convergent* to the limit $\hat{r} \in H$, and denoted by $\text{o-lim}_{k\to\infty} \{r_k\} = \hat{r}$, if two sequences $\{\underline{r}_k\}$ and $\{\bar{r}_k\}$ exist in H such that

$$\underline{r}_k \leqslant r_k \leqslant \bar{r}_k \qquad \forall k \in \mathbb{N}, \tag{2.1}$$

$$\{\underline{r}_k\} \nearrow \hat{r}, \qquad \{\bar{r}_k\} \searrow \hat{r}. \tag{2.2}$$

We now state

LEMMA 3. *Let $\{x_k\}$, $\{y_k\}$ be two order-convergent sequences in H with $\text{o-lim}_{k\to\infty}\{x_k\} = \hat{x}$, $\text{o-lim}_{k\to\infty}\{y_k\} = \hat{y}$, and let $\{c_k\}$ be a convergent sequence in $\mathbb{R}$ with $\lim_{k\to\infty} c_k = \hat{c}$. Then:*

$$\text{for every indefinite subsequence } \{x_{k_m}\}\text{: } \underset{m\to\infty}{\text{o-lim}}\{x_{k_m}\} = \hat{x}, \tag{2.3}$$

$$\underset{k\to\infty}{\text{o-lim}}\{x_k + y_k\} = \hat{x} + \hat{y}, \tag{2.4}$$

$$\underset{k\to\infty}{\text{o-lim}}\{c_k x_k\} = \hat{c}\hat{x}, \tag{2.5}$$

$$x_k \in H^+ \ \forall k \in \mathbb{N} \Rightarrow \hat{x} \in H^+, \tag{2.6}$$

$$\Theta \leqslant x_k \leqslant y_k \ \forall k \in \mathbb{N} \wedge \underset{k\to\infty}{\text{o-lim}}\{y_k\} = \Theta \Rightarrow \underset{k\to\infty}{\text{o-lim}}\{x_k\} = \Theta. \tag{2.7}$$

The statements of this lemma follow immediately from (1.3), (1.4), and Definition 3. Thus, the requirements on the concept of convergence are satisfied (see [13]).

Let $\{[x_k]\}$ be an interval sequence in $\mathbb{I}(H)$. This sequence can be associated elementwise with the sequence $\{\rho_k\} = \{\rho[x_k]\}$ (see (1.1)) in H^+.

DEFINITION 4. The interval sequence $\{[x_k]\}$ is called order-convergent to the interval limit $[x] = [\underline{x}, \bar{x}] \in \mathbb{I}(H)$, denoted by $\text{o-lim}_{k\to\infty}\{[x_k]\} = [x]$, if $\text{o-lim}_{k\to\infty}\{\underline{x}_k\} = \underline{x}$ and $\text{o-lim}_{k\to\infty}\{\bar{x}_k\} = \bar{x}$.

DEFINITION 5. The interval sequence $\{[x_k]\}$ is called *monotone* if $[x_{k+1}] \subseteq [x_k] \ \forall \ k \in \mathbb{N}$.

LEMMA 4. *If $\{[x_k]\}$ is a monotone interval sequence, then we have*

$$\{[x_k]\} \text{ is order-convergent}, \tag{2.8}$$

$$\{\rho_k\} \searrow . \tag{2.9}$$

The proof of Lemma 4 is left to the reader.

We will denote a monotone interval sequence by $\{[x_k]\} \searrow$.

DEFINITION 6. The interval sequence $\{[x_k]\}\searrow$ is called *point-convergent* if $\text{o-lim}_{k\to\infty}\{\rho_k\} = \Theta$.

DEFINITION 7. The interval sequence $\{[x_k]\}\searrow$ is called *linear point-convergent* if there exists a positive operator $C \in \mathfrak{L}$, independent of k, with $\text{o-lim}_{k\to\infty}\{C^k x\} = \Theta \; \forall x \in H^+$, such that $\rho_{k+1} \leqslant C\rho_k \; \forall k \in \mathbb{N}$.

LEMMA 5. *A linear point-convergent interval sequence* $\{[x_k]\}$ *is point-convergent.*

This follows immediately from (2.7) since

$$\Theta \leqslant \rho_{k+1} \leqslant C\rho_k \leqslant C^2\rho_{k-1} \leqslant \cdots \leqslant C^k\rho_1 ,$$

and according to Definition 7 we have $\text{o-lim}_{k\to\infty}\{C^k\rho_1\} = \Theta$.

DEFINITION 8. A point-convergent interval sequence $\{[x_k]\}\searrow$ is called *superlinear point-convergent* if there exists a sequence of positive operators $\{C_k\}$ in $\mathfrak{L}$ with $\text{o-lim}_{k\to\infty}\{C_k x\} = \Theta \; \forall x \in H^+$, such that $\rho_{k+1} \leqslant C_k\rho_k \; \forall k \in \mathbb{N}$.

For order-convergence of higher order we need a definition of a multiplication $\circ$ between positive elements in H^+ which satisfies the following conditions:

$$u, v \in H^+ \Rightarrow u \circ v = v \circ u \in H^+, \tag{2.10}$$

$$(u \circ v) \circ w = u \circ (v \circ w), \qquad u, v, w \in H^+, \tag{2.11}$$

$$u_1 \leqslant u_2 \wedge v_1 \leqslant v_2 \Rightarrow u_1 \circ v_1 \leqslant u_2 \circ v_2 , \qquad u_1 , v_1 , u_2 , v_2 \in H^+. \tag{2.12}$$

Furthermore $\rho^p = \rho^{p-1} \circ \rho$, $\rho \in H^+$, $p = 2, 3,\ldots$.

DEFINITION 9. A point-convergent interval sequence $\{[x_k]\}$ is called *point-convergent of order p* if there exists a positive operator $C \in \mathfrak{L}$, independent of k, such that $\rho_{k+1} \leqslant C\rho_k{}^p \; \forall k \in \mathbb{N}$.

Because H is a linear partially ordered space with order-convergence as the concept of convergence, we are able to introduce a pseudometric convergence in H using Definition 10 below (see also [13]) and a pseudometric convergence derived from (1.9). The pseudometric distance

$$d(x, y) = | x - y | \in H^+, \qquad x, y \in H, \tag{2.13}$$

satisfies all requirements of a pseudometric.

DEFINITION 10. The sequence $\{x_k\}$ is called *pseudometric convergent* to the limit $\hat{x} \in H$ if $\text{o-lim}_{k\to\infty}\{d(x_k , \hat{x})\} = \Theta$.

LEMMA 6. *The sequence* $\{x_k\}$ *is pseudometric convergent if and only if it is order-convergent.*

Proof. (1) Let $\{x_k\}$ be pseudometric convergent. With $d_k = | x_k - \hat{x} |$ we have, from Definition 10, o-$\lim_{k\to\infty}\{d_k\} = \Theta$. It follows from (1.11) that $\hat{x} - d_k \leqslant x_k \leqslant \hat{x} + d_k$, and therefore we obtain from Lemma 3 that o-$\lim_{k\to\infty}\{x_k\} = \hat{x}$.

(2) Let $\{x_k\}$ be order-convergent. Then we have from (2.1) that $\underline{x}_k - \hat{x} \leqslant x_k - \hat{x} \leqslant \bar{x}_k - \hat{x}$. This inequality remains valid if we add a negative element to the left side and a positive element to the right side. We then obtain $-(\bar{x}_k - \hat{x}) + (\underline{x}_k - \hat{x}) \leqslant x_k - \hat{x} \leqslant (\bar{x}_k - \hat{x}) - (\underline{x}_k - \hat{x})$. It follows then from (1.11) that $d_k = | x_k - \hat{x} | \leqslant \bar{x}_k - \underline{x}_k$, and so we obtain from Lemma 3 that o-$\lim_{k\to\infty} d_k = \Theta$.

3. *Equations in H*

Let there be given an interval $[x_0] \subseteq H$ and an operator T which maps $[x_0]$ into H. We want to find a solution $x^* \in [x_0]$ of the equation

$$Tx = \Theta. \tag{3.1}$$

Let the operator T satisfy:

ASSUMPTION I. To every interval $[x] \subseteq [x_0]$ there exists an $[L] = [\underline{L}, \bar{L}] \in \mathbb{I}(\mathfrak{L})$ and an $\tilde{L} \in [L]$ with the following properties:

$$\underline{L}(x_1 - x_2) \leqslant Tx_1 - Tx_2 \leqslant \bar{L}(x_1 - x_2) \text{ for all } x_1, x_2 \in [x] \text{ with } x_1 \geqslant x_2, \tag{3.2}$$

$$A = \tilde{L}^{-1} \text{ exists, and can be decomposed as } A = A^+ - A^-, \text{ where } A^+ \text{ and } A^- \text{ are positive operators.} \tag{3.3}$$

The linear operators $\underline{L}, \bar{L}$, and $\tilde{L}$, and therefore also A, are in general dependent on the interval $[x]$. Condition (3.2) is in fact just a generalized Lipschitz condition with respect to the given partial ordering. The Lipschitz operators $\underline{L}$ and $\bar{L}$ replace the usual Fréchet derivatives in the Newton method.

Furthermore let the following positive operators be defined:

$$| A | = A^+ + A^-, \qquad P = \bar{L} - \underline{L}. \tag{3.4}$$

Every $L \in [L]$ can be associated with a linear operator

$$R(L) = I - AL, \tag{3.5}$$

where A is the operator defined in (3.3) and I is the identity operator. With

$$\begin{aligned} \bar{R} &= I - A^+\underline{L} + A^-\bar{L}, \\ \underline{R} &= I - A^+\bar{L} + A^-\underline{L}, \end{aligned} \tag{3.6}$$

we obtain the inclusion

$$\underline{R} \leqslant R(L) \leqslant \bar{R} \qquad \text{for all } L \in [L]. \tag{3.7}$$

This follows immediately from $\underline{L} \leqslant L \leqslant \bar{L}$. Thus we have for all $x \in H^+$,

$$Rx = x - A^+Lx + A^-Lx \leqslant x - A^+\underline{L}x + A^-\bar{L}x = \bar{R}x,$$

and similarly $Rx \geqslant \underline{R}x$.

The difference of the operators $\bar{R}$ and $\underline{R}$ can be expressed as the product of two positive operators, namely,

$$\bar{R} - \underline{R} = (A^+ + A^-)(\bar{L} - \underline{L}) = |A| P. \tag{3.8}$$

As $R(\tilde{L}) = \mathfrak{O}$, it follows from (3.7) that

$$\underline{R} \leqslant \mathfrak{O} \leqslant \bar{R}, \tag{3.9}$$

which implies that $[\underline{R}, \bar{R}]$ is a null interval operator and that

$$R = \sup(-\underline{R}, \bar{R}) \tag{3.10}$$

is a positive operator.

Using the operators $\underline{R}$ and $\bar{R}$ defined in (3.6), we can derive a Lipschitz condition for the operator $I - AT$ corresponding to the Lipschitz condition for the operator T. To this end we introduce:

LEMMA 7. *Let $[x] \subseteq [x_0]$, and let the operator T satisfy Assumption* I. *Then the operator*

$$S = I - AT \tag{3.11}$$

maps the interval $[x]$ into H, and S satisfies the condition

$$\underline{R}(x_1 - x_2) \leqslant Sx_1 - Sx_2 \leqslant \bar{R}(x_1 - x_2) \qquad \textit{for all } x_1, x_2 \in [x] \textit{ with } x_1 \geqslant x_2 . \tag{3.12}$$

Proof. For $x_1, x_2 \in [x]$ with $x_1 \geqslant x_2$, (3.2) implies the inclusions

$$-A^+\bar{L}(x_1 - x_2) \leqslant -A^+(Tx_1 - Tx_2) \leqslant -A^+\underline{L}(x_1 - x_2),$$
$$A^-\underline{L}(x_1 - x_2) \leqslant A^-(Tx_1 - Tx_2) \leqslant A^-\bar{L}(x_1 - x_2).$$

Adding these inequalities and then adding the term $x_1 - x_2$ to the result, we obtain $x_1 - x_2 - A^+\bar{L}(x_1 - x_2) + A^-\underline{L}(x_1 - x_2) \leqslant x_1 - x_2 - ATx_1 + ATx_2 \leqslant x_1 - x_2 - A^+\underline{L}(x_1 - x_2) + A^-\bar{L}(x_1 - x_2)$. Using the notation of (3.6) and (3.11) we arrive then at (3.12).

To obtain an inclusion for $Sx_1 - Sx_2$ in the case when $x_1, x_2 \in [x]$ are chosen arbitrarily, we need:

LEMMA 8. *Let* $[x] \subseteq [x_0]$, *and let the operator* T *satisfy Assumption* I. *Then for* $x_1, x_2 \in [x]$ *with the decomposition* $x_1 - x_2 = (x_1 - x_2)^+ - (x_1 - x_2)^-$, *the operator* S *defined by* (3.11) *satisfies the condition*

$$\underline{R}(x_1 - x_2)^+ - \bar{R}(x_1 - x_2)^- \leqslant Sx_1 - Sx_2 \leqslant \bar{R}(x_1 - x_2)^+ - \underline{R}(x_1 - x_2)^-. \tag{3.13}$$

Proof. Let $x_1, x_2 \in [x] \subseteq [x_0]$. If we put $u = \inf(x_1, x_2)$, then, according to the lattice property $\underline{x} \leqslant u \leqslant \bar{x}$, we have $u \in [x]$. We now write $Sx_1 - Sx_2 = (Sx_1 - Su) - (Sx_2 - Su)$. Since $x_1 \geqslant u$ and $x_2 \geqslant u$, we can apply inequality (3.12) to both terms. We thus obtain

$$\underline{R}(x_1 - u) - \bar{R}(x_2 - u) \leqslant Sx_1 - Sx_2 \leqslant \bar{R}(x_1 - u) - \underline{R}(x_2 - u).$$

Furthermore, we have $x_1 - u = x_1 - \inf(x_1, x_2) = x_1 + \sup(-x_1, -x_2) = \sup(\Theta, x_1 - x_2) = (x_1 - x_2)^+$, and similarly $x_2 - u = x_1 - u - (x_1 - x_2) = (x_1 - x_2)^+ - (x_1 - x_2) = (x_1 - x_2)^-$. This concludes the proof of (3.13).

LEMMA 9. *Let* $\tilde{L} = \mu[L]$. *Then the operators* $\bar{R}$, $\underline{R}$, *and* R, *defined in* (3.6) *and* (3.10), *are related by*

$$\bar{R} = -\underline{R} = R. \tag{3.14}$$

Proof.

$$\begin{aligned} \bar{R} + \underline{R} &= (I - A^+\underline{L} + A^-\bar{L}) + (I - A^+\bar{L} + A^-\underline{L}) \\ &= 2I - (A^+ - A^-)\bar{L} - (A^+ - A^-)\underline{L} \\ &= 2I - A(\bar{L} + \underline{L}) = 2I - \tilde{L}^{-1}(2\tilde{L}) = \mathfrak{O}. \end{aligned}$$

Further, we obtain from (3.8) and (3.14)

$$R = \tfrac{1}{2} \mid A \mid P, \tag{3.15}$$

and from (3.13), (3.14), and (1.11) we obtain

$$\mid Sx_1 - Sx_2 \mid \leqslant R \mid x_1 - x_2 \mid. \tag{3.16}$$

In order to obtain methods with a higher speed of convergence, we replace Assumption I with the stronger

ASSUMPTION II. To every interval $[x] \subseteq [x_0]$ there exists an $[L] = [\underline{L}, \bar{L}] \in \mathbb{I}(\mathfrak{L})$ with the following properties:

$$\underline{L}(x - \tilde{x})^+ - \bar{L}(x - \tilde{x})^- \leqslant Tx - T\tilde{x} \leqslant \bar{L}(x - \tilde{x})^+ - \underline{L}(x - \tilde{x})^- \tag{3.17}$$

for $x \in [x]$ and $\tilde{x} = \mu[x]$,

$$A = \tilde{L}^{-1} \text{ exists, } \tilde{L} = \mu[L], \text{ and can be decomposed as}$$
$$A = A^+ - A^-, \text{ where } A^+ \text{ and } A^- \text{ are positive operators.} \qquad (3.18)$$

In a derivation similar to that of (3.16) we obtain from Assumption II the inequality

$$| Sx - S\tilde{x} | \leqslant R \, | x - \tilde{x} | \qquad \text{for} \quad x \in [x], \tilde{x} = \mu[x], \qquad (3.19)$$

where R is given by (3.15).

For every $x \in [x_0]$ we have hitherto had $Sx \in H$, where S is defined in (3.11). Now we will assume that S is also defined for a point interval $\underset{\cdot}{x} = [x, x] \subseteq [x_0]$. $S\underset{\cdot}{x}$, however, is not necessarily a point interval, but S of course has the property $Sx \in S\underset{\cdot}{x}$. For practical use we can in this way include the rounding errors.

4. *Iterative Methods*

Let φ be a mapping which maps every interval $[x] \in \mathbb{I}(H)$ into an element $\tilde{x} \in [x]$, and let the operator T satisfy Assumption I (or II).

With the initial interval $[x_0]$ the iteration scheme

$$[x_{k+1}] = [S_k\underset{\cdot}{\tilde{x}}_k + [R_k]\,[[x_k] - \underset{\cdot}{\tilde{x}}_k]] \cap [x_k], \qquad k = 0, 1, 2,..., \qquad (4.1)$$

will, according to Lemma 2, produce a monotone interval sequence

$$[x_0] \supseteq [x_1] \supseteq [x_2] \supseteq \cdots. \qquad (4.2)$$

Here $S_k = I - A_k T$, $\tilde{x}_k = \varphi[x_k]$, and $\underset{\cdot}{\tilde{x}}_k$ is a point interval. The meaning of $S_k\underset{\cdot}{\tilde{x}}_k$ is explained in Section 3. The linear operators A_k, defined by (3.3), as well as the null interval operators $[R_k] = [\underline{R}_k, \overline{R}_k]$ are associated with the intervals $[x_k]$. $[R_k][[x_k] - \underset{\cdot}{\tilde{x}}_k]$ is defined by (1.12), since $[x_k] - \underset{\cdot}{\tilde{x}}_k$ is a null interval, and by (3.10) $R_k = \sup(-\underline{R}_k, \overline{R}_k)$.

THEOREM 1 (Principle of inclusion). *If there exists a solution x^* of equation* (3.1), *then*

$$x^* \in [x_0] \Rightarrow x^* \in [x_k] \qquad \forall k \in \mathbb{N}.$$

Proof. We will prove that $x^* \in [x_0]$ implies $x^* \in [x_1]$. If we put $x_1 = x^*$, $x_2 = \tilde{x}_0$, $S = S_0$, $\underline{R} = \underline{R}_0$, and $\overline{R} = \overline{R}_0$ in (3.13), then we obtain

$$\underline{R}_0(x^* - \tilde{x}_0)^+ - \overline{R}_0(x^* - \tilde{x}_0)^- \leqslant S_0x^* - S_0\tilde{x}_0 \leqslant \overline{R}_0(x^* - \tilde{x}_0)^+ - \underline{R}_0(x^* - \tilde{x}_0)^-.$$

Now $S_0x^* = x^*$, so from (1.9) and (3.10) we obtain $S_0\tilde{x}_0 - R_0 \mid x^* - \tilde{x}_0 \mid \leqslant x^* \leqslant S_0\tilde{x}_0 + R_0 \mid x^* - \tilde{x}_0 \mid$. Because $\mid x^* - \tilde{x}_0 \mid = \sup(x^* - \tilde{x}_0, \tilde{x}_0 - x^*) \leqslant \sup(\bar{x}_0 - \tilde{x}_0, \tilde{x}_0 - \underline{x}_0)$ and $\underline{S_0\tilde{x}_0} \leqslant S_0\tilde{x}_0 \leqslant \overline{S_0\tilde{x}_0}$, we have from (1.12) $x^* \in [S_0\tilde{x}_0 + [R_0][[x_0] - \tilde{x}_0]]$, and therefore $x^* \in [x_1]$. By continuing this process we easily obtain an inductive proof of Theorem 1.

Corollary to Theorem 1. *If the iteration* (4.1) *terminates for some* k *because the intersection is empty, then no solution of Eq.* (3.1) *can be contained in* $[x_0]$.

Choosing $\varphi[x] = \mu[x]$, the iteration scheme (4.1) simplifies to

$$[x_{k+1}] = [S_k\tilde{x}_k + [R_k]\,[-\rho_k, \rho_k]] \cap [x_k], \qquad k = 0, 1, 2,..., \tag{4.3}$$

where $\rho_k = \rho[x_k]$.

If we put $A_k = A_0$, for all $k \in \mathbb{N}$, then we obtain $S_k = S_0$, $[R_k] = [R_0]$, and the iteration scheme (4.3) simplifies to

$$[x_{k+1}] = [S_0\tilde{x}_k + [R_0]\,[-\rho_k, \rho_k]] \cap [x_k], \qquad k = 0, 1, 2,.... \tag{4.4}$$

This iteration corresponds to the simplified Newton iteration (see [13]).

The following theorem says something about the convergence of the interval sequence (4.2) defined by (4.3).

Theorem 2. *Let the iteration scheme* (4.3) *satisfy*

(1) $\rho[S_k\tilde{x}_k] \leqslant \epsilon \; \forall \, k \in \mathbb{N}_0$.[3]

We assume that there exists at least one solution $x^* \in [x_0]$ *of Eq.* (3.1) *and a linear operator* $C \in \mathfrak{L}$, *independent of* k, *which satisfies the following conditions*:

(2) $R_k \leqslant C \; \forall \, k \in \mathbb{N}_0$,

(3) $(I - C)^{-1}$ *exists and is positive.*

For the interval sequence defined by (4.3) *we then have*

$$\underset{k\to\infty}{\text{o-lim}}\{[x_k]\} = [x] \qquad \text{with} \qquad x^* \in [x] \tag{4.5}$$

and

$$\rho[x] \leqslant (I - C)^{-1}\epsilon. \tag{4.6}$$

Proof. Since $x^* \in [x_0]$, it follows from Theorem 1 that the interval sequence $\{[x_k]\}$ is indefinite, and furthermore, because it is monotone, we can use Lemma 4,

[3] $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

and so prove (4.5). From (4.3), using Lemmas 1 and 2 as well as (1.12) and assumptions 1 and 2 of Theorem 2, it follows that

$$\rho_{k+1} = \rho[x_{k+1}] \leqslant \rho[S_k \tilde{x}_k] + R_k \rho_k \leqslant \epsilon + C\rho_k .$$

Applying Lemma 3 and the fact that o-$\lim_{k\to\infty} \rho_k = \rho[x] = \rho$, we obtain

$$\rho \leqslant \epsilon + C\rho \Rightarrow (I - C)\rho \leqslant \epsilon,$$

from which, according to assumption 3 of Theorem 2, (4.6) follows.

COROLLARY. *If we replace assumption* 1 *in Theorem* 2 *by*

(1a) $\rho[S_k \tilde{x}_k] \leqslant \epsilon_k \; \forall k \in \mathbb{N}_0$ *with* o-$\lim_{k\to\infty}\{\epsilon_k\} = \Theta$,

then (4.5) *and* (4.6) *sharpen to*

$$\text{o-}\lim_{k\to\infty}\{[x_k]\} = x, \tag{4.7}$$

and so $\rho[x] = \Theta$.

Thus from Theorem 1 it follows that $x^* = x$ and consequently the uniqueness of the solution x^* in $[x_0]$ of Eq. (3.1). The practical significance of assumption 1a is that the accuracy of the calculation increases with decreasing iteration error.

As the condition of convergence for the general iteration scheme (4.1) turns out to be less favorable when $\varphi[x] \neq \mu[x]$, we will confine ourselves to the iteration scheme (4.3) (or (4.4)).

In order to be able to say something about the speed of convergence we introduce:

DEFINITION 11. The iteration scheme (4.3) (or (4.4)) converges *linearly* (*superlinearly* or *with order p*) if the interval sequence (4.2) under assumption 1 of Theorem 2 with $\epsilon = \Theta$ is linearly (superlinearly or with order p) point convergent.

By choosing ϵ_k in assumption 1a in the corollary to Theorem 2 so that the speed of convergence (as in the case $\epsilon = \Theta$) is not changed, we see that Definition 11 is also meaningful for $\epsilon \geqslant \Theta$.

THEOREM 3. *If there exists an operator* $C \in \mathfrak{L}$, *independent of k, such that*

(1) $R_k \leqslant C \; \forall k \in \mathbb{N}_0$,

(2) o-$\lim_{k\to\infty}\{C^k x\} = \Theta \; \forall x \in H^+$,

then the iteration scheme (4.3) (*or* (4.4)) *will be at least linearly convergent.*

Proof. From (4.3) (or (4.4)) with $\epsilon = \Theta$, and in consideration of assumption 1 of Theorem 3, we obtain $\rho_{k+1} \leqslant C\rho_k \; \forall k \in \mathbb{N}_0$. Thus, according to assumption 2 of Theorem 3 and Definition 7, we have linear point convergence.

COROLLARY 1. *If assumption* 1 *of Theorem* 3 *is replaced by*

(1a) $\text{o-lim}_{k\to\infty}\{R_k x\} = \Theta \; \forall x \in H^+,$

then, according to Definitions 8 *and* 11, *the iteration scheme* (4.3) *will be at least superlinearly convergent.*

COROLLARY 2. *If there exists an operator* $C \in \mathfrak{L}$, *independent of* k, *such that*

(1b) $R_k\rho_k \leqslant C\rho_k{}^p \; \forall k \in \mathbb{N}_0$, $p \geqslant 2$, *and* $\text{o-lim}_{k\to\infty}\{\rho_k\} = \Theta$,

then, according to Definitions 9 *and* 11, *the iteration scheme* (4.3) *will be at least convergent of order* p.

Remark. For the simplified iteration scheme (4.4), assumption 1 of Theorem 3 is satisfied with $C = R_0$.

II. APPLICATIONS

In the following we show some examples of how the abstract analysis of Part I can be applied to concrete problems.

5. *Equations with One Unknown*

Let $H = \mathbb{R}$ with the usual partial ordering $\leqslant$. The lattice properties (1.5), (1.6) are trivially satisfied. Let there be given an interval $[x_0]$ and a real-valued function $f(x)$, defined for all $x \in [x_0]$, which in every $[x] \subseteq [x_0]$ satisfies the Lipschitz condition

$$\underline{L}(x_1 - x_2) \leqslant f(x_1) - f(x_2) \leqslant \bar{L}(x_1 - x_2) \qquad \text{for} \quad x_1, x_2 \in [x] \text{ with } x_1 \geqslant x_2. \tag{5.1}$$

With $Tx = f(x)$, condition (3.2) of Assumption I is fulfilled. With $\tilde{L} = \mu[L] = \frac{1}{2}(\underline{L} + \bar{L})$, one obtains

$$A = \frac{2}{\underline{L} + \bar{L}} \qquad \text{and} \qquad R = \frac{\bar{L} - \underline{L}}{\underline{L} + \bar{L}}. \tag{5.2}$$

Assumptions 2 and 3 of Theorem 2 (or assumptions 1 and 2 of Theorem 3) are satisfied if

$$\bar{L}_0 \underline{L}_0 > 0.\text{[4]} \tag{5.3}$$

It is readily seen that (5.3) implies $R_k \leqslant R_0 = C < 1\ \forall k \in \mathbb{N}$. Thus condition (5.3) guarantees linear convergence of the iteration scheme (4.3) (or (4.4)).

If $f(x)$ is continuously differentiable in $[x_0]$ and we put

$$\underline{L} = \min_{x \in [x]} f'(x), \qquad \bar{L} = \max_{x \in [x]} f'(x), \tag{5.4}$$

then (5.3) is equivalent to

$$|f'(x)| \geqslant \alpha > 0 \qquad \forall x \in [x_0]. \tag{5.5}$$

Also, assumption 1a of Corollary 1 to Theorem 4 is then satisfied, i.e., the iteration scheme (4.3) is superlinearly convergent.

This statement is not quite obvious. As Figs. 1 and 2 illustrate, (5.5) is not sufficient for convergence of the conventional Newton method.

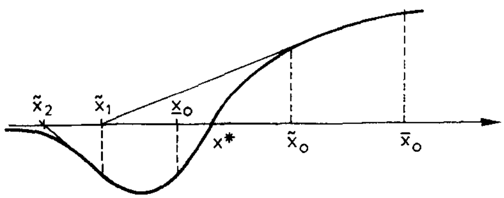

FIGURE 1

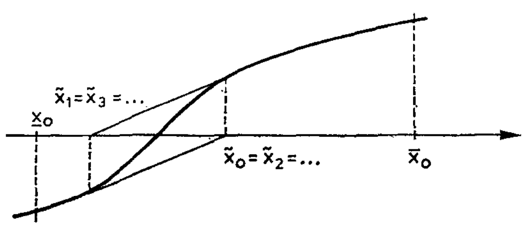

FIGURE 2

We now assume that $f(x)$ is twice continuously differentiable in $[x_0]$ and let

$$\beta = \max_{x \in [x]} |f''(x)|\,. \tag{5.6}$$

[4] The index 0 means that $\underline{L}_0$ and $\bar{L}_0$ are associated with the interval $[x_0]$. The same holds for R_0 and β_0.

For $Tx = f(x)$ and with

$$\underline{L} = f'(\tilde{x}) - \tfrac{1}{2}\beta\rho, \qquad \bar{L} = f'(\tilde{x}) + \tfrac{1}{2}\beta\rho,$$
$$(\tilde{x} = \tfrac{1}{2}(\underline{x} + \bar{x}), \qquad \rho = \tfrac{1}{2}(\bar{x} - \underline{x})), \tag{5.7}$$

condition (3.17) of Assumption II is satisfied and we obtain

$$A = \frac{1}{f'(\tilde{x})}, \qquad R = \frac{\beta\rho}{2\,|f'(\tilde{x})|}. \tag{5.8}$$

From (5.5) and the condition

$$\rho_0 < 2\alpha/\beta_0 \tag{5.9}$$

it follows that

$$R_k \leqslant R_0 = C < 1 \qquad \text{and} \qquad R_k \leqslant (\beta_0/2\alpha)\rho_k \quad \forall k \in \mathbb{N}.$$

Assumption 1b of Corollary 2 to Theorem 4 is then satisfied for $p = 2$, and so the iteration scheme (4.3) has quadratic convergence.

6. *Nonlinear Systems of Equations*

Let $H = \mathbb{R}^n$, and let $x \in \mathbb{R}^n$ denote a vector with ith component $x^{(i)}$. Then

$$x_1 \leqslant x_2 \Leftrightarrow x_1^{(i)} \leqslant x_2^{(i)} \qquad \text{for} \quad i = 1(1)n$$

defines a partial ordering in $\mathbb{R}^n$. The vector $x^+ \in \mathbb{R}^n$, according to (1.8), contains all nonnegative components of x (the negative components of x are replaced by zeros) and similarly $-x^- \in \mathbb{R}^n$ contains all nonpositive components (the positive components are replaced by zeros). We also have $|\,x\,|^{(i)} = |\,x^{(i)}\,|$ for $i = 1(1)n$, and the lattice properties (1.5), (1.6) are obviously satisfied.

Let there be given an interval $[x_0]$ (interval vector) and an n-dimensional vector function $f(x)$ with components f_i $(i = 1(1)n)$ which maps every $x \in [x_0]$ into H. We assume that $f(x)$ satisfies the following Lipschitz condition: For two vectors $x_1, x_2 \in [x] \subseteq [x_0]$ with

$$x_1^{(m)} = x_2^{(m)} \quad \text{for } m \neq j \qquad \text{and} \qquad x_1^{(j)} \geqslant x_2^{(j)} \quad (j = 1(1)\,n)$$

we have for $i = 1(1)\,n$: (6.1)

$$\underline{l}_{ij}(x_1^{(j)} - x_2^{(j)}) \leqslant f_i(x_1) - f_i(x_2) \leqslant \bar{l}_{ij}(x_1^{(j)} - x_2^{(j)}).$$

The operator T, with $Tx = f(x)$ and the $(n \times n)$ matrices $\underline{L} = \{\underline{l}_{ij}\}$, $\bar{L} = \{\bar{l}_{ij}\}$, then satisfies condition (3.2) of Assumption I. The matrices A, A^+, A^-, $\underline{R}$, and $\bar{R}$ are given by (3.3) and (3.6), or (3.15) if $\tilde{L} = \frac{1}{2}(\underline{L} + \bar{L})$. Thus A^+ contains all nonnegative coefficients of A and zeros elsewhere while $-A^-$ contains all negative coefficients of A and zeros elsewhere. Thus the coefficients of $|A|$ are the absolute values of the coefficients of A.

If $[L] = [\underline{L}, \bar{L}]$ denotes the interval matrix with bounds $\underline{L}$ and $\bar{L}$, and similarly $[R] = [\underline{R}, \bar{R}]$, then

$$[R] = \dot{I} - \dot{A}[L]. \tag{6.2}$$

That is, $[R]$ is the interval matrix computed from (6.2) by means of the interval operations discussed in [5, 7]. Condition (4.3) of Assumption I is satisfied for all $\tilde{L}_k$, provided L^{-1} exists for all $L \in [L_0]$.

Let $\sigma(R)$ denote the spectral radius of the matrix R. The following theorem gives some information regarding the assumptions under which condition (3.3) is satisfied.

THEOREM 4. *Let $A_0 = (\frac{1}{2}(\underline{L}_0 + \bar{L}_0))^{-1}$ exist, and let*

$$\sigma(R_0) < 1. \tag{6.3}$$

Then every matrix $L \in [L_0]$ is regular.

Proof. Assume Theorem 4 were false. Then there would exist a matrix $L' \in [L_0]$ and a vector $z \neq \Theta$ with $A_0 L' z = \Theta$. From (3.5), (3.7), and (3.14) the inclusion for $A_0 L'$ follows:

$$I - R_0 \leqslant A_0 L' \leqslant I + R_0 .$$

From these inequalities and from the equation with z above, we obtain by decomposition of $z = z^+ - z^-$ (by (1.7))

$$z^+ - R_0 z^+ \leqslant A_0 L' z^+ = A_0 L' z^- \leqslant z^- + R_0 z^-,$$

and analogously

$$z^- - R_0 z^- \leqslant z^+ + R_0 z^+.$$

Thus by (1.9) we have

$$-R_0 |z| \leqslant z \leqslant R_0 |z|,$$

and because of (1.11) we have

$$|z| \leqslant R_0 |z|.$$

By repeated use of this inequality we obtain

$$|z| \leqslant R_0^k |z| \qquad \forall k \in \mathbb{N}.$$

As $R_0^k \to \mathbb{O}$ for $k \to \infty$, where $\mathbb{O}$ is the null matrix, this inequality can only be satisfied if, in contradiction to the assumption, $z = \Theta$.

Also the uniqueness of a solution of $f(x) = \Theta$ in $[x_0]$ follows from the assumptions of Theorem 4, since to every pair $x_1, x_2 \in [x_0]$ there exists a regular matrix $L(x_1, x_2) \in [L]$ with the property

$$f(x_1) - f(x_2) = L(x_1, x_2) \cdot (x_1 - x_2).$$

Furthermore, condition (6.3) of Theorem 4 is crucial for the convergence of the simplified iteration scheme (4.4) since for $\sigma(R_0) < 1$ with R_0 positive we have

$$(I - R_0)^{-1} = \sum_{r=0}^{\infty} R_0^r > \mathbb{O},$$

so condition 3 of Theorem 2 is then satisfied.

Assumption (6.3), however, is not sufficient to prove convergence of the iteration scheme (4.3), because $\sigma(R_0) < 1$ does not necessarily imply $R_k \leqslant C = R_0$. Therefore the following theorem of convergence requires a sharper condition:

THEOREM 5. *Let there exist a solution $x^* \in [x_0]$ of the equation $f(x) = \Theta$, and let $\tilde{L}_0 = \frac{1}{2}(\underline{L}_0 + \bar{L}_0)$ be nonsingular, and*

$$\sigma(R_0) < \tfrac{1}{2}. \tag{6.4}$$

Then the iteration scheme (4.3) *is at least linearly convergent.*

Proof. From the existence of a solution it follows, just as in the proof of Theorem 2, that the interval sequence (4.2) is indefinite. Furthermore, with (6.4) the assumptions of Theorem 4, and therefore also Assumption I, are satisfied.

From (3.15) we have for all $k \in \mathbb{N}_0$

$$R_k = \tfrac{1}{2} \mid A_k \mid (\bar{L}_k - \underline{L}_k) \leqslant \tfrac{1}{2} \mid A_k \mid (\bar{L}_0 - \underline{L}_0).$$

With $A_k = \tilde{L}_k^{-1}$ and $B_k = A_0(\tilde{L}_k - \tilde{L}_0)$ we find

$$A_k = (I + B_k)^{-1} A_0,$$

from which we obtain

$$R_k \leqslant |(I + B_k)^{-1}| \, R_0.$$

Since $\mid B_k \mid \leqslant \mid A_0 \mid \mid \tilde{L}_k - \tilde{L}_0 \mid \leqslant \mid A_0 \mid \cdot \frac{1}{2}(\bar{L}_0 - \underline{L}_0) = R_0$ and $\sigma(R_0) < \frac{1}{2}$, then $\sigma(\mid B_k \mid) < \frac{1}{2}$ also, and thus

$$R_k \leqslant \left| \sum_{r=0}^{\infty} B_k^r \right| R_0 \leqslant \left(\sum_{r=0}^{\infty} \mid B_k \mid^r \right) R_0 \leqslant \sum_{r=1}^{\infty} R_0^r = (I - R_0)^{-1} R_0.$$

With $C = (I - R_0)^{-1}R_0$ and $\sigma(C) < 1$ the assumptions of Theorem 3 are satisfied, and so Theorem 5 is proved.

COROLLARY 1. *If the functional matrix of $f(x)$ exists and is continuous in $[x_0]$, then, under the assumptions of Theorem 5, the convergence will be at least superlinear.*

To see this we observe that with

$$\underline{l}_{ij}^{(k)} = \min_{x\in[x_k]} \left(\frac{\partial f_i}{\partial x^{(j)}}\right), \qquad \bar{l}_{ij}^{(k)} = \max_{x\in[x_k]} \left(\frac{\partial f_i}{\partial x^{(j)}}\right), \tag{6.5}$$

$$\underline{L}_k = \{\underline{l}_{ij}^{(k)}\}, \qquad \bar{L}_k = \{\bar{l}_{ij}^{(k)}\},$$

(3.2) is satisfied, and

$$\underset{k\to\infty}{\text{o-lim}}\{(\bar{L}_k - \underline{L}_k)\, x\} = \Theta \quad \forall x \in (\mathbb{R}^n)^+ \qquad \text{and} \qquad \underset{k\to\infty}{\text{o-lim}}\, \rho_k = \Theta. \tag{6.6}$$

Similar to the proof of Theorem 5 we can estimate R_k as follows:

$$R_k = \tfrac{1}{2} \mid A_k \mid (\bar{L}_k - \underline{L}_k) \leqslant \tfrac{1}{2}(I - R_0)^{-1} \mid A_0 \mid (\bar{L}_k - \underline{L}_k), \tag{6.7}$$

and so

$$\underset{k\to\infty}{\text{o-lim}}\{R_k x\} = \Theta \quad \forall x \in (\mathbb{R}^n)^+ \qquad \text{and} \qquad \underset{k\to\infty}{\text{o-lim}}\, \rho_k = \Theta.$$

COROLLARY 2. *If we assume that $f(x)$ is twice continuously differentiable, and if we put*

$$m_{jh}^{(i)} = \max_{x\in[x_0]} \left| \frac{\partial^2 f_i}{\partial x^{(j)}\, \partial x^{(h)}} \right|, \qquad i, j, h = 1(1)\, n, \tag{6.8}$$

then, with

$$\underline{l}_{ij}^{(k)} = \left(\frac{\partial f_i}{\partial x^{(j)}}\right)_{x=\tilde{x}_k} - \frac{1}{2} \sum_{h=1}^{n} m_{jh}^{(i)} \rho_k^{(h)}, \qquad \underline{L}_k = \{\underline{l}_{ij}^{(k)}\},$$
$$\bar{l}_{ij}^{(k)} = \left(\frac{\partial f_i}{\partial x^{(j)}}\right)_{x=\tilde{x}_k} + \frac{1}{2} \sum_{h=1}^{n} m_{jh}^{(i)} \rho_k^{(h)}, \qquad \bar{L}_k = \{\bar{l}_{ij}^{(k)}\}, \tag{6.9}$$

Assumption (3.17) *is satisfied. If, further, we make the assumptions of Theorem* 5, *then Assumption* (3.18) *with $A_k = \tilde{L}_k^{-1}$, where $\tilde{L}_k$ denotes the functional matrix of $f(x)$ for $x = \tilde{x}_k$, is also satisfied.*

Let $M_i = \{m_{jh}^{(i)}\}$ denote the $(n \times n)$ matrix with elements from (6.8), let (x, y) denote the vector inner product of the vectors $x, y \in \mathbb{R}^n$, and let $(\rho_k, M_i\rho_k) \leqslant \lambda_i(\rho_k, \rho_k)$ for $i = 1(1)n$, $k \in \mathbb{N}_0$ (λ_i is the spectral radius of the positive symmetric matrix M_i). With $\rho_k{}^2 = \{(\rho_k^{(i)})^2\}$,

$$Q = \begin{pmatrix} \lambda_1 & \lambda_1 & \cdots & \lambda_1 \\ \lambda_2 & \lambda_2 & \cdots & \lambda_2 \\ \cdots & \cdots & \cdots & \cdots \\ \lambda_n & \lambda_n & \cdots & \lambda_n \end{pmatrix}$$

and $C = \frac{1}{2}(I - R_0)^{-1} \mid A_0 \mid Q$, (6.7) then implies

$$R_k\rho_k \leqslant C\rho_k{}^2 \qquad \forall\, k \in \mathbb{N}_0 \tag{6.10}$$

as the condition for quadratic convergence (see Definition 9). If the assumptions of Theorem 5 are satisfied for the matrices $\underline{L}_0$ and $\bar{L}_0$ defined by (6.9) for $k = 0$ and the matrix R_0 derived from these, then, according to Theorem 4, the matrices $A_k \;\forall\, k \in \mathbb{N}_0$ exist and the iteration scheme (4.3) is by virtue of (6.10) at least quadratically convergent.

Remark 1. In the case $[S_0\tilde{x}_0 + [R_0][[x_0] - \tilde{x}_0]] \subseteq [x_0]$, we have $S_0[x_0] = \{S_0x \mid x \in [x_0]\} \subseteq [x_0]$. According to the Brower fixed-point theorem, the existence of a solution $x^* \in [x_0]$ follows from the continuity of S_0.

Remark 2. The simplified iteration scheme (4.4) is in general sufficient for practical use. In contrast to (4.1) (or (4.3)), there is the advantage here that not only is the condition for convergence (6.3) weaker, but, most important, there is considerably less computational work since the matrices $\underline{L}_0$ and $\bar{L}_0$ are computed from (6.1) just once. The same holds for A_0 and R_0.

If ρ_k or $\rho_k/\tilde{x}_k$ after k steps does not decrease further, the iteration is terminated. The domain of rounding errors has then been reached, and ρ_k is, according to Theorem 1, an exact error bound. Thus, in contrast to most other methods, we obtain an exact error estimation without using bounds for the inverse functional matrix and the second derivatives. Very often $\sigma(R_0) \lll 1$. Practical experience has shown that in this case the number of steps required by the iteration for the simplified scheme (4.4) is hardly greater than that for a superlinearly or quadratically convergent method.

Hitherto it has been common for similar interval methods to compute the interval matrix $[R]$ by (6.2). Nevertheless, if $\underline{R} = -\bar{R}$ then this is superfluous. For the computation of $-\underset{\bullet}{A}[L]$ we only require an arithmetic by which, depending upon whether $(-a_{ij})$ is positive or negative, we can perform the multiplication $(-a_{ij})\bar{l}_{jh}$ (or $(-a_{ij})\underline{l}_{jh}$) with upward rounding. The same is true for the addition of $\underset{\bullet}{I}$ and $-\underset{\bullet}{A}[L]$.

7. *The Algebraic Eigenvalue Problem*

Let there be given the eigenvalue problem

$$\begin{aligned} (G - \lambda I)\,\xi &= 0, \\ (\xi, l) - 1 &= 0, \end{aligned} \tag{7.1}$$

with a real $(n \times n)$ matrix G and a vector $l \in \mathbb{R}^n$, with respect to which the eigenvector ξ is normalized. ((ξ, l) denotes the vector inner product of ξ and l.) Let λ denote a real eigenvalue. With $x = \binom{\xi}{\lambda}$ and the operator T defined by

$$Tx = \begin{pmatrix} (G - \lambda I)\,\xi \\ (\xi, l) - 1 \end{pmatrix} \tag{7.2}$$

we can interpret (7.1) as a nonlinear system of equations in $\mathbb{R}^{n+1}$; i.e., (7.1) is actually just a special case of the problem of Section 6.

Assuming we have an initial interval $[x_0] = \binom{[\xi_0]}{[\lambda_0]}$, where $[\lambda_0]$ contains a real eigenvalue λ^* and $[\xi_0]$ the corresponding eigenvector ξ^*, the iteration scheme (4.3) (or (4.4)) will lead to an improvement of the interval limits of $[\lambda_0]$ and $[\xi_0]$. With the $(n+1) \times (n+1)$ matrices

$$\underline{L} = \begin{pmatrix} G - \bar{\lambda} I & -\bar{\xi} \\ l' & 0 \end{pmatrix}, \quad \bar{L} = \begin{pmatrix} G - \underline{\lambda} I & -\underline{\xi} \\ l' & 0 \end{pmatrix}, \tag{7.3}$$

(3.2) of Assumption I will be satisfied. (l' denotes the transposed vector of l.)

We achieve a substantial improvement of the iteration for enclosing an eigenpair of (7.1), however, if we put

$$\underline{L} = \begin{pmatrix} G - \bar{\lambda} I & -\tilde{\xi} \\ l' & 0 \end{pmatrix}, \quad \bar{L} = \begin{pmatrix} G - \underline{\lambda} I & -\tilde{\xi} \\ l' & 0 \end{pmatrix}. \tag{7.4}$$

Condition (3.17) of Assumption II is then satisfied since

$$\begin{pmatrix} (G - \lambda I)\,\xi - (G - \tilde{\lambda} I)\,\tilde{\xi} \\ (l', \xi - \tilde{\xi}) \end{pmatrix} = \begin{pmatrix} G - \lambda I & -\tilde{\xi} \\ l' & 0 \end{pmatrix} \begin{pmatrix} \xi - \tilde{\xi} \\ \lambda - \tilde{\lambda} \end{pmatrix}.$$

Estimating λ from above and below, we can verify the inclusion (3.17). If E denotes the $(n+1) \times (n+1)$ matrix

$$E = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix},$$

with I as the nth-order identity matrix. It follows then from (7.4) that

$$\bar{L} - \underline{L} = (\bar{\lambda} - \underline{\lambda}) \cdot E = 2\rho[\lambda] \cdot E$$

and that

$$\tilde{L} = \tfrac{1}{2}(\bar{L} + \underline{L}) = \begin{pmatrix} G - \tilde{\lambda} I & -\tilde{\xi} \\ l' & 0 \end{pmatrix}.$$

We now assume that the assumptions of Theorem 4 are satisfied and hence all $\tilde{L}_k$ are nonsingular.

By (3.15) we obtain for R the expression

$$R = \rho[\lambda] \begin{pmatrix} A_{11} & 0 \\ a_{21} & 0 \end{pmatrix} \quad \text{with} \quad |A| = \begin{pmatrix} A_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \tag{7.5}$$

The inequality

$$\sigma(R_0) = \rho[\lambda_0] \cdot \sigma(|\tilde{A}_0|) < 1 \qquad (\text{or} < \tfrac{1}{2})$$

with

$$|\tilde{A}_0| = \begin{pmatrix} (A_0)_{11} & 0 \\ (a_0)_{21} & 0 \end{pmatrix} \tag{7.6}$$

will then be a sufficient criterion for convergence of the simplified iteration scheme (4.4) (or the general scheme (4.3)).

Consequently, the convergence depends on the limits of the eigenvalues but not on those of the eigenvectors. Since by virtue of the normalization it is often possible to give rough bounds on the eigenvector (for example, the eigenvector corresponding to the largest eigenvalue of a nonnegative matrix), the iterative method (4.3) (or (4.4)) is also able to determine the eigenvector provided that an inclusion of the eigenvalue is known. Further, the process automatically provides an error estimate for the eigenvector.

If the interval $[\lambda_0]$ contains more than one eigenvalue, or if there exists a multiple eigenvalue such that (7.1) has no unique solution, then (7.6) cannot be satisfied, since this, because of the guaranteed convergence of the interval sequence, would lead to a contradiction.

The method described here can be extended to include complex eigenvalues and eigenvectors. This is due to the fact that a complex eigenvalue problem can always be reformulated as a nonlinear system of equations in the real Euclidean space. A detailed description of such a method can be found in [26].

References

1. E. Adams and W. F. Ames, "On Contracting Interval Iteration for Nonlinear Problems in $\mathbb{R}^n$," Center for Applied Mathematics No. 7, University of Georgia 1978.
2. G. Alefeld, Über die aus monoton zerlegbaren Operatoren gebildeten Iterationsverfahren, *Computing* (*Arch. Elektron. Rechnen*) **6** (1970), 161–172.

3. U. ALLENDÖRFER, "Abbrechkriterium und numerische Konvergenz in halbgeordneten Mengen," Diplomarbeit, Institut für Praktische Mathematik der Universität Karlsruhe.
4. N. APOSTOLATOS, U. KULISCH, R. KRANCZYK, B. LORTZ, K. NICKEL, AND H.-W. WIPPERMAN, The algorithmic language triplex ALGOL 60, *Numer. Math.* 11 (1968), 175–180.
5. N. APOSTOLATOS AND U. KULISCH, Grundlagen einer Maschinenintervallarithmetik, *Computing (Arch. Elektron. Rechnen)* 2 (1967), 89–104.
6. N. APOSTOLATOS AND U. KULISCH, Approximation der erweiterten Maschinenintervallarithmetik, *Computing (Arch. Elektron. Rechnen)* 2 (1967), 181–194.
7. N. APOSTOLATOS AND U. KULISCH, Grunzüge einer Intervallrechnung für Matrizen und einige Anwendungen, *Elektron. Rechenanlagen* **10** (1968), 73–83.
8. G. BIRKHOFF, "Lattice-Theory," Amer. Math. Soc., Providence, R. I., 1967.
9. O. CAPRANI AND K. MADSEN, Contraction mappings in interval analysis, *BIT* **15** (1975), 362–366.
10. O. CAPRANI AND K. MADSEN, Iterative methods for interval inclusion of fixed points, *BIT* **18** (1978), 42–51.
11. L. COLLATZ AND J. SCHRÖDER, Einschließen der Lösungen von Randwertaufgaben, *Numer. Math.* **1** (1959), 61–72.
12. L. COLLATZ, Monotonie und Extremalprinzipien beim NEWTONschen Verfahren, *Numer. Math.* **3** (1961), 99–106.
13. L. COLLATZ, "Funktionalanalysis und numerische Mathematik," Springer–Verlag, Berlin/Göttingen/Heidelberg, 1964.
14. L. COLLATZ, Einige abstrakte Begriffe in der numerischen Mathematik (Anwendung der Halbordnung), *Computing (Arch. Elektron. Rechnen)* **1** (1966), 233 ff.
15. R. DUSSEL AND B. SCHMITT, Die Berechnung von Schranken für den Wertebereich eines Polynoms in einem Intervall, *Computing (Arch. Elektron. Rechnen)* **6** (1970), 35–60.
16. E. HANSEN, On the solution of linear algebraic equations with interval coefficients, *Linear Algebra and Appl.* **2** (1969), 153–165.
17. E. HANSEN, On solving systems of equations using interval arithmetic, *Math. Comp.* **22** (1968), 374–384.
18. H. HERMES, "Einführung in die Verbandstheorie," Springer–Verlag, Berlin/Heidelberg/New York, 1967.
19. A. S. HOUSEHOLDER, "The Theory of Matrices in Numerical Analysis," Blaisdell, New York/Toronto/London, 1964.
20. L. V. KANTOROVITCH, Funktionalanalysis und Angewandte Mathematik, *Uspehi Mat. Nauk. N.S.* **3** (1948), 89–185. [English transl., C. C. Benster, National Bureau of Standards].
21. R. KRAWCZYK, Fehlerabschätzung reeller Eigenwerte und Eigenvektoren von Matrizen, *Computing (Arch. Elektron. Rechnen)* **4** (1969), 281–293.
22. R. KRAWCZYK, NEWTON-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken, *Computing (Arch. Elektron. Rechnen)* **4** (1969), 187–201.
23. R. KRAWCZYK, Iterative Verbesserung von Schranken für Eigenwerte und Eigenvektoren reeller Matrizen, *Z. Angew. Math. Mech.* **48** (1968), 80–83.
24. R. KRAWCZYK, Iterationsverfahren zur Bestimmung komplexer Nullstellen, *Z. Angew. Math. Mech.* **50** (1970), 59–61.
25. R. KRAWCZYK, Einschließung von Nullstellen mit Hilfe einer Intervallarithmetik, *Computing (Arch. Elektron. Rechnen)* **5** (1970), 356–370.
26. R. KRAWCZYK, Verbesserung von Schranken für Eigenwerte und Eigenvektoren von Matrizen, *Computing (Arch. Elektron. Rechnen)* **5** (1970), 200–206.

27. R. Krawczyk, Gleichungen in halbgeordneten Räumen, *Abh. Math. Sem. Univ. Hamburg* **36** (1971), 150–165.
28. R. Krawczyk, Zur Einschließung inverser linearer Operatoren, *in* "Methoden und Verfahren der mathematischen Physik 7," pp. 1–19, Bibl. Inst. Mannheim, 1972.
29. K. Madsen, On the solution of nonlinear equations in interval arithmetic, *BIT* **13** (1973), 428–433.
30. O. Mayer, "Über die in der Intervallrechnung auftretenden Räume und einige Anwendungen," Dissertation, Karlsruhe, 1968.
31. O. Mayer, Algebraische und metrische Strukturen in der Intervallrechnung und einige Anwendungen, *Computing* (*Arch. Eelektron. Rechnen*) **5** (1970), 144–162.
32. R. E. Moore, "Interval Analysis," Prentice–Hall, Englewood Cliffs, N. J., 1966.
33. R. E. Moore, Two-sided approximation to solutions of nonlinear operator equations: A comparison of methods from classical analysis, functional analysis, and intervall-analysis, *in* "Lecture Notes in Computer Science No. 29," pp. 31–46, Springer–Verlag, New York/Berlin, 1975.
34. R. E. Moore, A test for existence of solutions to nonlinear systems, *SIAM J. Numer. Anal.* **14** (1977), 611–615.
35. K. Nickel, "Triplex-ALGOL and Applications," Interner Bericht des Instituts für Informatik der Universität Karlsruhe.
36. K. Nickel, Verbandstheoretische Grundlagen der Intervall-Mathematik, *in* "Lecture Notes in Computer Science No. 29," pp. 251–262, Springer–Verlag, New York/Berlin, 1975.
37. K. Nickel, Einige Bemerkungen über die Intervall-Newton-Methode, *in* "Vortrag auf der Tagung über Intervall-Mathematik, Oberwolfach, 1976."
38. H. H. Schäfer, "Topological Vektor Spaces," Macmillan Series in Advanced Mathematics and Theoretical Physics, Macmillan Co., New York, 1966.
39. G. Scheu, Über eine Wahl des Parameters beim Parallelenverfahren (parallel-chord-method), *Computing* (*Arch. Eelektron. Rechnen*) **20** (1978), 17–26.
40. J. Schröder, Über das NEWTONsche Verfahren, *Arch. Rational Mech. Anal.* **1** (1957), 154–180.
41. J. Schröder, Fehlerabschätzung bei linearen Gleichungssystemen mit dem Brouwerschen Fixpunktsatz, *Arch. Rational Mech. Anal.* **3** (1959), 28–44.
42. H. Unger, Nichtlineare Behandlung von Eigenwertaufgaben, *Z. Angew. Math. Mech.* **30** (1950), 281–282.
43. M. Urabe, A posteriori component-wise error estimation of approximate solutions to nonlinear equations, *in* "Lecture Notes in Computer Science No. 29," pp. 99–117, Springer–Verlag, New York/Berlin, 1975.
44. J. Weissinger, Zur Theorie und Anwendung des Iterationsverfahrens, *Math. Nachr.* **8** (1952), 193–212.